Advances in Intelligent Systems and Computing 998

Kohei Arai
Rahul Bhatia
Supriya Kapoor *Editors*

Intelligent Computing

Proceedings of the 2019 Computing Conference, Volume 2

Springer

Advances in Intelligent Systems and Computing

Volume 998

Series Editor

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Advisory Editors

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
Rafael Bello Perez, Faculty of Mathematics, Physics and Computing, Universidad Central de Las Villas, Santa Clara, Cuba
Emilio S. Corchado, University of Salamanca, Salamanca, Spain
Hani Hagras, School of Computer Science & Electronic Engineering, University of Essex, Colchester, UK
László T. Kóczy, Department of Automation, Széchenyi István University, Gyor, Hungary
Vladik Kreinovich, Department of Computer Science, University of Texas at El Paso, El Paso, TX, USA
Chin-Teng Lin, Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan
Jie Lu, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW, Australia
Patricia Melin, Graduate Program of Computer Science, Tijuana Institute of Technology, Tijuana, Mexico
Nadia Nedjah, Department of Electronics Engineering, University of Rio de Janeiro, Rio de Janeiro, Brazil
Ngoc Thanh Nguyen, Faculty of Computer Science and Management, Wrocław University of Technology, Wrocław, Poland
Jun Wang, Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong

fridaya@uef.fi

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink ****

More information about this series at http://www.springer.com/series/11156

fridaya@uef.fi

Kohei Arai · Rahul Bhatia ·
Supriya Kapoor
Editors

Intelligent Computing

Proceedings of the 2019 Computing Conference, Volume 2

Springer

fridaya@uef.fi

Editors
Kohei Arai
Faculty of Science and Engineering
Saga University
Saga, Japan

Rahul Bhatia
The Science and Information
(SAI) Organization
Bradford, West Yorkshire, UK

Supriya Kapoor
The Science and Information
(SAI) Organization
Bradford, West Yorkshire, UK

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-030-22867-5 ISBN 978-3-030-22868-2 (eBook)
https://doi.org/10.1007/978-3-030-22868-2

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

fridaya@uef.fi

Editor's Preface

On behalf of the Organizing Committee and Program Committee of the Computing Conference 2019, we would like to welcome you to the Computing Conference 2019 which will be held from July 16 to 17, 2019, in London, UK.

Despite the short history of computer science as a formal academic discipline, it has made a number of fundamental contributions to science and society—in fact, along with electronics, it is a founding science of the current epoch of human history called the Information Age and a driver of the Information Revolution. The goal of this conference is to give a platform to researchers with such fundamental contributions and to be a premier venue for industry practitioners to share new ideas and development experiences. It is one of the best respected conferences in the area of computer science.

Computing Conference 2019 began with an opening ceremony, and the conference program featured welcome speeches. It was a two-day conference, and each day started with keynote speeches from experts in the field. During the span of two days, a total of 18 paper presentation sessions and 4 poster presentation sessions were organized giving the opportunity to the authors to present their papers to an international audience. The conference attracted a total of 563 submissions from many academic pioneering researchers, scientists, industrial engineers, and students from all around the world. These submissions underwent a double-blind peer-review process. Of those 563 submissions, 170 submissions have been selected to be included in this proceedings. The published proceedings has been divided into two volumes covering a wide range of conference tracks, such as technology trends, computing, intelligent systems, machine vision, security, communication, electronics, and e-learning to name a few.

Deep appreciation goes to the keynote speakers for sharing their knowledge and expertise with us and to all the authors who have spent the time and effort to contribute significantly to this conference. We are also indebted to the Organizing Committee for their great efforts in ensuring the successful implementation of the conference. In particular, we would like to thank the Technical Committee for their constructive and enlightening reviews on the manuscripts in the limited timescale.

fridaya@uef.fi

We hope that all the participants and the interested readers benefit scientifically from this book and find it stimulating in the process. We are pleased to present the proceedings of this conference as its published record.

Hope to see you in 2020, in our next Computing Conference, with the same amplitude, focus, and determination.

Kohei Arai

fridaya@uef.fi

Contents

Check^2: A Framework for Fake Check-in Detection 1
Dina Abbas Abd El Moniem and Hoda M. O. Mokhtar

***n*-means: Adaptive Clustering Microaggregation of Categorical Medical Data** 13
Malik Imran-Daud

An Efficient Density-Based Clustering Algorithm Using Reverse Nearest Neighbour 29
Stiphen Chowdhury and Renato Cordeiro de Amorim

We Know What You Did Last Sonar: Inferring Preference in Music from Mobility Data 43
José C. Carrasco-Jiménez, Fernando M. Cucchietti, Artur Garcia-Saez, Guillermo Marin, and Luz Calvo

Characterization and Recognition of Proper Tagged Probe Interval Graphs 62
Sanchita Paul, Shamik Ghosh, Sourav Chakraborty, and Malay Sen

Evolvable Media Repositories: An Evolutionary System to Retrieve and Ever-Renovate Related Media Web Content 76
Marinos Koutsomichalis and Björn Gambäck

Ensemble of Multiple Classification Algorithms to Predict Stroke Dataset 93
Omesaad Rado, Muna Al Fanah, and Ebtesam Taktek

A Parallel Distributed Galois Lattice Approach for Data Mining Based on a CORBA Infrastructure 99
Abdelfettah Idri and Azedine Boulmakoul

A Community Discovery Method Based on Spectral Clustering and Its Possible Application in Audit 118
Hu Jibing, Ge Hongmei, and Yang Pengbo

fridaya@uef.fi

Refinement and Trust Modeling of Spatio-Temporal Big Data 132
Lei Zhang

Data Analytics: A Demographic and Socioeconomic Analysis of American Cigarette Smoking 145
Ah-Lian Kor, Reavis Mitch, and Sanela Lazarevski

Designing a System for Integration of Macroeconomic and Statistical Data Based on Ontology 157
Olga N. Korableva, Olga V. Kalimullina, and Viktoriya N. Mityakova

A Modified Cultural Algorithm for Feature Selection of Biomedical Data 166
Oluwabunmi Oloruntoba and Georgina Cosma

Exploring Scientists' Research Behaviors Based on LDA 178
Benji Li, Weiwei Gu, Yahan Lu, Chensheng Wu, and Qinghua Chen

Hybrid of Filters and Genetic Algorithm - Random Forests Based Wrapper Approach for Feature Selection and Prediction 190
Pakizah Saqib, Usman Qamar, Andleeb Aslam, and Aleena Ahmad

A New Arabic Dataset for Emotion Recognition 200
Amer J. Almahdawi and William J. Teahan

Utilise Higher Modulation Formats with Heterogeneous Mobile Networks Increases Wireless Channel Transmission 217
Heba Haboobi and Mohammad R. Kadhum

Performance Analysis of Square-Shaped and Star-Shaped Hierarchically Modulated Signals Generated Using All-Optical Techniques 230
Anisa Qasim and Salman Ghafoor

Dynamic Bit Loading with the OGFDM Waveform Maximises Bit-Rate of Future Mobile Communications 242
Mohammad R. Kadhum, Triantafyllos Kanakis, and Robin Crockett

The Two Separate Optical Fibres Approach in Computing with 3NLSE–Domain Optical Solitons 253
Anastasios G. Bakaoukas

Intra-channel Interference Avoidance with the OGFDM Boosts Channel Capacity of Future Wireless Mobile Communication 281
Mohammad R. Kadhum, Triantafyllos Kanakis, and Robin Crockett

Decomposition of a Non-stationary Process Using Time-Dependent Components 294
Abdullah I. Al-Shoshan

fridaya@uef.fi

Open Source Firmware Customization Problem: Experimentation and Challenges 306
Thomas Djotio Ndie and Karl Jonas

An Entity-Based Black-Box Specification Approach for Modeling Wireless Community Network Services 318
Thomas Djotio Ndié

Downlink High Quality 72 GHz Millimeter 336
Fawziya Al-Wahaibi and Hamed Al-Rwashidi

Study the Performance of Optical Millimetre Wave Based on Carrier Suppressed by Using an Inverted Optical Filter 345
Fawziya Al-Wahaibi, Rajagopal Nilavalan, and Hamed Al-Rwashidi

Visualization: A Tool to Study Nodes in Multi-dimensional Networks 358
Fozia Noor, Muhammad Usman Akram, Asadullah Shah, and Shoab Ahmad Khan

The Nigerian E-Waste Problem: Way Forward 368
Victor Ndako Adama, Ibrahim Shehi Shehu, Solomon Adelowo Adepojur, and A. Fatima Sulayman

High-Throughput Electronic Band Structure Calculations for Hexaborides 386
Zhenxi Pan, Yong Pan, Jun Jiang, and Liutao Zhao

Switched-Current Sampled and Hold Circuit with Digital Noise Cancellation Circuit for 2+2 MASH ΣΔ Modulator 396
Guo-Ming Sung, Leenendra Chowdary Gunnam, and Shan-Hao Sung

Towards the Identification of Context in 5G Infrastructures 406
Chrysostomos Symvoulidis, Ilias Tsoumas, and Dimosthenis Kyriazis

Researches on Time-Domain Relay Feedback Identification Approaches for High-Acceleration Linear Servo Systems 419
Xiaoli Shi, Yong Han, Jianhua Wu, and Chao Liu

Multimodal Biometrics Using Features from Face, Iris, and Conjunctival Vasculature 435
Noorjahan Khatoon and Mrinal Kanti Ghose

Evaluating Machine Learning Models on the Ethereum Blockchain for Android Malware Detection 446
Md. Shohel Rana, Charan Gudla, and Andrew H. Sung

Blockchain-Based Authentication for Network Infrastructure Security 462
Chi Ho Lau, Kai-Hau Yeung, and Kai Man Kwong

fridaya@uef.fi

Blockchain: Analysis of the New Technological Components as Opportunity to Solve the Trust Issues in Supply Chain Management 474
Adnan Imeri, Nazim Agoulmine, Christophe Feltus, and Djamel Khadraoui

Password Manager Combining Hashing Functions and Ternary PUFs 494
Bertrand Cambou

Capabilities of Email Forensic Tools 514
Ahmad Ghafarian

Random Sampling on an High-Dimensional Sphere for Solving SVP . . . 529
Ling Qin, Xue Zhang, Xiaoyun Wang, Zhongxiang Zheng, and Wenfeng Qi

Hybrid Dependencies Between Cyber and Physical Systems 550
Sandra König, Stefan Rass, Benjamin Rainer, and Stefan Schauer

Design and Implementation of a Lightweight Encryption Scheme for Wireless Sensor Nodes 566
Rutuja Salunke, Gaurav Bansod, and Praveen Naidu

Deep Random Based Key Exchange Protocol Resisting Unlimited MITM 582
Thibault de Valroger

Dynamic Password Protocol for User Authentication 597
H. Channabasava and S. Kanthimathi

Side-Channel Attacks on the Mobile Phones: Applicability and Improvements 612
Roman Mostovoy, Pavel Borisenko, Daria Sleptsova, Alla Levina, and Igor Zikratiov

Improving Learning Efficiency and Evaluation Fairness for Cyber Security Courses: A Case Study 622
Emin Çalışkan, Risto Vaarandi, and Birgy Lorenz

Multi-Confirmations and DNS Graph Mining for Malicious Domain Detection 639
Hau Tran, Chuong Dang, Hieu Nguyen, Phuong Vo, and Tu Vu

On Non-commutative Cryptography with Cubical Multivariate Maps of Predictable Density 654
V. Ustimenko and M. Klisowski

A Stealth Key Exchange Protocol 675
Khan Farhan Rafat

fridaya@uef.fi

Evaluation of the Multifactor Authentication Technique for Mobile Applications 696
Tianyi Zhang, Le Yang, and Yan Wu

A Doubly-Masked-Ballot Multi-questions Voting Scheme with Distributed Trust 708
Robert Szabo and Andras Telcs

Advantages of the PaySim Simulator for Improving Financial Fraud Controls 727
Edgar A. Lopez-Rojas and Camille Barneaud

Secure Cryptosystem Using Randomized Rail Fence Cipher for Mobile Devices 737
Amit Banerjee, Mahamudul Hasan, and Him Kafle

Network Protocol Security: A Linear Bitwise Protocol Transformation 751
Suzanna Schmeelk

Mesh-Based Encryption Technique Augmented with Effective Masking and Distortion Operations 771
Muhammed Jassem Al-Muhammed and Raed Abuzitar

Photos and Tags: A Method to Evaluate Privacy Behavior 797
Roba Darwish and Kambiz Ghazinour

The Analysis of the Socio-Technical Environment (STE) of Online Sextortion Using Case Studies and Published Reports from a Cybersecurity Perspective 817
Alex Cadzow

Lightweight Datapath Implementation of ANU Cipher for Resource-Constrained Environments 834
Vijay Dahiphale, Gaurav Bansod, and Ankur Zambare

An Evolutionary Mutation Testing System for Java Programs: eMuJava 847
Muhammad Bilal Bashir and Aamer Nadeem

"STEPS" to STEM 866
Esther Pearson

Named Entity Enrichment Based on Subject-Object Anaphora Resolution 873
Mary Ting, Rabiah Abdul Kadir, Azreen Azman, Tengku Mohd Tengku Sembok, and Fatimah Ahmad

fridaya@uef.fi

Tour de Force: A Software Process Model for Academics 885
Zeeshan Haider Malik, Habiba Farzand, Muhammad Ahmad, and Awais Ashraf

Virtual Testbeds with Meta-data Propagation 902
Thomas Kuhn, Pablo Oliveira Antonino, and Andreas Morgenstern

Challenges and Mitigation Strategies in Reusing Requirements in Large-Scale Distributed Agile Software Development: A Survey Result 920
Syeda Sumbul Hossain

Harvesting and Informal Representation of Software Process Domain Knowledge 936
R. O. Oveh and F. A. Egbokhare

A Methodology for Performing Meta-analyses of Developers Attitudes Towards Programming Practices 948
Thomas Butler

Modelling and Analysis of Partially Stochastic Time Petri Nets Using Uppaal Model Checkers 975
Christian Nigro, Libero Nigro, and Paolo F. Sciammarella

Strengthening Virtual Learning Environments by Incorporating Modern Technologies 994
Kashif Laeeq and Zulfiqar Ali Memon

Knowledge Management System: Designing a Virtual Community of Practice for Higher Education 1009
Gabriela Ziegler

The Influence of Online and Traditional Computer Laboratories on Academic Performance of Students 1030
Ahmed O. A. Ismail, Ahmad K. Mahmood, Ammar E. Babiker, and Abdelzahir Abdelmaboud

Smart, Social, Flexible and Fun: Escaping the Flatlands of Virtual Learning Environments 1047
Mike Brayshaw, Neil A. Gordon, and Simon Grey

Smart Mobile Learning Environment for Programming Education in Nigeria: Adaptivity and Context-Aware Features 1061
Friday Joseph Agbo and Solomon Sunday Oyelere

An Integrated Approach for Educating the Marine Forces Reserve Inspector-Instructor Cadre 1078
Alejandro S. Hernandez and Lisa R. Spence

fridaya@uef.fi

Application of Web 2.0 Tools in the Pedagogical Process of Mathematics: A Case Study with Students of Basic Education 1097
Diego Avila-Pesantez, Mónica Vaca-Cárdenas, L. Miriam Avila P., and Leticia Vaca-Cárdenas

Designing an Educational Multimedia System for Supporting Learning Difficulties in Arabic 1117
Moutaz Saleh and Jihad M. Alja'am

Straight Line Detection Through Sub-pixel Hough Transform 1129
Guillermo J. Bergues, Clemar Schürrer, and Nancy Brambilla

Microcontroller-Based Speed Control Using Sliding Mode Control in Synchronize Reluctance Motor 1138
Wei-Lung Mao, Gia-Rong Liu, Chao-Ting Chu, and Chung-Wen Hung

A Breath Switch for Secure Password Input 1150
Natsuki Sayama, Naoka Komatsu, Kana Sawanobori, and Manabu Okamoto

Don't Sweat It: Mobile Instruments for Clinical Diagnosis 1155
Hellema Ibrahim and Perry Xiao

Technical Analysis Toolkit for Neural Networks in Finance and Investing 1170
Chun-Yu Liu, Shu-Nung Yao, and Ying-Jen Chen

Computer Graphics Based Analysis of Loading Patterns in the Anterior Cruciate Ligament of the Human Knee 1175
Ahmed Imran

Gesture Recognition Using an EEG Sensor and an ANN Classifier for Control of a Robotic Manipulator 1181
Rocio Alba-Flores, Fernando Rios, Stephanie Triplett, and Antonio Casas

An Eye Tracking Study: What Influences Our Visual Attention on Screens? 1187
Arwa Mashat

Sports Vision Based Tennis Player Training 1193
Kohei Arai

No-Reference Image Quality Assessment Based on Multi-scale Convolutional Neural Networks 1202
Peikun Chen, Yuzhen Niu, and Dong Huang

Motion Detection with IoT-Based Home Security System 1217
Pei Kee Tiong, Nur Syazreen Ahmad, and Patrick Goh

fridaya@uef.fi

A Simple Parameter Estimation Method for Permanent Synchronous Magnet Motors 1230
Dong-Myung Lee

A Bug-Inspired Algorithm for Obstacle Avoidance of a Nonholonomic Wheeled Mobile Robot with Constraints 1235
Sing Yee Ng and Nur Syazreen Ahmad

Obstacle Avoidance Strategy for Wheeled Mobile Robots with a Simplified Artificial Potential Field 1247
Sing Yee Ng and Nur Syazreen Ahmad

Immersive Analytics Through HoloSENAI MOTOR Mixed Reality App 1259
André L. M. Ramos, Thiago Korb, and Alexandra Okada

Explainable Artificial Intelligence Applications in NLP, Biomedical, and Malware Classification: A Literature Review 1269
Sherin Mary Mathews

Author Index 1293

fridaya@uef.fi

Check^2: A Framework for Fake Check-in Detection

Dina Abbas Abd El Moniem(✉) and Hoda M. O. Mokhtar

Information Systems Department, Faculty of Computers and Information, Cairo University, Giza, Egypt
{d.abbas,h.mokhtar}@fci-cu.edu.eg

Abstract. Location-based social networks (LBSNs) have increased rapidly over the past several years due to the proliferation of mobile devices. As people continue to use their phones everywhere and all the time - even while they shop and dine, a lot of LBSN services such as Foursquare (https://foursquare.com/), Facebook place (https://www.facebook.com/places/) have been introduced. These LBSNs provide new types of services that are location based for social network users. LBSNs integrate the functionalities of both location services and social networks. Consequently, this empowers people to use location data with online social networks in several ways including: location recommendation, mobile-advertising, etc. Mainly, LBSN is widely used in sharing location between users. In addition, LBSNs make use of the notion of "check-in" to enable users to share their location with other users. To attract more users, LBSN service providers offer rewards to users who check-in at a certain place, however, this caused some users (i.e. attackers) to cheat regarding their true locations in order to get either monetary or virtual rewards. Nevertheless, fake check-ins can cause monetary losses to service providers. In this paper, we propose a novel framework for analyzing users' check-ins and detecting fake check-ins.

Keywords: Location-based social networks · Location-based social services · Fake check-in · Detection · Monetary cheaters · Game cheaters · Venue · Check-in · Global Positioning System (GPS) · Reward · Detection

1 Introduction

Location dimension helps bridging the gap between the physical world and the digital online social networking services. Location based social networks (LBSNs) utilize location information from mobile phones obtained through Global Positioning Systems (GPSs), enabling users to interact in "real time" with the places and people around them. These networks use mobile applications to connect with people while they are out, and currently, many of these networks operate with a "check-in" feature, where the user logs into the application and broadcasts whatever location they have just arrived at or they claim to visit. While being checked in, users can usually leave tips, write reviews, upload photos, or interact with other users. Other networks pull users' data continuously and place their users' location on the map according to their traveling GPS location.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1–12, 2019.
https://doi.org/10.1007/978-3-030-22868-2_1

fridaya@uef.fi

A lot of LBSN applications were recently introduced such as Foursquare[1], Facebook place[2], Gowalla[3], Brightkite[4], Loopt[5]. These applications [1] offer users monetary and virtual rewards. For example, a user who checks in Starbucks for 3 times a month, will be rewarded a free large drink. Users can also explore restaurants, cafes and stores with discounts/coupons [2]. They can also share their experience(s) with their friends. In addition, users can rate the places they visit and share those ratings with friends for future recommendation. In the remainder of the paper we will use word "check-in" to refer to an event when the time and the location of a particular user is recorded. For location-based social networks, this means that a user checked-in to a specific location using the online social network website/application.

Despite the fact that LBSNs attract users for the above reasons, business owners are attracted as well to LBSNs to promote their businesses. From a business owner perspective, to have your business benefit from LBSNs, you have to create a venue in LBSN application with all the information needed about your business. The basic information only will be public to users as venue name, location, rating, and the offers offered by the venue. In addition, business owners can collect statistics about their users' behavior as well as attract new users to their venues through providing coupons/discounts [3].

In general, the key entities in LBSNs are the users and the venues. Both entities interact together in different ways, a user can simply check-in in a venue and rate it; and on the other hand, a venue provides a service to the user with a certain quality. Nevertheless, trust is a crucial feature for the success of LBSNs. When users enter fake check-ins for venues they didn't actually visit, this can easily mislead the business owner making him think that his venue is highly visited or that a certain "fake" user deserves rewards. In addition, a user with a fake check-in can also introduce forged rating for a venue that can eventually affect the venues reputation and even result in false recommendations.

Detecting and identifying false check-ins is a critical point to guarantee the success of LBSNs [4]. Ensuring trustfulness affects the whole LBSN system resulting in: accurate recommendations, rewards to people who deserve them, and real business revenue increase, etc.

In this paper we propose a framework for detecting fake check-ins. The main contributions of this paper are as follows:

- Presenting different types of cheatings and how cheaters use LBSNs to introduce fake check-ins.
- Proposing a framework for fake check-in detection (CHECK^2). This framework integrates different approaches to effectively detect fake check-ins and cheaters. The proposed framework employs data mining techniques rather than relying on GPS related systems.

[1] https://foursquare.com/.

[2] https://www.facebook.com/places/.

[3] http://blog.gowalla.com/.

[4] https://brightkite.com/.

[5] http://www.looptmix.com/.

fridaya@uef.fi

The rest of the paper is organized as follows: In Sect. 2, we provide an overview of location-based social networks. In Sect. 3, we discuss the related work. In Sect. 4, the proposed framework is introduced for fake check-in detection. In Sect. 5, our Experiments. In Sect. 6, we conclude our work and propose directions for future work.

2 Overview of LBSNs

In this section we briefly discuss two main points, namely, location based social networks and its main components, and types of cheaters occurring in location based social networks.

First, Location-Based Social Networks (LBSNs) do not only mean adding a location to an existing social network so that people in the social structure can share location-embedded information, but rather consists of the new social structure made up of individuals connected by the interdependency derived from their locations in the physical world as well as their location-tagged media content, such as photos, video, and texts [5]. Here, the physical location consists of the instant location of an individual at a given timestamp and the location history that an individual had accumulated over a certain time period. Furthermore, the interdependency includes not only that two persons co-occurring in the same physical location or sharing similar location histories, but also the knowledge, e.g., common interests, behavior, and activities, inferred from an individual's location (history) and location-tagged data [6]. On the other hand, a Social Network is a social structure made up of individuals connected by one or more specific types of interdependency, such as friendship, common interests, and shared knowledge. Generally, a social networking service builds on and reflects the real-life social networks among people through online platforms such as a website, providing ways for users to share ideas, activities, events, and interests over the Internet [6, 7]. In location based social networks, a location is usually represented either in absolute representation (i.e. latitude-longitude coordinates), or relative (e.g. 100 meters North of the Space Needle), or symbolic (e.g. home, office, or shopping mall). Alongside, a location usually has three kinds of geospatial representations: (1) a point location, (2) a region, and (3) a trajectory [6, 7].

The second point we discuss in this section is the types of cheaters and their characteristics [8] Cheaters are divided into two major cheating types, namely: (i) monetary cheaters and (ii) gamer cheaters. In the following discussion we elaborate those types in more details.

Monetary Cheaters. Those are cheaters who are attracted by venues that offer special deals. Monetary cheaters are attracted to real world rewards. He et al. [9] have reported that the majority of the special offers (more than 90%) in Foursquare - the major LBSN service to date - require multiple check-ins (e.g., X times) to the venue. Hence, cheaters have to generate a number of fake check-ins in order to obtain the offer faster and with less cost. Clearly, monetary cheaters can consequently lead to revenue losses [10].

Gamer Cheaters. Those are cheaters who are attracted by venues that can facilitate their goal for as many virtual rewards as possible. In other words, they do not care for the specific characteristics of the venues as long as the latter satisfy their goals. A large

fridaya@uef.fi

fraction of users view these virtual rewards as means to prove their social status (e.g., more mayor ships translate to more outgoing, social personality etc.), hence, form an important reason for them to continue using the system. For instance, in Foursquare a user is able to earn points for every check-in, badges for a specific series of check-ins, "mayor ships" of venues once he has the most check-ins in the venue within the last two months etc.

In the rest of this paper - we will elaborate how we can use those characteristics of cheaters to identify them and eventually prevent and check for fake "check-ins".

3 Related Work

Due to the importance of location based social networks and more specifically, their vital role in location based recommendation [11–13] several research work targeted this topic from different perspectives. In the following discussion we highlight some of the main research efforts that were done in this domain and are related to the work presented in this paper. Zhang [14] proposed a "Honeypot Scheme" for detecting check-ins in location-based services. This scheme is mainly designed as a filtering mechanism, flagging users with high or low levels of suspicious behavior. The main focus of the paper is gamer cheater as well as monetary cheaters. The main idea is as follows: LBSN service provider can create fake venues – namely, the honeypots - that appear attractive to gamer cheaters (e.g., for the case of Foursquare a possible honeypot venue is one that appears to be easy to obtain its "mayor ship"). Given that under honest use of the system no one should be present in that locale, users that check-in in honeypot venues are automatically flagged as (potential) gamer cheaters. However, the authors did not perform evaluations of the proposed schemes. Thorough evaluations would require the creation of a large number of honeypot venues to measure the efficiency of the proposed approach. Yu et al. [15] presented "SybilGuard", a novel protocol for limiting the corruptive influences of sybil attacks. In a Sybil attack, a malicious user obtains multiple fake identities and pretends to be multiple, distinct nodes in the system. By controlling a large fraction of the nodes in the system, the malicious user is able to "out vote" the honest users in collaborative tasks such as Byzantine failure defenses. The protocol is based on the "social network" among user identities, where an edge between two identities indicates a human-established trust relationship. Malicious users can create many identities but few trust relationships. Thus, there is a disproportionately-small "cut" in the graph between the Sybil nodes and the honest nodes. Iasonas et al. [16] created a platform for testing the feasibility of fake-location attacks. Their experimental results validate that detection-based mechanisms are not effective against fake check-ins, and new directions should be taken for designing countermeasures. Hence, they implemented a system that employs near field communication (NFC) hardware and a check-in protocol that is based on delegation and asymmetric cryptography, to eliminate fake-location attacks. Authors in [17–19] presented Sybil defender, a mechanism that identifies sybil nodes and detect the community around a Sybil node. In OSNs, a Sybil is a fake account with which a user attempts to create multiple identities to make as many friends as possible with legitimate accounts. A Sybil account can lead to many malicious activities in an OSN.

fridaya@uef.fi

For example, a Sybil can control some accounts by using the fake identities to provide misleading information. Moreover, false reputation can be created based on Sybil accounts.

Inspired by the previous attempts, in the following discussion we present a novel framework for face check-in detection that uses some features from previous approaches along with employing cheaters characteristics to enhance fake check-in detection.

4 Proposed Framework

In this section, a fake check-in detection and checking framework is proposed (check^2) that considers both historical check-ins and new check-ins. The proposed framework integrates different techniques to detect fake check-ins, among those techniques is unsupervised data mining techniques to classify users and detect fake ones. One of the key challenges in this work is the lack of real datasets for users with check-ins information. The main reason behind this is to preserve users' privacy. Hence, to overcome this problem, in this work we use unlabeled dataset from "Weeplaces" [20]. The proposed framework proceeds in the following phases as shown in Fig. 1.

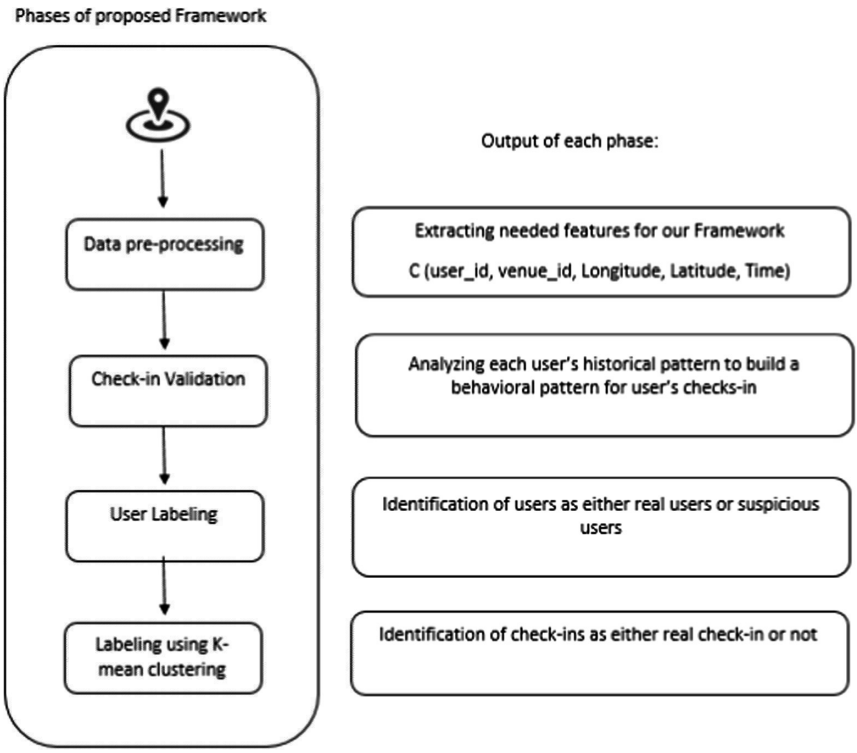


Fig. 1. "Check 2" - the proposed framework

fridaya@uef.fi

Phase 1: Data Pre-processing
In this phase we pre-process the data and design the database schema that we will use for the rest of the work. Relational database model is used for data modeling. Data pre-processing is an important and critical step in the data mining process and it has a huge impact on the success of a data mining project. Through this step, the nature of the data is better understood and the data analysis is performed more accurately and efficiently. For each check-in, we perform data preprocessing techniques to make sure that the data are correct and can be injected into the framework. There are a number of different tools and methods used for pre-processing, including: sampling, which selects a representative subset from a large population of data; transformation, which manipulates raw data to produce a single input; de-noising, which removes noise from data; normalization, which organizes data for more efficient access; and feature extraction, which pulls out specified data that is significant in some particular context.

Our database schema is divided into three entities as shown below:

User (UserId, UserName, UserLocation)
Venue (VenueId, VenueName, VenueLocation)
Check-in (UserId, VenueId, Date_Time, Longitude, Latitude, Category)

Table 1 below shows a sample of our dataset for different users.

Table 1. Sample of our dataset

UserId	VenueId	Date_Time	Longitude	Latitude	Category
ben-parr	muir-woods-national-monument-mill-valley	2017-07-17 19:48:28	37.89256	−122.572	Parks & Outdoors: Hiking Trail
bijan-sabet	spruce-san-francisco	2017-03-26 02:29:22	37.7876	−122.453	Food: American
bijan-sabet	palace-of-fine-arts-san-francisco	2017-04-14 03:02:34	37.80743	−122.444	Parks & Outdoors: Harbor/Marina
bijan-sabet	st-francis-yacht-club-san-francisco	2017-04-15 01:10:36	37.80043	−122.441	Food: Breakfast/Brunch
eric-wu	nijo-sushi-seattle	2017-10-23 04:36:53	47.60506	−122.338	Food: Sushi
eric-wu	trinity-seattle	2017-10-23 06:01:04	47.6018	−122.334	Nightlife: Nightclub/Discotheque

Below is a formal annotation to be used in our framework:

U_i: A unique User identifier (UserId) for the user who performs a check-in
V_i: A unique venue identifier (VenueId) that a user $U_i$checked-in
Lat: Geographical latitude for Vi
Lon: Geographical longitude for Vi
D: Distance between coordinates of check-in1 and check-in 2 for the same user Ui

fridaya@uef.fi

T_{exp}: Expected time of the performed check-in
T_{act}: Actual Time of the performed check-in

Phase 2: Check-in Validation
In this phase historical check-ins for each user are first analyzed. The aim is to understand the behavior of the user (Ui) through his previous check-ins. This facilitates the identification process for new check-ins. For each user, we create a profile for the user. Then, start analyzing each check-in. For each check-in for user i, we extract all the needed fields (user id, venue id, longitude, latitude, Time). Then we start taking every 2 successive check-ins and perform some calculations on each pair of check-ins. Those calculation are performed to check the validity and feasibility of different check-ins. The following discussion elaborates the checks performed:

1. ***Time***: (i.e. Time allowed for a user to perform the second check-in)

It is important to set a baseline for the allowed time between 2 consecutive check-ins for the same user. This will reduce successive check-ins done by malicious users. So, in the proposed framework we set a 30 min as a time factor between a check-in and the next one. Nevertheless, this could be infeasible or non-realistic for all cases; as a user can check-in in 2 different venues near each other within 30 min. Hence, we introduce a new measure and calculate the expected and actual time between 2 successive check-ins.

2. ***Expected Time between 2 check-ins***

Expected time between 2 check-ins is calculated according to Greenwich Mean Time [21] as in Eq. (1). Assuming that the speed for a user from a venue to another riding a vehicle is 60 km/hr (average speed), and the speed for a user walking is 10 km/hr. Equation (1) helps us to expect the amount of time needed from one venue to another.

The time standard for celestial navigation is Greenwich Mean Time, GMT. GMT is based upon the GHA of the mean sun (an imaginary sun which move at a constant speed):

$$GMT[h] = \frac{GHA_{\,Mean\,Sum}\,[^\circ]}{15} + 12$$

(If GMT is greater than 24 h, subtract 24 hours)

Equation 1: Greenwich Mean Time

In other words, GMT is the angle, expressed in hours, between the lower branch of the Greenwich meridian and the hour circle through the mean sun. The GHA of the mean sun increases by exactly 15° per hour, completing a 360° cycle in 24 h. Celestial coordinates tabulated in the Nautical Almanac refer to GMT.

3. ***Actual Time between 2 check-ins***

Actual time is calculated by computing the difference between the time of check-in (i + 1) and the previous check-in (i). Then, we compare between expected time and actual time. If the difference is in the acceptable range, we mark check-in as a valid one. Else, we mark the check-in as a suspicious one.

fridaya@uef.fi

4. ***Actual Distance between 2 check-ins***

For each check-in, we transformed longitude & latitude to distance in kilometers using Great Circle Distance [22, 23] equation as stated in Eq. (2). The Great Circle Distance method with two equations, Vincenty and Haversine, is used to calculate the accurate positioning.

```
ACOS(COS(RADIANS(90-Lat1))*COS(RADIANS(90-Lat2))
+SIN(RADIANS(90-Lat1))        *SIN(RADIANS(90-Lat2))
*COS(RADIANS(Long1-Long2))) *6371
```

Equation 2: Great Circle Distance

5. ***Actual Co-ordinates for a venue***

For each venue V, there exists the actual coordinates for this venue, we defined '∂' as an acceptable range for a user to check-in this venue. '∂' can be in the range of 500 meters away from the actual defined coordinates for the venue.

6. ***Number of check-ins per day***

Number of check-ins performed by a user is one of the important factors to be taken into consideration. Malicious users always perform a lot of check-ins per day to gain rewards from venue owners. Therefore, we set a basis for no. of check-ins per a day as follows:

a. 0–5 check-ins is considered normal,
b. 5–15 check-ins is considered suspicious,
c. >15 check-ins is considered fake.

Phase 3: User Labelling

In this phase, we are concerned by studying the user behavior. The historical pattern of a user's check-in behavior has two properties in LBSNs. Firstly, a user's check-in history approximately follows a power-law distribution [5], i.e., a user goes to a few places many times and to many places a few times. Secondly, historical pattern has short-term effect. i.e., a user arrives at the airport and then takes a shuttle to the hotel. After his dinner, he drinks a cup of coffee. The historical pattern of the previous check-ins at the airport, shuttlestop, hotel and restaurant have different strengths with respect to the latest check-in at the coffee shop. Furthermore, historical tie strength decreases over time.

In location-based social networking sites, users explore various POIs and checkin at places that interest them. The power-law property of users' check-in behavior indicates that users do visit new places, resulting in the "cold-start" check-in problem. Predicting the "cold-start" check-in locations (i.e., predicting a user's next location where he has never been before) exacerbates the already difficult problem of location prediction as there is no historical information on the user for the new place, hence, traditional prediction models relying on the observation of historical check-ins would fail to predict the "cold-start" check-ins. In this scenario, social network information could be utilized to help address the "cold-start" problem since social theories (e.g., social correlation [1]) suggest that human movement is usually affected by their social networks.

fridaya@uef.fi

In our framework, we start building profiles for all current users. For each user, a detailed behavioral analysis is conducted. This detailed profile includes check-ins for a user daily, weekly and monthly. Mainly, we are concerned in our framework by 3 attributes as follows; Number of check-ins per day, type of check-ins and time frame for a check-in. Then, we start analyzing the user behavior and detect any suspicious actions done by any user.

Consider for example the behavior of the 2 users, a normal and a suspicious one: User1 has around 3 to 5 check-ins daily. The analysis shows that he starts his day by going to a coffeeshop to get his coffee, then he goes to his work. In most of the afternoons, he goes to the gym and then hangout with his friends. So, this is the normal behavior for User1. The other scenario is that User2 starts doing several check-ins within small time frames. Also, he is not following the normal behavior he is used to follow. In addition, the number of check-ins exceeds five. He checks-in in his work place then after a small-time stamp, he checks-in a restaurant that is far from work, then he repeats the same check-in pattern after a while, the following check-in is in a coffeeshop that is also far from the previous check-in. This will be detected by the proposed system as suspicious. Suspicious users are reviewed for the following successive days to check their check-ins pattern. Taking into consideration the number of check-ins, time stamps, geographical locations of venues and distance between check-ins, a user is then identified as either a real user or suspicious user that may end up being detected as a cheater.

Phase 4: Check-in Labelling Using K-means Clustering
After building a detailed profile for each user with a detailed information about a user's check-in. The objective of this phase is to process each check-in to identify if there is any suspicious behavior about it. The output of phase 2 eases the identification process. For each check-in, we build a vector of 3 parameters named as follows (Time, Location, & Type of check-in). Our first parameter is Time; as mentioned above, we start calculating the difference between actual time and expected time for a check-in to be performed. If Actual Time for a check-in > Expected time for a check-in by α, where α is defined to be 1800 s all day except in rush hours, it can extend to be 5400 s. The second parameter is Location coordinates. We start comparing the coordinates of the check-in performed with the actual coordinates for a venue that is defined in our data. Our framework accepts a small difference between the 2 coordinates. We defined this difference as β. β can be up to 0.5 km from the actual coordinates. The third parameter is a combination between time of check-in and type of the performed check-in. As discussed above, we built profiles for each user behavior so now we can detect if any new behavior is performed by the user.

The output of these parameters is then injected into a k-means clustering algorithm. The objective of the clustering process is to identify the check-ins either as real ones or suspicious ones. The suspicious ones are then analyzed more to be labelled as suspicious check-ins or not.

5 Experiments

LBSN is an interesting research area however more research needs to be conducted in this field. So we implemented Honeypot scheme that was stated in related work section to be able to compare our work to others. We injected a honeypot venue with specific

fridaya@uef.fi

geographical coordinates in our dataset. We made honeypot venue an attractive one to attract cheaters. This venue offers a lot of rewards for users who checks-in it. So, whenever a user checks-in this venue, he/she is marked as a fake user. However, we permit a user to check-in a honeypot venue x-times as we assume that this can be done by mistake. So, a user who checks-in a honeypot is put under observation to determine his state as a cheater or not. [Experiment 1]

All our experiments are shown in Table 2.

Table 2. Experiments done on our dataset

Experiment 1	Honeypot venue
Experiment 2	Our framework + honeypot venue
Experiment 3	Our framework without K means

We also added the honeypot implemented part to our framework and run our work with this part integrated with the other factors stated above. [Experiment 2]

5.1 Datasets

This dataset is collected from Weeplaces, a website that aims to visualize users' check-in activities in location-based social networks (LBSN). It is now integrated with the APIs of other location-based social networking services, e.g., Facebook Places, Foursquare, and Gowalla. Users can login Weeplaces using their LBSN accounts and connect with their friends in the same LBSN who have also used this application. All the crawled data is originally generated in Foursquare. This dataset contains 7,658,368 check-ins generated by 15,799 users over 971,309 locations. In the data collection, we can't get the original Foursquare IDs of the Weeplaces users. We can only get their check-in history, their friends who also use Weeplaces, and other additional information about the locations.

5.2 Results and Discussions

The performance of the proposed method is evaluated and compared to Honeypot method, with and without k-means clustering. The best result was achieved with the whole framework including all heuristics, k-means and honeypot. The output of our framework was more accurate than other techniques as shown in Table 3. These results are represented graphically in Fig. 2. However we need to run our experiments on labeled data to be able to determine our accuracy level on labeled data.

Table 3. Expected results vs. actual results

	Expected results	Actual results
# of valid check-ins	1915	1712
# of suspicious check-ins	9580	9783

fridaya@uef.fi

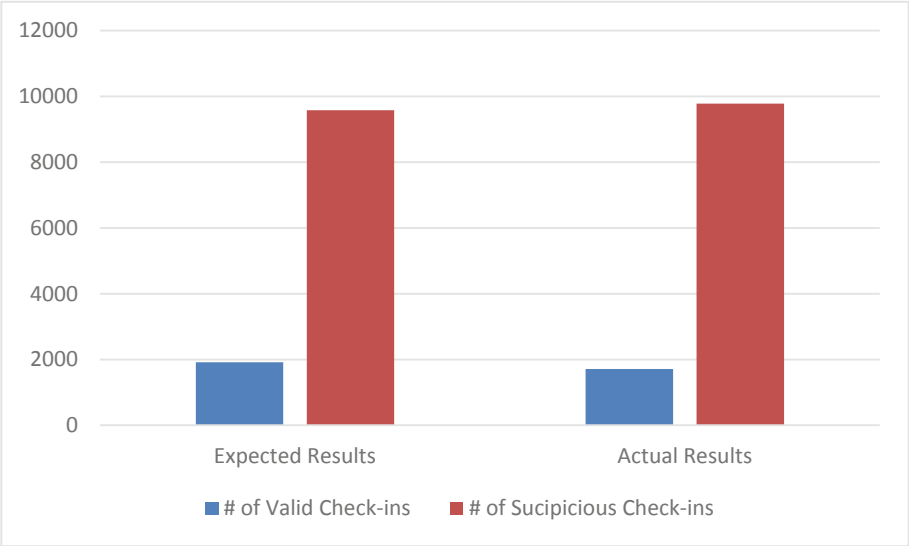


Fig. 2. Graph representation for both results

6 Conclusion and Future Work

Applications that facilitate location information to provide a number of novel services have emerged during the last few years. However, these applications have mainly focused on providing users with an easy way to generate huge volumes of data, more specifically, through the action of check-in without checking the trustfulness of the system. This has left the floor open for misbehaving users to game around the system and even overwhelm it with fake geographical information. In this paper, we propose a framework for detecting fake check-ins in location-based services. The proposed framework is based on analyzing the misbehaving users and integrating different techniques to detect fake check-ins. The proposed framework also classifies users as cheaters or not. This feature helps in tracking future users' behaviors and consequently in determining future cheating actions. This is for sure an important step before recommendation places to users. For future work, the proposed framework using other machine learning techniques which is still open for further investigation will be experimented.

References

1. Rachuri, K.K., Hossmann, T., Mascolo, C., Holden, S.: Beyond location check-ins: exploring physical and soft sensing to augment social check-in apps. In: 2015 IEEE International Conference on Pervasive Computing and Communications, PerCom 2015, pp. 123–130 (2015)
2. Yu, Z., Yang, Y., Zhou, X., Zheng, Y., Xie, X.: Investigating how user's activities in both virtual and physical world impact each other leveraging LBSN data. Int. J. Distrib. Sens. Networks **10**, 461780 (2014)

fridaya@uef.fi

3. Ester, M., Kriegel, H., Xu, X., Miinchen, D.: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise (1996)
4. He, W.: Location Cheating: A Security Challenge to Location-based Social Network Services (2011)
5. Gao, H., Liu, H.: Data analysis on location-based social networks. In: Mobile Social Networking, pp. 165–194 (2013)
6. By, E., Zheng, Y., Zhou, X.: Computing with Spatial Trajectories (2011)
7. Zheng, Y.: Tutorial on Location-Based Social Networks. In: Proceeding of the 21st International conference on World Wide Web, no. 5 (2012)
8. Rossi, L.: It's the Way you Check-in: Identifying Users in Location-Based Social Networks Categories and Subject Descriptors (2011)
9. He, W., Liu, X., Ren, M.: Location Cheating: A Security Challenge to Location-based Social Network Services (2011)
10. Adikari, S., Systems, I.: Identifying fake profiles in Linkedin. In: Pacific Asia Conference on Information Systems, pp. 1–15 (2014)
11. Liu, B.: Point-of-interest recommendation in location based social networks with topic and location awareness. In: SDM, pp. 396–404 (2013)
12. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. **17**(6), 734–749 (2005)
13. Wang, H., Terrovitis, M., Mamoulis, N.: Location Recommendation in Location-based Social Networks using User Check-in Data Categories and Subject Descriptors. In: Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - SIGSPATIAL'13 (2013)
14. Pelechrinis, K., Krishnamurthy, P., Zhang, K.: Gaming the game: honeypot venues against cheaters in location-based social networks, p. 3. CoRR arXiv:1210 (2012)
15. Kaminsky, M., Gibbons, P.B., Flaxman, A.: SybilGuard: Defending Against Sybil Attacks via Social Networks. In: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications, vol. 16, no. 3, pp. 267–278, (2006)
16. Polakis, I., Volanis, S., Athanasopoulos, E., Markatos, E.P.: The Man Who Was There: Validating Check-ins in Location-Based Services. In: Proceedings of the 29th Annual Computer Security Applications Conference on - ACSAC '13, pp. 19–28 (2013)
17. Wei, W., Xu, F., Tan, C.C., Li, Q., Member, S.: SybilDefender: a defense mechanism for sybil attacks in large social networks. IEEE Trans. Parallel Distrib. Syst. **24**(12), 2492–2502 (2013)
18. Al-Qurishi, M., Member, S.: Sybil defense techniques in online social networks: a survey. IEEE Access **5**, 1200–1219 (2017)
19. Chang, W., Wu, J.: A survey of sybil attacks in networks. In: Publications of computer and Information Sciences, vol. 98. Temple University, Philadelphia (2013)
20. Datasets: http://www.yongliu.org/datasets/. Accessed 15 Aug 2018
21. Xiao, L., Hung, E.: An efficient distance calculation method for uncertain objects. In: Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2007, no. i, pp. 10–17 (2007)
22. Kifana, B.D.: Great circle distance methode for improving operational control system based on GPS tracking system. Int. J. Comput. Sci. Eng. **5**, 1200–1219 (2012)
23. Bullock, R.: Great Circle Distances and Bearings Between Two Locations, pp. 3–5 (2007)

fridaya@uef.fi

n-means: Adaptive Clustering Microaggregation of Categorical Medical Data

Malik Imran-Daud(✉)

Department of Software Engineering,
Foundation University Rawalpindi Campus, Islamabad, Pakistan
malikimran.daud@urv.cat

Abstract. Huge amount of information is managed and shared publically by the individuals and data controllers. Publically shared data contains information that can reveal identity of users, thus affecting privacy of individuals. To palliate these disclosure risks, Statistical Disclosure Control (SDC) methods are applied to the data before it is released. Microaggregation is one of the SDC methods that aggregate similar records into clusters, and then transform them into *m* indistinguishable records. K-means is a famous data mining clustering algorithm for continuous data, which iteratively maps similar elements into k-cluster until they all converge. However, adapting k-means algorithm for categorical multivariate is a challenging task due to high dimensionality of attributes. In this paper, we extend k-means clustering algorithm to achieve notion of microaggregation of structured data. Moreover, to preserve data utility, we extend fixed clustering nature of this algorithm to adaptive size clusters. For this purpose, we introduce *n*-means clustering approach that construct clusters based on the semantics of the datasets. In experiments, we proved significance of our proposed system by measuring cohesion of clusters and information loss for utility purpose.

Keywords: Anonymity · Microaggregation · K-means · Privacy

1 Introduction

Individuals and organizations use internet as a platform to manage their data. Such data is managed and shared publically in unstructured or structured format, for example: query logs, medical records, documents, social network posts, etc. Release of such data implies some serious implications from privacy perspective [1]. However, privacy can be preserved by anonymizing data prior to its release.

Microaggregation [2] is one of the techniques of Statistical Disclosure Control (SDC) that achieves notion of k-anonymity [3] to ensure privacy of users (i.e., it enables *m* indistinguishable records within a cluster to ensure anonymization of records). Initially, this method was proposed for continuous data, later it was adapted for categorical data by many researchers [1, 4–6]. In this method, the dataset is partitioned in clusters based on their mutual properties, and then records of such clusters are anonymized with an aggregate value determined from the cluster (usually with the centroid of the cluster). However, constructing clusters for categorical data is a challenging task to perform due

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 13–28, 2019.
https://doi.org/10.1007/978-3-030-22868-2_2

fridaya@uef.fi

to the high dimensionality of attributes [7]. Moreover, data generated as a result of microaggregation methods must retain data utility, which is used for analytical purposes by the researchers.

K-means is a well-known distance-based clustering algorithm designed for continuous datasets that maps n records into k clusters. Researchers [8–10] extended this algorithm to manipulate categorical datasets by considering dissimilarities of data items instead of their distances (known as k-modes). The core of these algorithm is to construct clusters through the cost function (known as distance measure), and then compute centroid of the clusters to aggregate data items that converge to the respective centroids. However, these solution rely on fixed number of clusters (i.e., k), as a result, non-homogeneous data items are also mapped to the clusters. Moreover, identifying semantically coherent multivariate data items, and then anonymizing such data by retaining utility is also a concern [11].

1.1 Contributions

In this paper, we propose a new solution (i.e., n-means) to microaggregate multivariate categorical data that maps formal semantics of the records through the lexical database (i.e., WordNet[1]). For this purpose, we extend features of k-means clustering algorithm and adapt them in our proposed approach in order to achieve notion of microaggregation. Hence, following contributions are proposed:

- We propose a new microaggregation method (i.e., n-means) that aggregate homogeneous categorical records in clusters by semantically analyzing data through the ontology. Later, these clusters are anonymized through the proposed method that relies on the lexical database.
- Contrary to existing solutions [9, 10] (which rely on fix number of clusters), we propose an adaptive clustering mechanism that construct clusters according to the semantics of the datasets. As a result, the clusters are semantically cohesive, and the anonymized data generated as a result of microaggregation method emits less information loss.
- We extend *cost function* of k-means algorithm (which is required to construct clusters) to deal with the categorical data by exploiting lexical database in order to drive ontological relation of the data items (this database holds conceptual semantics of data items in an ontology). Moreover, this proposed method aggregate records in partial clusters by analyzing multivariate of the records. These partial clusters add more randomness in records to de-identify data, as a result, guarantees privacy of the users.
- Finally, we propose a novel microaggregation algorithm (i.e., n-means) that incorporates above mentioned methods.

Rest of the paper is organized in the following manner. Section 2 illustrates state of the art. In Sect. 3, we detail our contributions and newly proposed algorithm to microaggregate data. We evaluate our proposed model in Sect. 4. Finally, conclusion is drawn in Sect. 5.

[1] https://wordnet.princeton.edu/.

fridaya@uef.fi

2 Related Work

The scientific community significantly contributed to improve clustering problem for categorical data. In this regard, Wei et al. [12] proposed a model that analyze semantics of data items through their interrelations derived from the taxonomy. In this model, the main focus is to derive true lexical semantics of data and to address dimensionality issues. For this purpose, they proposed measures to derive semantics of data from the lexical database, and semantically disambiguate them to address dimensionality issues. Based on it, they measure distance of data elements through the ontological branches to derive similarity between the concepts. In this approach, clusters are formed based on semantics only but it does not specify number of clusters, thus, it requires modifications to be considered for microaggregation. In another approach, Ben Salem et al. [13] proposed a model that is inspired in k-means algorithm, which is extended for categorical datasets. For this purpose, the data items are assigned numeric values determined through the dissimilarity measure (i.e., 1 for similar data items and 0 for non-similar), which is used to measure frequency of data items in the later step. Based on the frequency determined, mode of datasets is calculated that measures centroid of the clusters. This proposed mechanism bear some limitations, as the number of clusters are fixed (i.e., k clusters). Thus, it hampers the homogeneity of clusters due to addition of non-similar data items in order to form k-cluster.

Cluster formation has also been a significant process in microaggregation methods. In a similar approach, Batet et al. [1] proposed a framework to anonymize query logs using semantic microaggregation method. This framework relies on MDAV method [14] to anonymize set-valued data by analyzing semantics of query logs using Open Directory Project as knowledge source. In this method, records are arranged in clusters based on their semantic distance from the centroid of the cluster, which is the aggregated distance of data items of the query. Later, such records are anonymized by choosing appropriate concepts for each data elements adhering centroid of corresponding data items within the cluster. However, due to fixed number of clusters, the distant records are also accommodated in the cluster that may not be semantically coherent with other records. In another semantic-based approach, Han et al. [11] addressed limitations of approaches relying on k-mode method [9] by incorporating generalization and microaggregation methods. To achieve this, distance measure for categorical data computes a common node for attributes from value generalization hierarchies (VGH) after semantic analysis. Similarly, distance for continuous data is determined through the Euclidean distance measure. These both distances are used to determine centroid of the mixed data clusters, which is later used for anonymizing data. This scheme is constrained to structured data, and semantics are derived from VGH.

3 Proposed Model

In this section, we propose a novel approach (i.e., *n*-means) to microaggregate categorical medical data. Usually, microaggregation-based methods aggregate homogeneous records in clusters, and later such records are replaced with the centroid record of the cluster in order to achieve notion of k-anonymity [15]. For this purpose, we

fridaya@uef.fi

construct clusters by extending k-means algorithm that accumulates categorical data. As already discussed in Sect. 1, k-means algorithm was initially proposed to aggregate *continuous data* (i.e., numeric data) in k number of clusters. To adapt this algorithm, we extend its following functions: (i) *mean function* that computes centroid of the data values, and (ii) a *cost function* that measures Euclidean distance between the data items and the estimated centroid of the cluster in order to aggregate data items to their closely matched clusters. However, while dealing with *categorical data* (e.g., medical records), such features need to examine semantics of the records before aggregating them in their respective clusters. For this purpose, we extend *cost function* with a taxonomy-based feature that computes semantic distance between categorical data items (method detailed in Sect. 3.1). Similarly, *mean* (or centroid) of categorical data can also be computed by exploiting taxonomic relations of the data items, and an item is chosen as centroid of the cluster that is semantically least distant to other data items (method detailed in Sect. 3.2).

Moreover, k-means is constrained to fixed number of clusters (i.e., k), hence, a data item can be appended to such cluster that is far distant to this data item (distance is measured between the centroid of the cluster and the data item). Thus, this practice of k-means hampers homogeneity of clusters due to addition of non-semantically similar records. As a result, data sustains huge information loss after the redaction methods are enforced. To tackle this issue, we introduce a notion of adaptive clusters that accumulates homogeneous datasets in n clusters, where the size n is not a fix number but determined after the semantic analysis of the dataset.

For semantic analysis, we rely on lexical database (i.e., WordNet) that holds conceptual semantics (i.e., synonyms or concepts) of categorical data items in a taxonomy. Such concepts are taxonomically arranged through a relation in which top level concepts are generalized items and lower levels are their specialized items. These modified features are explained in the following subsections.

3.1 Cost Function

In a medical record database, health records are usually stored in relations, where each relation has a set of distinct attributes usually represented in columns. In this relation, a tuple illustrates a distinct record of patient. Such attributes encompass several values (we consider categorical datasets only), for example: an attribute *Disease* can have set of values such as *typhoid*, *hepatitis*, *HIV* representing patients' records. Therefore, the microaggregation methods generate huge information loss if they focus to replace data items with the centroid value of the *disease* attribute (e.g., if *hepatitis* is the centroid then all data items (i.e., *HIV*, *typhoid*) may be replaced with *hepatitis*). In view of this, we propose to construct clusters for such attributes, where each cluster aggregates semantically coherent data items for a given attribute (e.g., a cluster c_j holds all types of liver diseases only). As a result of redaction methods, all data values of the cluster are substituted with the centroid of the cluster that only holds semantically coherent records, thus it retains data utility. The process to construct such clusters is explained in Sect. 3.2.

fridaya@uef.fi

For this purpose, we extend *cost function* of k-means algorithm to deal with categorical data by relying on taxonomic relations of lexical database (i.e., WordNet) (this cost function was initially designed to operate on continuous data). Proposed function computes semantic distance between the data items of a dataset in order to aggregate semantically-coherent items within same clusters. In k-means algorithm, initially k centroids are chosen at random from the dataset (k also indicate number of clusters), and the distance between a data item and all centroids is measured. Hence, a data item that is least distant to the centroid forms a mutual cluster. These processes (i.e., centroid selection and computing cost of data items) are repeated until all data items converge to the centroid of the clusters. Although, this method is trivial for continuous datasets, but it requires special handling for categorical datasets that has diverse semantics. Therefore, it is significant to drive level of dissimilarity between categorical data items before aggregating them in clusters.

For this purpose, we rely on taxonomy-based measure [16] (Eq. (1)) that computes semantic distance between medical records from their taxonomic relations derived through the ontological database (i.e., WordNet). This measure also states the degree of dissimilarity between such attributes. In this equation, the distance of each data item 'x' is measured with respect to the centroid '' of the clusters to relocate it in least-distant cluster comprising of similar items (method explained in Sect. 3.4). The distance is computed by calculating uncommon ancestors between two concepts divided by the total number of ancestors they are sharing in a taxonomy (where $T(x)$ represents generalized concepts in a taxonomic branch of data items x). Whereas, the logarithm adds smoothing factor to the difference of the compared concepts, and to avoid the condition of Log(0) factor (1+) is added to the equation.

$$dist(x, \mu) = \log_2\left(1 + \frac{|T(x) \cup T(\mu) - T(x) \cap T(\mu)|}{|T(x) \cup T(\mu)|}\right) \tag{1}$$

In this system, a database relation D has set of attributes ($D = \{m_1, m_2, m_3 \ldots. m_z\}$), and each attribute m_a has a set of categorical data values (i.e., $m_a = \{x_1, x_2 \ldots. x_i\}$) that is represented in columns. Tuples of this relation holds multivariate representing a complete record of a patient. Though, due to diverse semantics of attributes, it is rare that such tuples has common values for all multivariate. For example, in Table 1, an attribute m_1 is common (due to semantic similarity of data values) in records R_1, R_2 and R_3, whereas an attribute m_2 is only common in R_1&R_2. Due to this reason, it is challenging to aggregate homogeneous records in clusters due to the diversity in the values of attributes. To tackle this issue, we aggregate records in clusters in the following steps: (i) first we aggregate semantically similar univariate in partial clusters, and then based on this, (ii) we construct clusters holding multivariate of records.

For univariate, we construct set of partial clusters of each attribute m_a (i.e., set of clusters $\acute{C}m_a = \{c_1, c_2, \ldots. c_y\}$) that aggregate semantically similar data items of an attribute m_a. These clusters are adaptive in nature, which determine number of clusters and their data items after semantic analysis of a dataset. To achieve this, initially two

fridaya@uef.fi

random centroids (i.e., set of centroids $Cen(m_a) = \{\mu_1, \mu_2\}$) are chosen from the data values of each attribute m_a. Then, distance of each data item x_i (such that $x_i \in m_a$ attribute) is measured with each estimated set of centroids (i.e., $Cen(m_a)$). Hence, a data item that is least distant to the respective centroid μ_j (distance measured through Eq. (1)) are coupled to its corresponding cluster c_j (this notion is illustrated in Eq. (2)). In order to ensure homogeneity of records in clusters, new clusters are constructed for the data items that are far distant to the centroids. Moreover, during each iteration new centroids are also added to the set of centroids $Cen(m_a)$ (process to compute centroid is detailed in Sect. 3.3). Hence, this process of creating clusters and computing centroid is repeated iteratively until all data items converge to the centroid of their respective clusters (details in Sect. 3.4). As a result of this process, each cluster aggregate homogeneous data items.

$$\arg \min_{\mu_j \in c_j, x_i \in m_a} \sum_{j=1}^{k} dist(x_i, \mu_j) \tag{2}$$

Example 1: Following Table 1 illustrates records of several patients in tuples, and each record R_i has set of attributes (i.e., m_1, m_2, m_3, m_4) along with their categorical data values (represented in column), which are aggregated in clusters as a result of cost function (i.e. Eq. (2)). Each attribute has set of clusters holding semantically coherent data items, which are represented through different pattern (e.g., an attribute m_1 has clusters $c_1 = \{a_1, a_2, a_3\}$ and $c_2 = \{\acute{a}_1, \acute{a}_2\}$, similarly, an attribute m_2 has clusters $c_1 = \{b_1, b_2\}$ and $c_2 = \{ß_1, ß_2, ß_3\}$).

Table 1. Univariate cluster formation of a database relation D

	m_1	m_2	m_3	m_4
R_1	a_1	b_1	$\acute{c}_1$	d_1
R_2	a_2	b_2	$\acute{c}_2$	d_2
R_3	a_3	$ß_1$	$\acute{c}_3$	$\acute{d}_1$
R_4	$\acute{a}_1$	$ß_2$	$\acute{c}_1$	$\acute{d}_2$
R_5	$\acute{a}_2$	$ß_3$	$\acute{c}_2$	$\acute{d}_3$

This cost function construct set of clusters for distinct attributes of a database relation D, now we require a mechanism to construct clusters comprising of semantically similar records that are holding multivariate. This method is discussed in following Sect. 3.2.

fridaya@uef.fi

3.2 Multivariate Cluster Formation

As illustrated in Table 1, consider that the tuples of a database relation holds medical record of distinct patients (i.e., R_1,…,$R_{5)}$. In the next step, we aggregate such records in clusters that accumulate multivariate of these records. In Table 1, it is important to note that the size of clusters of univariate attributes is not uniform. For example: a cluster size for attributes m_1 and m_3 is 3 that hold records R_1, R_2 and R_3, whereas, clusters for attributes m_2 and m_4 lacks a record R_3. Hence, we cannot construct a multivariate cluster comprising of records R_1, R_2 and R_3 nor a cluster with records R_1 and R_2, because in each selection some of the attribute items are missing or not semantically similar. Thus, constructing clusters for such states may yield to information loss due to aggregation of non-semantically similar data items. To deal with this situation, we split records in several partial clusters holding multivariate, and then accumulate them for redaction methods. For this purpose, we aggregate univariate clusters (already holding semantically similar data items) in groups that are equal in size (i.e., number of records).

To achieve this, we use Eq. (3) that measure cardinality of clusters holding univariate of records. This equation measures cardinality of clusters for each attribute m_a that holds set of univariate clusters $\acute{C}m_a = \{c_1, c_2, \ldots . c_y\}$, where each cluster c_i holds set of data items. As a result, cardinality score of each respective cluster is measured in a Φm_a (i.e., $\Phi m_a = \{\bar{\bar{c}}_1, \bar{\bar{c}}_2 \ldots ... \bar{\bar{c}}_y\}$ such that $\bar{\bar{c}}_p$ holds cardinalities of a cluster c_p of all univariate attributes).

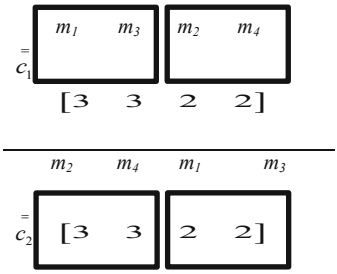
$$\Phi m_a = card(\acute{C}m_a) \tag{3}$$

As a result of Eq. (3), we construct following multivariate matrix (determined from example (1) that holds cardinality score of each cluster of attributes. In this matrix, rows hold corresponding clusters' scores. For example, cluster c_1 of an attribute m_1 has cardinality score 3 (i.e., $\bar{\bar{c}}_1$).

$$\begin{array}{c} \\ \bar{\bar{c}}_1 \\ \bar{\bar{c}}_2 \end{array} \begin{array}{c} \begin{array}{cccc} m_1 & m_2 & m_3 & m_4 \end{array} \\ \begin{bmatrix} 3 & 2 & 3 & 2 \\ 2 & 3 & 2 & 3 \end{bmatrix} \end{array} \tag{11}$$

Now we divide cardinal matrix in row vectors (as shown below), and construct clusters that have similar cardinality holding multivariate of records (e.g., for row vector $\bar{\bar{c}}_1$, attributes m_1 & m_3 share a common cluster (as both have same cardinality), and attributes m_2 & m_4 share a cluster).

fridaya@uef.fi



Now, we substitute actual data items in this matrix, as a result, we have following Table 2 illustrating partial clusters holding multivariate of records. Hence, we have a set of multivariate clusters $\hat{C}_m = \{(c'_{1a}, c'_{1b}), (c'_{2a}, c'_{2b}) \ldots\ldots (c'_{ia}, c'_{ib})\}$ of a dataset D, where each pair (i.e., (c'_{ia}, c'_{ib})) represents partial cluster of a distinct attribute m_n. A cluster c'_{ia} denotes the i^{th} cluster holding some of the attributes of a relation D, and 'a' represents a segment of i^{th} cluster.

Table 2. Multivariate clusters formation for a relation D

	$\hat{c}_1$			
	c'_{1a}		c'_{2a}	
	m_1	m_3	m_2	m_4
R_1	a_1	$ç_1$	b_1	d_1
R_2	a_2	$ç_2$	b_2	d_2
R_3	a_3	$ç_3$		

	$\hat{c}_2$			
	c'_{1b}		c'_{2b}	
	m_1	m_3	m_2	m_4
R_3	$á_1$	$ć_1$	$ß_1$	$đ_1$
R_4	$á_2$	$ć_2$	$ß_2$	$đ_2$
R_5			$ß_3$	$đ_3$

Based on this formation, we can calculate centroid of these multivariate clusters that is required for microaggregation-based systems. The method to calculate centroid is discussed in the following section.

3.3 Centroid of Clusters

As k-means was initially proposed to compute mean of continuous dataset. However, computing mean of categorical datasets is not a straightforward procedure due to diversity in semantics. Therefore, we compute mean of clusters through the ontological relations of the data items retrieved from the WordNet database. Our solution compute following two types of clusters, which are (i) univariate clusters (Sect. 3.1) and (ii) multivariate clusters (Sect. 3.2). Therefore, it is essential to compute centroid for both of these situations.

fridaya@uef.fi

To compute centroid of univariate clusters, we choose a data item as centroid of the cluster that is least distant to other data items of the cluster. In order to compute this distance, we rely on taxonomic tree that maps all data items of the cluster. Hence, a least distant data item is the one that lies in the center of all data items in a taxonomic tree. To achieve this, Eq. (4) determines centroid μ_j of the clusters by computing a data item that is least distant to all other data items of the cluster. For this purpose, we rely on distance measure discussed in Eq. (1) that computes distance of the data items (in this case another data item is used in place of centroid μ).

$$\mu_j = \arg \min_{(x_i, x_j) \in c_j} \sum_{i=1, j=i+1}^{i=n-1, j=n} dist(x_i, x_j) \tag{4}$$

In another case, we have set of multivariate clusters holding adaptable size records that are determined after semantic analysis of a dataset. Now, we require centroid (or mean) of these clusters that is another important feature of microaggregation-based algorithms. For this purpose, we compute centroid of each multivariate cluster of the cluster set $\hat{C}_{\eta}$. To achieve this, following steps are performed: (i) we compute semantic similarity score of each tuple of the cluster c_{jp}^{*} comprising of multivariate, then, (ii) from each partial set c_{jp}^{*}, we chose a tuple as centroid that has maximum score of similarity within a cluster.

In order to compute semantic similarity score of tuples, first we compute similarity score of each data item with the other items of an attribute. To do so, we determining sum of semantic distance of each data item x_i with all possible distinct combination of other n items of partial cluster c_{jp}^{*} (distance is measured using Eq. (1)). Hence, the similarity is determined by taking a complement of the distance computed for the paired data values (i.e., x_i & x_j) (illustrated in Eq. (5)). As a result, each data item has its similarity score. In the next step, we compute cumulative score of tuples by aggregating similarity score of data items of multivariate in respective tuples. Thus, a tuple that has maximum aggregated similarity score is chosen as the centroid $\acute{\eta}_{jp}$ of the multivariate cluster.

$$score(x_i) = \sum_{x_i, x_j \in c_j}^{n} (1 - dist(x_i, x_j)) \tag{5}$$

3.4 Clustering and Anonymization of Data

As discussed earlier, it is vital to adhere homogeneity of data items while constructing clusters in order to minimize information loss during anonymization methods. For this purpose, we introduce a notion of adaptive clusters where the size of clusters is not fixed, but determined after the semantic analysis of the dataset. Hence, for a given set of multivariate categorical attributes D = {m_1, m_2, m_3….. m_z}, adaptive n-means algorithm partitions these attributes in n clusters (where $n < z$), such that n is determined after performing lexical semantics analysis of given population (determined through the lexical database).

fridaya@uef.fi

Algorithm-1 illustrates the procedure to construct such type of clusters. The input to this algorithm is the set of categorical attributes (i.e., D) that require to be partitioned in n clusters comprising of homogeneous data items (lines 1–2). First, we construct univariate clusters that are processed further and aggregated in multivariate clusters in the later part of the algorithm. For univariate clusters, each attribute m_r of a dataset D is processed individually (line 3). Initially, two centroid (i.e., $k = 2$) are chosen at random from the given data items of an attribute m_r (line 4). Hence, distance of each data item x_i is computed against the set of centroids (i.e., Cen(m_r)), as a result, a data items x_i that is least distance to the centroid μ_j is coupled to the respective cluster of the centroid (using Eq. (2)) (line 7–8).

In addition, the algorithm tracks all data items that are far distant to the set of centroids. Conceptually, such data items are not semantically coherent to all centroids, thus, retaining in any of the clusters will hamper the homogeneity of records. For this reason, the distance of each paired items (i.e., x_i and μ_j) is compared with the threshold distance factor ω (line 9–10), which denotes the maximum distance that two distances may differ. This factor ranges between $0 < \omega < 1$, however, the distance of closely matched data items (which need to form a common cluster) must be less than ω (but it must be close to '0' to guarantee homogeneity). Hence, such data items are added to the set of centroid Cen(m_r) (line 11), as a result, new clusters are constructed in next iteration of the algorithm.

In the next step, we compute new centroids of the clusters (line 13) (procedure discussed in Sect. 3.3), and repeat the same process (i.e., lines 6–13) until all data items converge to the centroid of their respective clusters (this condition occurs when new clusters are not constructed) (line 14). At this stage, we have constructed univariate clusters that form a base for the multivariate clusters. For multivariate clusters, each attribute m_r is processed individually and cardinality of univariate clusters is measured through Eq. (3) (line 16–18). Hence, multivariate clusters are constructed for the attributes that are equal in size and added in a clusters set $\hat{C}_m$ (line 19–20) (details in Sect. 3.2).

fridaya@uef.fi

Algorithm 1: Adaptive n-means Microaggregation

1. Input: set of attributes D={m_1, m_2, m_3 ….. m_z}
2. initially k=2
   // univariate clusters
3. **for** each attribute m_r
4. choose k random centroids(i.e., $Cen(m_r)=\{\mu_1,\mu_2,\ldots\mu_k\}$ from the data items of m_r
5. **do**
6. **for** each data item x_i of an attribute m_r
7. find nearest centroid μ_j such that

$$\arg \min_{\mu_j \in C_j, x_i \in m_r} \sum_{j=1}^{k} dist(x_i, \mu_j)$$

8. assign data item x_i to cluster c_j (as $c_j \in$ cluster set $\acute{C}m_a$)
9. $dist_{xi} \leftarrow$ distance between x_i and μ_j
10. **if** $dist_{xi} > \omega$
11. add x_i as centroid in a set $Cen(m_r)$
12. **end for**
13. compute new centroids of all clusters using

$$\mu_j = \arg \min_{(x_i, x_j) \in c_j} \sum_{i=1, j=i+1}^{i=n-1, j=n} dist(x_i, x_j)$$

14. **until** (convergence)
15. **end for**
    // multivariate clusters
16. **for** each attribute m_r
17. compute cardinality of each cluster of m_r using

$$\Phi m_a = card(\acute{C}m_a)$$

18. **end for**
19. Construct cardinal matrix from Φm_a and construct multivariate partial clusters of equal size
20. store partial cluster in a cluster set $\hat{C}_m$
    // centroid of multivariate clusters
21. **for** each partial cluster c_{jp} (as $c_{jp} \in \hat{C}_m$)
22. compute semantic similarity score of each tuple of a cluster c_{jp} by using

$$score(x_i) = \sum_{x_i, x_j \in c_j}^{n} (1 - dist(x_i, x_j))$$

23. **end for**
24. chose a tuple as centroid that has maximum score of similarity within a cluster
25. **for** each partial cluster c_{jp}
26. replace all tuples with the centroid record($\acute{\eta}_{jp}$)
27. **end for**

fridaya@uef.fi

Once we have multivariate clusters, it is important to compute centroid of the tuples that is required for the redaction method. For this purpose, we compute similarity score of each tuple of the partial clusters holding multivariate through the Eq. (5) (lines 21–23). As a result, a tuple that has maximum score is chosen as the centroid of the partial cluster (line 24). The details are provided in Sect. 3.3.

At this stage, we have a set of homogeneous clusters that can be anonymized to ensure privacy of users. For this purpose, the centroid $\acute{\eta}_{jp}$ calculated for each cluster $c^{'}_{jp}$, in the above mentioned algorithm, is chosen to be replaced with all records of the cluster $c^{'}_{jp}$ (lines 25–27). As a result, following Table 3 illustrates the anonymized version derived from the clusters of Table 2 (similar colors shows semantically similar data items). In this table each record R_i is similar to multiple records, and due to the variety of semantically similar attributes within records increases more randomness to de-identify the data. For example: R_1 is similar to records R_2 and R_3 at the same time, and it is difficult to identify a person due to increased randomness in the attributes.

Table 3. Anonymized version of the records

	m_1	m_2	m_3	m_4
R_1	a_2	b_2	$\acute{c}_2$	d_2
R_2	a_2	b_2	$\acute{c}_2$	d_2
R_3	a_2	$ß_1$	$\acute{c}_2$	$đ_2$
R_4	$á_1$	$ß_1$	$\acute{c}_1$	$đ_2$
R_5	$á_1$	$ß_1$	$\acute{c}_2$	$đ_2$

4 Evaluation

In order to measure effectiveness of our proposed system, we ground our analysis on two important factors, which are (i) cohesiveness of clusters, and (ii) information loss. Cohesiveness of clusters signifies that how closely data items are coherent to each other within the cluster. Whereas, information loss measure the degree of data lost after applying microaggregation method. In order to measure cohesion of clusters that hold categorical data, we determine the degree of data items deviate from the centroid of the cluster. For this purpose, we consider univariate clusters for this analysis, as these cluster as constructed based on the semantic similarity of data items, which also form a base for multivariate clusters. To do so, following measure (Eq. (6)) is proposed to quantify the degree of dispersion of data items from the centroid μ_j of the cluster c_j, where N is the total number of elements within the clusters c_j. A low cohesive factor indicates that the data items are more close to the centroid and the cluster is more cohesive, whereas a cluster is less cohesive in the other case.

fridaya@uef.fi

$$\underset{\mu_j \in c_j, x_i \in c_j}{Cohesion} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} dist\left(x_i, \mu_j\right)} \tag{6}$$

Similarly, information loss can be measured as the ratio between the *Sum of Square Errors* (SSE) and the *Total Sum of Squares* (SST) (Eq. (7)), which is widely used by many researchers [17, 18].

$$IL = \frac{SSE}{SST} \times 100 \tag{7}$$

In this method, SSE is measured the as sum of square of distances between the centroid and the data items of each cluster (Eq. (8)), where m is the total number of clusters and n_j is the total number of elements in cluster c_j.

$$SSE = \sum_{j=1}^{m}\sum_{i=1}^{n_j}\left(dist(x_i, \mu_j)\right)^2 \tag{8}$$

Whereas, SST measures sum of square of distances between the centroid of overall dataset and each data item (Eq. (9)).

$$SST = \sum_{j=1}^{m}\sum_{i=1}^{n_j}\left(dist(x_i, \mu_D)\right)^2 \tag{9}$$

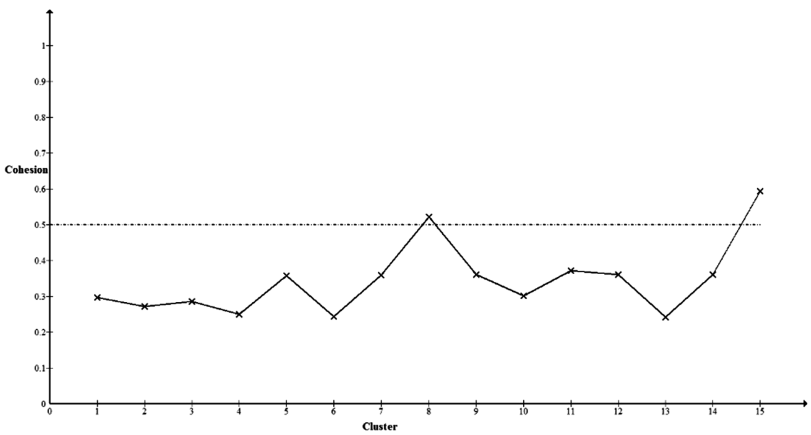


Fig. 1. Cohesion of clusters

fridaya@uef.fi

To achieve above mentioned objectives, we took medical records from online social network[2] that is publically available. We took 1000 medical records of 500 patients and extracted 2073 noun phrases for our study. Figure 1 illustrates cohesion of clusters measured for 15 different clusters. This figure implies that the majority of the clusters have cohesive factor less than the average dissimilarity factor (i.e., 0.5), thus they hold homogeneous data elements.

In our solution, the clusters are self-adaptive; therefore, we measured information loss for each distinct cluster. We compared information loss emitted by our solution with existing solutions (i.e., Abril et al. [18] and Martinez et al. [4]), and results are shown in Fig. 2. In this illustration, our system out performs when the size of the clusters (no of element i.e., k) is less than 14 as compare to large size clusters. One of the reasons behind the hike in projection (for $k > 14$) is due to the merging of non-semantically similar items within the last few clusters that were leftovers, and had no concrete semantic coherence with any cluster.

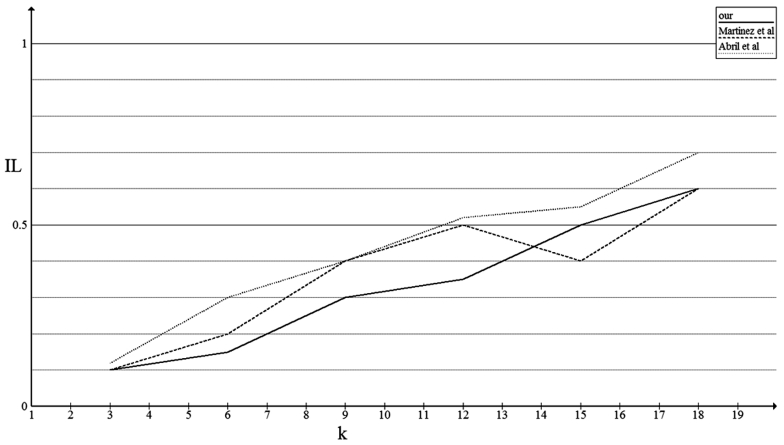


Fig. 2. Comparison of information loss with existing solutions

5 Conclusion

In this paper, we present a novel approach to microaggregate univariate categorical data by adapting k-means clustering algorithm. For this purpose, k-means algorithm is extended to drive ontological relations of concepts from the lexical database (i.e., WordNet) to form clusters. Moreover, number of clusters is not fixed, but adapted after analyzing semantics of dataset, which results in more cohesive clusters with homogeneous attributes. Evaluation proved that the proposed solution performs better than the existing approaches by retaining utility and emitting less information loss. Moreover, clusters aggregate semantically coherent records and are more cohesive.

[2] https://www.patientslikeme.com/.

fridaya@uef.fi

As a future work, cluster formation of existing system can be extended to incorporate multivariate of unstructured data, and based on this propose a new microaggregation model. Moreover, such system can be evaluated to study the impact on information loss as a result of multivariate microaggregation.

Acknowledgment. We acknowledge Higher Education Commission Pakistan and Foundation University Islamabad for their support to publish this research work.

References

1. Batet, M., Erola, A., Sánchez, D., Castellà-Roca, J.: Utility preserving query log anonymization via semantic microaggregation. Inf. Sci. **242**, 49–63 (2013)
2. Domingo-Ferrer, J.: Microaggregation. In: Liu, L., Özsu, M.T. (eds.) Encyclopedia of Database Systems, pp. 1736–1737. Springer, Boston (2009)
3. Sweeney, L.: k-anonymity: a model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **10**, 557–570 (2002)
4. Martínez, S., Sánchez, D., Valls, A.: Semantic adaptive microaggregation of categorical microdata. Comput. Secur. **31**, 653–672 (2012)
5. Domingo-Ferrer, J., Sánchez, D., Rufian-Torrell, G.: Anonymization of nominal data based on semantic marginality. Inf. Sci. **242**, 35–48 (2013)
6. Erola, A., Castellà-Roca, J., Navarro-Arribas, G., Torra, V.: Semantic microaggregation for the anonymization of query logs. In: Privacy in Statistical Databases, pp. 127–137. Springer, Heidelberg (2010)
7. Ahmad, A., Dey, L.: A k-means type clustering algorithm for subspace clustering of mixed numeric and categorical datasets. Pattern Recogn. Lett. **32**, 1062–1069 (2011)
8. Huang, Z.: Extensions to the k-means algorithm for clustering large data sets with categorical values. Data Min. Knowl. Disc. **2**, 283–304 (1998)
9. Jiang, F., Liu, G., Du, J., Sui, Y.: Initialization of K-modes clustering using outlier detection techniques. Inf. Sci. **332**, 167–183 (2016)
10. Kuo, R.J., Potti, Y., Zulvia, F.E.: Application of metaheuristic based fuzzy K-modes algorithm to supplier clustering. Comput. Ind. Eng. **120**, 298–307 (2018)
11. Han, J., Yu, J., Mo, Y., Lu, J., Liu, H.: MAGE: a semantics retaining K-anonymization method for mixed data. Knowl.-Based Syst. **55**, 75–86 (2014)
12. Wei, T., Lu, Y., Chang, H., Zhou, Q., Bao, X.: A semantic approach for text clustering using WordNet and lexical chains. Expert Syst. Appl. **42**, 2264–2275 (2015)
13. Ben Salem, S., Naouali, S., Chtourou, Z.: A fast and effective partitional clustering algorithm for large categorical datasets using a k-means based approach. Comput. Electr. Eng. **68**, 463–483 (2018)
14. Domingo-Ferrer, J., Mateo-Sanz, J.M.: Practical data-oriented microaggregation for statistical disclosure control. IEEE Trans. Knowl. Data Eng. **14**, 189–201 (2002)
15. Templ, M., Meindl, B., Kowarik, A., Chen, S.: Introduction to Statistical Disclosure Control (SDC). IHSN Working Paper No. 007 (2014)
16. Sánchez, D., Batet, M., Isern, D., Valls, A.: Ontology-based semantic similarity: a new feature-based approach. Expert Syst. Appl. **39**, 7718–7728 (2012)

fridaya@uef.fi

17. Domingo-Ferrer, J., Martínez-Ballesté, A., Mateo-Sanz, J.M., Sebé, F.: Efficient multivariate data-oriented microaggregation. VLDB J. **15**, 355–369 (2006)
18. Abril, D., Navarro-Arribas, G., Torra, V.: Towards semantic microaggregation of categorical data for confidential documents. In: Modeling Decisions for Artificial Intelligence, pp. 266–276. Springer, Heidelberg (2010)

fridaya@uef.fi

An Efficient Density-Based Clustering Algorithm Using Reverse Nearest Neighbour

Stiphen Chowdhury[1(✉)] and Renato Cordeiro de Amorim[2]

[1] School of Computer Science, University of Hertfordshire, College Lane Campus, Hatfield AL10 9AB, UK
s.chowdhury8@herts.ac.uk

[2] School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK
r.amorim@essex.ac.uk

Abstract. Density-based clustering is the task of discovering high-density regions of entities (clusters) that are separated from each other by contiguous regions of low-density. DBSCAN is, arguably, the most popular density-based clustering algorithm. However, its cluster recovery capabilities depend on the combination of the two parameters. In this paper we present a new density-based clustering algorithm which uses reverse nearest neighbour (RNN) and has a single parameter. We also show that it is possible to estimate a good value for this parameter using a clustering validity index. The RNN queries enable our algorithm to estimate densities taking more than a single entity into account, and to recover clusters that are not well-separated or have different densities. Our experiments on synthetic and real-world data sets show our proposed algorithm outperforms DBSCAN and its recent variant ISDBSCAN.

Keywords: Density-based clustering · Reverse nearest neighbour · Nearest neighbour · Influence space

1 Introduction

Clustering algorithms aim to reveal natural groups of entities within a given data set. These groups (clusters) are formed in such a way that each contains homogeneous entities, according to a pre-defined similarity measure. This grouping of similar entities is usually data-driven and by consequence it does not require information regarding the class label of the entities. Detecting, analysing, and describing natural groups within a data set is of fundamental importance to a number of scientific fields. Thus, it is common to see clustering algorithms being applied to problems in various fields such as: bioinformatics, image processing, astronomy, pattern recognition, medicine, and marketing [1–4].

There are indeed a number of different approaches to clustering. Some algorithms were designed so they could be applied to data sets in which each entity is

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 29–42, 2019.
https://doi.org/10.1007/978-3-030-22868-2_3

fridaya@uef.fi

described over a number of features. Others, take as input a dissimilarity matrix or even the weights of edges in a graph. There are different formats for the final clustering as well. The clusters may be a partition of the original data set, or they may present overlaps so that an entity belongs to more than one cluster (usually at different degrees, adding to one). They may also be non-exhaustive so that not every entity belongs to a cluster, which can be particularly helpful if the data set contains noise entities. We may also have hierarchical clusterings, which may be generated following a top-down or bottom-up approach. We direct readers interested in more details to the literature (see for instance [1,3,4] and references therein). In this paper we focus on density-based clustering. This approach defines clusters as areas of higher density separated by areas of lower density. Clearly, such loose definition may raise a number of questions regarding what exactly a cluster is (or is not!). However, given there is no generally accepted definition for the term *cluster* that works in all scenarios, one can raise similar questions even if using non density-based algorithms. Defining 'true' clusters is particularly difficult and may also depend on other factors than the data set alone (for a discussion see [5] and references therein). The major advantage a density-based algorithm has is that the impact of a similarity measure on the shape bias of clusters is considerably reduced.

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [6] is arguably the most popular density-based clustering algorithm. A recent search in Google Scholar for the term "DBSCAN" returned a total of 21,500 entries. Most importantly searches for the years of 2014, 2015, and 2016 returned 2,190, 2,710, and 3,550, respectively. These numbers support the growing popularity of DBSCAN. Unfortunately, as popular as it may be, DBSCAN is not without weaknesses. For instance: (i) it requires two parameters (for details see Sect. 2); (ii) it is a non-deterministic algorithm, so it may produce different partitions under the same settings; (iii) it is not particularly suitable for data sets whose clusters have different densities.

There have been some advancements in the literature. For instance, OPTICS [7] has been designed to deal with clusters of different densities. Using the concept of k-Influence Space [8,9], ISDBSCAN [10] can also deal with clusters of different densities, and requires a single parameter to be tuned. ISDBSCAN algorithm significantly outperforms DBSCAN and OPTICS [7].

In this paper we make a further advancement in density-based clustering research. Here, we introduce Density-based spatial clustering using reverse nearest neighbour (DBSCRN) a new method capable of matching or improving cluster recovery in comparison to ISDBSCAN (and by consequence DBSCAN and OPTICS), but being orders of magnitude faster. Our method has a single parameter for which we show a clear estimating method.

2 Related Work

The purpose of any clustering algorithms is to split a data set Y containing n entities $y_i \in \mathbb{R}^m$ into K clusters $S = \{S_1, S_2, ..., S_K\}$. Here, we are particularly

fridaya@uef.fi

interested in hard-clustering so that a given entity y_i can be assigned to a single cluster $S_k \in S$. Thus, the final clustering is a partition subject to $S_k \cap S_l = \emptyset$ for $k, l = 1, 2, ..., K$ and $k \neq l$.

It is often stated that density-based clustering algorithms are capable of recovering clusters of arbitrary shapes. This is a very tempting thought, which may lead to some disregarding the importance of selecting an appropriate distance or similarity measure. This measure is the key to produce homogeneous clusters as it defines homogeneity. Selecting a measure will have an impact on the actual clustering. Most likely the impact will not be as obvious as if one were to apply an algorithm such as k-means [11] (where the measure in use leads to a bias towards a particular cluster shape). However, the impact of this selection will still exist at a more local level. If this was not the case, DBSCAN would produce the same clustering regardless of the distance measure in place.

Arguably, the most popular way of calculating the dissimilarity between two entities y_i, y_j each described over m features is given by the squared Euclidean distance, that is

$$d(y_i, y_j) = \sum_{v=1}^{m} (y_{iv} - y_{jv})^2. \tag{1}$$

DBSCAN classifies each entity $y_i \in Y$ as either a core entity, a reachable entity, or an outlier. To do so, this algorithm applies (1) together with two parameters: a distance threshold (ϵ), and the minimum number of entities required to form a dense region ($MinPts$). The ϵ-neighbourhood of an entity $y_i \in Y$ is given by

$$\mathcal{N}(y_i) = \{y_j \in Y \mid d(y_i, y_j) \leq \epsilon\}, \tag{2}$$

so that $\mathcal{N}(y_i) \subseteq Y$. Clearly, $\mathcal{N}(y_i) = Y$ would be an indication the value of ϵ is too high. An entity $y_i \in Y$ is classified as a core entity iff

$$|\mathcal{N}(y_i)| \geq MinPts, \tag{3}$$

in this case each entity in $\mathcal{N}(y_i)$ is said to be directly reachable from y_i. No entity can be directly reachable from a non-core entity. An entity y_i is classified as a reachable entity if there is a path $y_j, y_{j+1}, y_{j+2}, ..., y_i$ in which each entity is directly reachable from the previous. If these two cases (core and reachable) do not apply, then y_i is classified as an outlier. Given a core entity y_i, DBSCAN can form a cluster of entities (core and non-core) that are reachable from y_i. The general idea is, of course, very intuitive but one may find difficult to set ϵ and $MinPts$ as they are problem-dependent.

The ISDBSCAN outperforms the above and OPTICS. Probably, the major reason for this is the use of the k-influence space (IS_k) to define the density around a particular entity. IS_k is based on the k-nearest neighbour (NN_k) [12] and reverse k-nearest neighbour (RNN_k) [13] methods.

$$NN_k(y_i) = \{y_1, y_2, ..., y_j, ..., y_k \in Y \mid d(y_j, y_i) \leq d(y_t, y_i) \forall y_t \in Y'\}, \tag{4}$$

where $Y' = Y \backslash \{y_1, y_2, ..., y_j, ..., y_k\}$, and k is the number of nearest neighbours. The reverse k-nearest neighbours is given by the set

$$RNN_k(y_i) = \{y_j \in Y \mid y_i \in NN_k(y_j)\}, \tag{5}$$

fridaya@uef.fi

leading to the k-influence space

$$IS_k(y_i) = NN_k(y_i) \cap RNN_k(y_i). \tag{6}$$

With the above we can now describe ISDBSCAN.

ISDBSCAN(Y, k)
Input
Y: Data set to be clustered;
k: Number of nearest neighbours;
Output
S: A clustering $S = \{S_1, S_2, ..., S_c, ..., S_K\}$;
S_{noise}: A set of entities marked as noise;
Algorithm:

1. while $Y \neq \emptyset$
2. Randomly select y_i from Y;
3. $S_c \leftarrow$ MakeCluster(Y,y_i,k);
4. $Y \leftarrow Y \backslash S_c$;
5. if $|S_c| > k$ then
6. Add S_c to S;
7. else
8. Add y_i to S_{noise};
9. end if
10. end while
11. return S;
12. *MakeCluster(Y,y_i,k)*
13. $S_c \leftarrow \emptyset$;
14. if $|IS_k(y_i)| > 2/3k$ then
15. for each $y_j \in IS_k(y_i)$ do
16. $S_c \leftarrow S_c \cup \{y_j\}$;
17. $S_c = S_c \cup$ MakeCluster(Y,y_j,k);
18. end for
19. endif
20. return S_c

3 Density-Based Spatial Clustering Using Reverse Nearest Neighbour (DBSCRN)

The algorithm we introduce in this paper, DBSCRN, has some similarities to DBSCAN. They are both density-based clustering algorithms which need to determine whether an entity $y_i \in Y$ is core or non-core. Section 2 explains how this is done by DBSCAN. In the case of DBSCRN this is determined using a reverse nearest neighbour query. Given an entity $y_i \in Y$ we apply Eq. (5) to find

fridaya@uef.fi

the set of entities to which y_i is one of their k-nearest neighbours. We find this to be a more robust method to estimate density because it uses more than just one core entity to find nearest neighbours. We present the DBSCRN algorithm below.

DBSCRN(Y, k)
Input
Y: Data set to be clustered.
k: Number of nearest neighbours.
Output
S : A clustering $S = \{S_1, S_2, \cdots, S_K\}$
Algorithm:

1. for each $y_i \in Y$ do
2. if $|RNN_k(y_i)| < k$ then
3. Add y_i to $S_{non-core}$;
4. else
5. Add y_i to S_{core};
6. $S \longleftarrow S \cup$ *expandCluster(y_i,k, S)*;
7. end if
8. end for
9. Assign each $y_j \in S_{non-core}$ to the cluster of the nearest $y_i \in S_{core}$, using Equation (1);
10. return S;
11. *expandCluster(y_i,k, S)*
12. $S_{y_i} \leftarrow \{y_i\}$;
13. $S_{tmp} \leftarrow \{y_i\}$;
14. for each $y_j \in RNN_k(y_k \in S_{tmp})$ do
15. if $|RNN_k(y_j)| > 2k/\pi$ then
16. $S_{tmp} \leftarrow S_{tmp} \cup RNN_k(y_j)$;
17. end if
18. If $y_j \notin S_{tmp}$ and y_j is not assigned to any cluster in S.
19. Add y_j to S_{y_i};
20. end if
21. end for
22. return S_{y_i};

In the above the quantity of nearest neighbours (k) is a user-defined parameter. The quantity of clusters (K) is automatically found by the algorithm.

4 Estimating Parameters

Here, we take the view that parameter estimation can be accomplished using a Clustering Validity Index (CVI). Validation is one of the most challenging aspects of clustering. It raises the question: how can one measure the quality of

fridaya@uef.fi

a clustering when labeled data is non-existent? the simple fact an algorithm produced a clustering says nothing about the quality of that clustering. Clustering algorithms will produce a clustering even if the data has no cluster structure. A number of CVIs have been proposed to measure the quality of clusterings obtained using distance-based algorithms such as k-means (for a review see [14] and references therein). Selecting a CVI to use is not a trivial matter, it should take into account the definition of cluster in use and any other requirement that may exist. CVIs suitable for density-based clustering algorithms are not as popular. However, they are particularly important as all algorithms we experiment with have at least one parameter that needs to be estimated.

In this paper we do not focus on finding and comparing CVIs suitable for density-based clustering algorithms. One could apply any such CVI to estimate the parameters of the methods we experiment with. With this in mind we leave such comparison for future work. Here, we have experimented with Density-Based Clustering Validation (DBCV) [15]. This CVI measures clustering quality based on the relative density connection between pairs of entities. This index is formulated on the basis of a new kernel density function, which is used to compute the density of entities and to evaluate the within and between-cluster density connectedness of clustering results. This is well aligned to the definition we use of cluster (see Sect. 1).

Using density-based clustering algorithms, DBCV has unsurprisingly outperformed the Silhouette Width [16], the Variance Ratio Criterion [17], and Dunn's index [18]. These three CVIs are not well-aligned with the definition of cluster used by density-based clustering algorithms. DBCV has also outperformed Maulik-Bandyopadhyay [19] and CDbw [20].

5 Setting of Experiments

We experimented with synthetic and real-world data sets, all obtained from the UCI machine learning repository [21]. We selected the data sets described in Table 1, these are rather popular and have been used in a number of publications [22–30]. The clusters in real-world data sets, like Iris, tend to have a globular shape aligned to Gaussian distributions. The synthetic data sets contain arbitrarily shaped clusters of different sizes and densities. All of these data sets allow us to scrutinize the cluster recovery of the clustering algorithms we experiment with.

We have the set of correct labels for each of the data sets we experiment with. This allows us to measure the cluster recovery of each algorithm in relation to the correct labels. In each experiment we generate a set of labels from a clustering solution using a confusion matrix. We then compare the labels of the clustering solution with the correct labels using the adjusted Rand Index (ARI) [31].

$$ARI = \frac{\sum_{ij}\binom{n_{ij}}{2} - \left[\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}\right]/\binom{n}{2}}{\frac{1}{2}\left[\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}\right] - \left[\sum_i \binom{a_i}{2}\sum_j \binom{b_j}{2}\right]/\binom{n}{2}},$$

fridaya@uef.fi

Table 1. Data sets used in our experiments.

	Entities	Clusters	Features
Aggregation	788	7	2
Compound	399	6	2
Pathbased	300	3	2
Spiral	200	2	2
Mixed	1479	5	2
Toy	373	2	2
Flame	240	2	2
R15	600	15	2
Soya	47	4	58
Iris	150	3	4

where $n_{ij} = |S_i \cap S_j|$, $a_i = \sum_{j=1}^{K} |S_i \cap S_j|$ and $b_i = \sum_{i=1}^{K} |S_i \cap S_j|$.

We have standardised the features of each data set by their respective ranges:

$$y_{iv} = \frac{y_{iv} - \bar{y_v}}{max(y_v) - min(y_v)}, \tag{7}$$

where $\bar{y_v} = n^{-1} \sum_{i=1}^{n} y_{iv}$. We chose to use (7) rather than the popular z-score because the latter favours unimodal distributions. For instance, consider two features: a unimodal v_1 and a bimodal v_2. The standard deviation of v_2 will be higher than that of v_1. By consequence the z-score (and the contribution to the clustering) of v_2 will be lower than that of v_1. However, we would be usually interested in the cluster structure present in v_2.

We experiment with three algorithms: DBSCAN, ISDBSCAN, and DBSCRN. Each of these algorithms require the use of parameters, we have estimated these using DBCV. In the case of DBSCAN we run experiments with values for $MinPts$ from 3 to 20 in steps of 1, and ϵ from the minimum pairwise distance to the maximum pairwise distance in steps of 0.1. We selected as final clustering that with the best DBCV index. For ISDBSCAN, we run experiments setting the number of nearest neighbours from 5 to 25 in steps of 1. In the case of DBSCRN we experiment with values of k (the number of nearest neighbours) from 3 to 30, in steps of 1.

All experiments were run on a PC with Intel(R) Core(TM) i7-2670QM CPU 2.20 GHz and 8.00 GB RAM. The operating system was Windows 7 (64-bits). The algorithms were implemented using MATLAB 2016a.

fridaya@uef.fi

Table 2. Experiments comparing k-means, DBSCAN, ISDBSCAN, and DBSCRN. This table reports the best possible ARI each of the algorithms can achieve at each data set. Non-deterministic algorithms were run 100 times.

	k-means			DBSCAN			ISDBSCAN			DBSCRN	
	Mean	Std dev	Max	Mean	Std dev	Max	Mean	Std dev	Max	Mean	Max
Aggregation	0.74	0.03	0.78	0.98	0.002	0.98	0.91	0.02	0.94	–	**0.99**
Compound	0.57	0.10	0.78	0.83	0.00	0.83	0.88	0.01	0.91	–	**0.96**
Pathbased	0.46	0.001	0.46	0.89	0.01	0.9	0.85	0.01	0.89	–	**0.92**
Spiral	0.05	0.01	0.06	**1.00**	0.00	1.00	0.98	0.00	1.00	–	**1.00**
Mixed	0.39	0.02	0.42	**1.00**	0.00	1.00	**1.00**	0.00	1.00	–	**1.00**
Toy	0.31	0.01	0.32	0.96	0.00	0.96	**1.00**	0.00	1.00	–	**1.00**
Flame	0.46	0.02	0.51	**0.96**	0.00	0.96	0.90	0.00	0.90	–	0.93
R15	0.88	0.07	0.99	**0.99**	0.00	0.99	0.94	0.00	0.94	–	**0.99**
Soya	0.80	0.20	1.00	**1.00**	0.00	1.00	0.96	0.00	0.96	–	**1.00**
Iris	**0.67**	0.10	0.71	0.36	0.01	0.37	0.40	0.01	0.47	–	0.45

6 Results and Discussion

In this section we present the results, and discussion, of our experiments. We compare k-means, DBSCAN, ISDBSCAN, and DBSCRN on the data sets presented in Table 1. Our comparison is mainly focused on cluster recovery, measured using the ARI, but we also discuss the amount of time the algorithms take to complete.

In our first set of experiments we aim to show the best possible cluster recovery for each algorithm. Given an algorithm, we set its parameters to those producing a clustering with the highest ARI. This scenario is not realistic as it requires the user to know the correct labels for each data set. However, it allows us to analyse the best possible result for each algorithm. Table 2 presents the results for this set of experiments. Each non-deterministic algorithm was run 100 times.

Table 2 shows that in the vast majority of cases our method is competitive or superior to others in average. The noticeable exception is given by k-means in the Iris data set. In this case none of the density-based clustering algorithms performs well. Most likely, the definition of cluster used in k-means (a globular set of entities in the Euclidean space) is more well-aligned to the clusters in this particular data set. This should remind us that one should define what a cluster is before choosing a clustering algorithm.

Let us analyse in more details some of the results in Table 2. The Compound data set contains three difficult clustering problems: (i) nested clusters with approximately the same density; (ii) nested clusters with different densities; (iii) clusters separated by local minimum density regions. This data set contains two clusters for each of these problems. Figure 2 presents the best possible results for each of the algorithms we experiment with. The k-means algorithm searches for globular clusters in the data set, so it is unable to deal with problems (i) and (ii). Probably the major weakness of DBSCAN is its inability to detect clusters of

fridaya@uef.fi

different densities, leading to 51 out of 399 entities being classifies as noise (red cross, labelled as zero). ISDBSCAN was designed to deal with clusters of different densities, but does not deal well with problems (ii) and (iii) on this occasion. Our method does produce misclassification, but there are considerably less of them than in other methods.

The Flame data set contains two clusters of similar densities separated by either a low density region or a soft boundary. Figure 3 presents the best possible clusterings for each algorithm. We can see k-means is unable to correctly separate these clusters, as they are not Gaussian. DBSCAN does perform particularly well in this data set, but as well as ISDBSCAN it wrongly classifies a few entities as noise. In the case of ISDBSCAN this happens because there is a lower cluster density near the boundary region, leading to the misclassification of entities as noise.

Figure 4 presents the best possible clusterings for each algorithm in the Path-based data set. This data set contains three clusters of equal cardinality in close proximity. These are separated by uneven low density regions. The clustering task is particularly difficult in this data set because two of the clusters are nested inside the third one. Unfortunately, k-means cannot deal with this type of scenario. DBSCAN and ISDBSCAN seem to find noise entities where there should not be any. The clusterings for the Toy data set can be seen in Fig. 5. This data set contains two half-moon clusters of different densities. In these we can see that ISDBSCAN and DBSCRN were the only to correctly recover the two clusters.

Given the data sets we selected for our experiments it is hardly surprising that the density-based algorithms outperformed k-means in most cases. This result should not be interpreted as meaning that density-based algorithms tend to outperform distance-based algorithms. Before clustering what one ought to do is to define the objective of the clustering and then decide what method to use. Finally, in terms of conversion time we can see that DBSCAN is undoubtedly

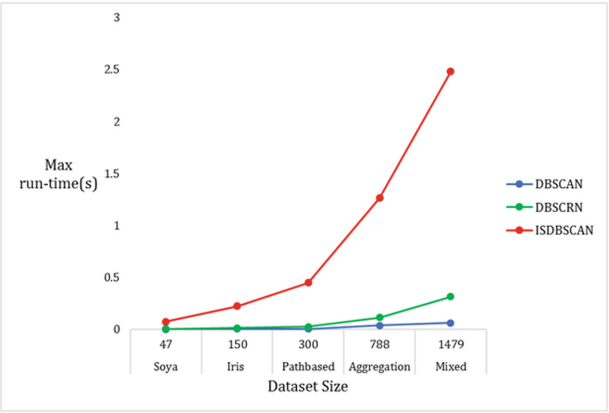
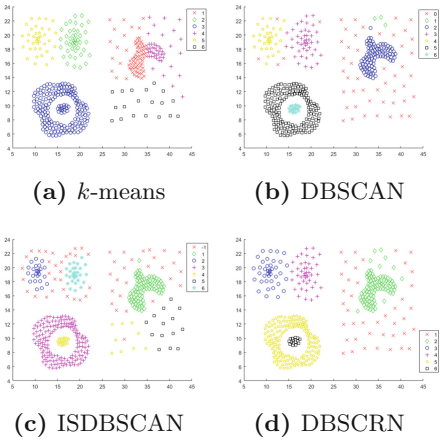


Fig. 1. Maximum run-time for DBSCAN, ISDBSCAN, and DBSCRN.

fridaya@uef.fi



(a) k-means **(b)** DBSCAN

(c) ISDBSCAN **(d)** DBSCRN

Fig. 2. Best possible cluster recovery as measured by the ARI on the Compound data set.

the fastest density-based algorithm we experiment with (see Fig. 1). However, DBSCAN has the worst cluster recovery and it is outperformed by ISDBSCAN and DBSCRN. DBSCRN outperforms ISDBSCAN in terms of cluster recovery and it is orders of magnitude faster than the latter.

We find the results of our previous set of experiments very enlightening, but we feel we need to evaluate the algorithms in a realistic clustering scenario. We know DBSCRN has the best possible cluster recovery in most cases, but now we need to establish whether we can successfully estimate its parameters. With this in mind we ran a new set of experiments in which the parameters of each

Table 3. Experiments comparing DBSCAN, ISDBSCAN, and DBSCRN. The final clustering of each algorithm is that with the highest DBCV index.

	DBSCAN			ISDBSCAN			DBSCRN	
	Mean	Std dev	Max	Mean	Std dev	Max	Mean	Max
Aggregation	0.95	0.01	0.97	0.95	0.01	0.97	–	**0.99**
Compound	0.75	0.00	0.75	0.89	0.00	0.89	–	**0.96**
Pathbased	0.8	0.00	0.8	0.55	0.02	0.6	–	**0.92**
Spiral	**1.00**	0.00	1.00	**1.00**	0.00	1.00	–	**1.00**
Mixed	**1.00**	0.00	1.00	**1.00**	0.00	1.00	–	**1.00**
Toy	0.36	0.00	0.36	**1.00**	0.00	1.00	–	**1.00**
Flame	0.88	0.01	0.9	0.92	0.01	0.94	–	**0.93**
R15	**0.98**	0.00	0.98	**0.98**	0.00	0.98	–	0.96
Soya	0.98	0.00	1.00	0.98	0.00	1.00	–	**1.00**
Iris	0.36	0.01	0.37	0.42	0.01	0.44	–	**0.45**

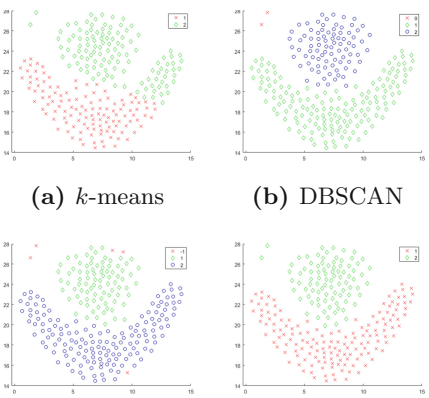
fridaya@uef.fi

Table 4. Experiments comparing the run time in seconds of DBSCAN, ISDBSCAN, and DBSCRN. The below includes the time-elapse for DBCV.

	DBSCAN				ISDBSCAN				DBSCRN			
	Mean	Std dev	Max	Min	Mean	Std dev	Max	Min	Mean	Std dev	Max	Min
Aggregation	0.0237	0.0065	0.0374	0.0125	1.2481	0.0073	1.2655	1.2307	0.1131	0.0004	0.1149	0.1126
Compound	0.0071	0.0017	0.0101	0.0037	0.6043	0.0200	0.6819	0.5964	0.0490	0.0004	0.0508	0.0485
Flame	0.0032	0.0005	0.0065	0.0018	0.3458	0.0121	0.3709	0.3377	0.0194	0.0001	0.0201	0.0192
Iris	0.0018	0.0001	0.0036	0.0016	0.2221	0.0006	0.2247	0.2217	0.0130	0.0001	0.0135	0.0128
Mixed	0.0442	0.0037	0.0637	0.0349	2.4784	0.0019	2.4820	2.4759	0.3075	0.0053	0.3150	0.3034
Pathbased	0.0042	0.0010	0.0062	0.0022	0.4474	0.0003	0.4480	0.4467	0.0260	0.0001	0.0263	0.0258
R15	0.0177	0.0029	0.0217	0.0093	0.9157	0.0005	0.9167	0.9149	0.1948	0.0006	0.1986	0.1940
Soya	0.0005	0.0001	0.0012	0.0003	0.0747	0.0001	0.0750	0.0745	0.0042	0.0000	0.0044	0.0042
Spiral	0.0029	0.0004	0.0044	0.0025	0.2983	0.0012	0.3010	0.2966	0.0169	0.0001	0.0174	0.0168
Toy	0.0058	0.0014	0.0085	0.0032	0.5616	0.0006	0.5624	0.5605	0.0488	0.0002	0.0497	0.0485

algorithm were those optimising the DBCV index. This is a truly unsupervised scenario. Table 3 presents the results in terms of cluster recovery. This time we decided not to run experiments with k-means because we have empirically demonstrated this is not well-aligned with the type of data sets we experiment with, and because DBCV was designed to be used by density-based algorithms. The experiments clearly demonstrate that in all cases DBSCRN is competitive or better than the other density-based algorithms we experiment with.

In our experiments we have shown that ISDBSCAN outperforms DBSCAN in terms of cluster recovery, and that DBSCRN outperforms both of them in the same measure. Table 4 summarises the running time for each algorithm in seconds. This table includes the computational time required to run DBCV. We can clearly see DBSCAN is the fastest algorithm we experiment with. However,

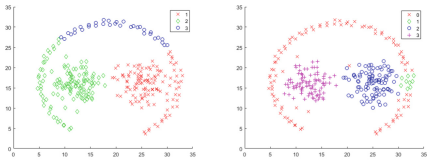


(a) k-means **(b)** DBSCAN

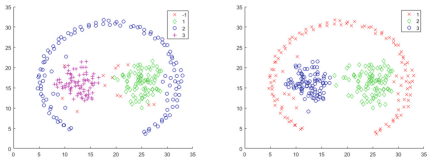
(c) ISDBSCAN **(d)** DBSCRN

Fig. 3. Best possible cluster recovery as measured by the ARI on the Flame data set.

fridaya@uef.fi

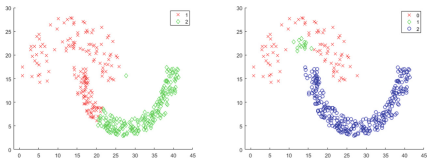


(a) k-means **(b)** DBSCAN

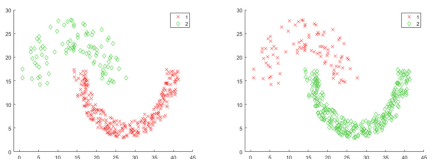


(c) ISDBSCAN **(d)** DBSCRN

Fig. 4. Best possible cluster recovery as measured by the ARI on the Pathbased data set.



(a) k-means **(b)** DBSCAN



(c) ISDBSCAN **(d)** DBSCRN

Fig. 5. Best possible cluster recovery as measured by the ARI on the Toy data set.

DBSCRN outperforms DBSCAN and ISDBSCAN in terms of cluster recovery, and it is orders of magnitude faster than the latter.

7 Conclusion

In this paper we have introduced a new density-based clustering algorithm, Density-based spatial clustering using reverse nearest neighbour (DBSCRN). We have run a number of experiments clearly showing our algorithm to outperform DBSCAN and ISDBSCAN in terms of cluster recovery. These experiments also established that we can indeed estimate a good parameter for DBSCRN which

fridaya@uef.fi

leads to better cluster recovery than that of other algorithms in a truly unsupervised scenario. Our experiments also show DBSCRN to be orders of magnitude faster than ISDBSCAN.

The experiments have also shown k-means not to perform well in most cases. Given the data sets we experiment with, this is hardly surprising. These results should not lead to conclusion that k-means is inferior to density-based algorithms, but rather that one should pay considerable attention when selecting a clustering algorithm.

In our future research we intend to establish whether DBCV is indeed the best CVI to use in our case, and whether we can introduce the concept of feature weights to our method. These feature weights should model the degree of relevance of each feature in the data set.

References

1. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Comput. Surv. **31**(3), 264–323 (1999)
2. Hou, J., Gao, H., Li, X.: DSets-DBSCAN: a parameter-free clustering algorithm. IEEE Trans. Image Process. **25**(7), 3182–3193 (2016)
3. Jain, A.K.: Data clustering: 50 years beyond k-means. Pattern Recogn. Lett. **31**(8), 651–666 (2010). Award winning papers from the 19th International Conference on Pattern Recognition (ICPR)
4. Mirkin, B.: Clustering: A Data Recovery Approach. CRC Press, Boca Raton (2012)
5. Hennig, C.: What are the true clusters? Pattern Recogn. Lett. **64**, 53–62 (2015)
6. Ester, M., Kriegel, H.-P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD 1996, pp. 226–231. AAAI Press (1996)
7. Ankerst, M., Breunig, M.M., Kriegel, H.-P., Sander, J.: OPTICS: ordering points to identify the clustering structure. In: ACM Sigmod Record, vol. 28, pp. 49–60. ACM (1999)
8. Hinneburg, A., Keim, D.A., et al.: An efficient approach to clustering in large multimedia databases with noise. In: KDD, vol. 98, pp. 58–65 (1998)
9. Jin, W., Tung, A.K.H., Han, J., Wang, W.: Ranking outliers using symmetric neighborhood relationship. In: PAKDD, vol. 6, pp. 577–593. Springer (2006)
10. Cassisi, C., Ferro, A., Giugno, R., Pigola, G., Pulvirenti, A.: Enhancing density-based clustering: parameter reduction and outlier detection. Inf. Syst. **38**(3), 317–330 (2013)
11. MacQueen, J., et al.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Oakland, CA, USA, vol. 1, pp. 281–297 (1967)
12. Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. Am. Stat. **46**(3), 175–185 (1992)
13. Korn, F., Muthukrishnan, S.: Influence sets based on reverse nearest neighbor queries. SIGMOD Rec. **29**(2), 201–212 (2000)
14. Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J.M., Perona, I.: An extensive comparative study of cluster validity indices. Pattern Recogn. **46**(1), 243–256 (2013)

fridaya@uef.fi

15. Moulavi, D., Jaskowiak, P.A., Campello, R.J.G.B., Zimek, A., Sander, J.: Density-based clustering validation. In: Proceedings of the 2014 SIAM International Conference on Data Mining, Philadelphia, Pennsylvania, USA, 24–26 April 2014, pp. 839–847 (2014)
16. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. **20**, 53–65 (1987)
17. Caliński, T., Harabasz, J.: A dendrite method for cluster analysis. Commun. Stat. Theory Methods **3**(1), 1–27 (1974)
18. Dunn, J.C.: Well-separated clusters and optimal fuzzy partitions. J. Cybern. **4**(1), 95–104 (1974)
19. Maulik, U., Bandyopadhyay, S.: Performance evaluation of some clustering algorithms and validity indices. IEEE Trans. Pattern Anal. Mach. Intell. **24**(12), 1650–1654 (2002)
20. Halkidi, M., Vazirgiannis, M.: A density-based cluster validity approach using multi-representatives. Pattern Recogn. Lett. **29**(6), 773–786 (2008)
21. Bache, K., Lichman, M.: UCI machine learning repository (2013)
22. Limin, F., Medico, E.: Flame, a novel fuzzy clustering method for the analysis of dna microarray data. BMC Bioinformatics **8**(1), 3 (2007)
23. Jain, A.K., Law, M.H.C.: Data clustering: a user's dilemma. PReMI **3776**, 1–10 (2005)
24. Veenman, C.J., Reinders, M.J.T., Backer, E.: A maximum variance cluster algorithm. IEEE Trans. Pattern Anal. Mach. Intell. **24**(9), 1273–1280 (2002)
25. Chang, H., Yeung, D.-Y.: Robust path-based spectral clustering. Pattern Recogn. **41**(1), 191–203 (2008)
26. Zahn, C.T.: Graph-theoretical methods for detecting and describing gestalt clusters. IEEE Trans. Comput. **100**(1), 68–86 (1971)
27. Gionis, A., Mannila, H., Tsaparas, P.: Clustering aggregation. ACM Trans. Knowl. Discov. Data (TKDD) **1**(1), 4 (2007)
28. Fisher, R.A.: The use of multiple measurements in taxonomic problems. Ann. Eugenics **7**(2), 179–188 (1936)
29. Tan, M., Eshelman, L.: Using weighted networks to represent classification knowledge in noisy domains. In: Proceedings of the Fifth International Conference on Machine Learning, pp. 121–134 (1988)
30. Fisher, D.H., Schlimmer, J.C.: Concept simplification and prediction accuracy. In: Proceedings of the Fifth International Conference on Machine Learning, pp. 22–28 (2014)
31. Hubert, L., Arabie, P.: Comparing partitions. J. Classif. **2**(1), 193–218 (1985)

fridaya@uef.fi

We Know What You Did Last Sonar: Inferring Preference in Music from Mobility Data

José C. Carrasco-Jiménez(✉), Fernando M. Cucchietti, Artur Garcia-Saez, Guillermo Marin, and Luz Calvo

Barcelona Supercomputing Center, Barcelona, Spain
jose.carrasco@bsc.es
https://www.bsc.es/viz/whatyoudid

Abstract. Digital music platforms allow us to collect feedback from users in the form of ratings. This type of information is *explicit* about the users musical preferences. In contrast, *intrinsic feedback* provides contextual information from which preferences can be inferred, some quite obvious like amount of playing activity, playlists, and others less direct like activity in or components of social networks. Here we focus on physical *intrinsic feedback* in the form of mobility traces on a music festival with multiple stages, and analyze it to infer music preferences. To the best of our knowledge, this is the first research work that exploits physical contextual clues and human mobility behavior from WiFi traces to approximate ratings that are later used to: (1) measure musical similarity among musical bands, and (2) estimate the effect of loyalty of a physical audience (i.e. going further than measuring sheer number of attendees). As part of this work, we developed a novel metric to measure weighted user rating from the mobility patterns of individuals during a music festival, incorporating physical contextual data in addition to factors commonly used in ranking systems. The experiments reveal groups of people with similar musical preference, and adjusted rankings that identify acts that, even if small, were highly successful in terms of their effect on the audience.

Keywords: Social computing · Social behavior · Collective intelligence · Music similarity

1 Introduction

Understanding human interaction and behavior has led to a number of benefits in different fields, a better insight into the needs of users and, as a consequence, more personalized access to services [4,9,14,21,23]. In this paper, we use mobility data and analyze human behavior to infer music preferences, which may lead to offer personalized recommendations of music during live music festivals. We focus on music festivals with multiple stages, which provide opportunities for

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 43–61, 2019.
https://doi.org/10.1007/978-3-030-22868-2_4

fridaya@uef.fi

the audience to choose from dissimilar artists and styles and allow us to collect physical mobility data that includes musical taste information.

Our objective is two-fold. First, to use WiFi traces and mobility information to extract musical and user preference information, and second, to use mobility information to measure the performance of artists and their relation with the audience.

Unlike the much more extensively researched explicit feedback [12], implicit feedback is harder to extract. In our case, we interpret music preference for each user from the mobility patterns exhibited during the festival. This is analogous to the way in which most music recommender systems work, where listening habits are analyzed to extract music preferences. Although there are other approaches in which implicit feedback is collected from the context, most of the research relies on contextual information extracted from digital media such as online social networking websites [20].

In our work, a number of factors were carefully crafted in order to estimate user rating from mobility patterns. We propose a network-based approach to compute similarity between the audiences of artists, and from this, we infer user music preferences. Our approach differs from the typical content-based approach, which relies on the information contained either in the sound waves or song lyrics, extracting physical clues to uncover related bands.

Lastly, we devise a metric to measure the performance of musical bands using the behavior of the audience instead of measuring just the size of the audience. The metric measures the loyalty of the audience, with which we can construct a ranking value based on the mobility of the attendees. It is very common to rank popularity of music bands based on number of sales, and in the case of live music by the size of the audience. Our findings indicate that loyalty leads to different rankings and can provide useful insight, incorporating information like the significance of each band for each user (calculated from the behavior of each individual), the size of hall in which the band performed, and the time allocated to each performance.

The remainder of this paper is organized as follows. Section 2 describes related work. Section 3 details the methodology, including the experimental settings, datasets, and some challenges encountered during the analysis. Section 4 describes some insights about audience behavior inferred from the analysis of mobility patterns. Section 5 elaborates on the estimation of weighted ratings from mobility patterns exhibited by attendees of the music festival. Furthermore, Sects. 6 and 7 describe a method to discover similar bands from musical preferences and the effect of loyalty in the ranking, respectively. We conclude with a discussion of the most notable findings and future work in Sect. 8.

2 Related Work

Several ways for exploiting music features in the field of music information retrieval (MIR) and music similarity have been introduced. Initially defined to recommend songs or artists, music features have served other purposes such as

fridaya@uef.fi

uncovering listening patterns, inferring artist similarity and ranking artists [25]. There is interesting research in MIR especially in artist similarity computation, with different approaches to this problem in which alternative types of features are exploited to compute music/artist similarity. On one hand, acoustic features, also known as content-based features, are used to identify similar bands based on the content of the acoustic wave [8]. In this regard, [10] and [11] suggest two different methods to compute music similarity based on chord progression. In addition to chord progression analysis, tempo and timbre features [27] are also employed to classify artists by describing music genre and emotion. Panteli et al. developed a singing style similarity based on pitch structure and melodic embellishments [19], which uncover similarity based on language and culture proximity.

Although acoustic-based features offer a good approximation to artist similarity computation, other sources of information have been explored in recent years. Text-based features, including knowledge originating from web pages, user tags and song lyrics [13] have been studied, leading to different levels of accuracy. Although song lyrics are typically a major source of information for identifying similar songs/artists, the analysis of lyrics by itself has exhibited worst similarity results compared to those obtained from acoustic-based features [13]. Other approaches include the analysis of collaborative tags. In their work, Knees and Schedl [13], compute artist similarity based on overlapping tags, while Lin et al. [17] combine user tags with other attributes such as music genre, timbre, and era. Both authors assume that the more characteristics two artists share, the more similar they are.

In other works such as the one described in [20], the authors establish a different approach to artist similarity in which Twitter hashtags are used as the source of information of what users are listening to. In their methodology, Schedl and Hauger, estimate artist similarity by computing a co-occurrence matrix where two artists are said to co-occur if they are in the play-list of the same user. Such information is extracted from Twitter using the hashtags *#nowplaying* and *#itunes*.

Furthermore, some works have studied the effect of combining both, text-based and content-based features [16], but there is still a lack of ways to exploit physical contextual information in order to estimate preferences in music and artist similarity.

3 Methods

In this section we describe the methodology proposed to reach our objectives, described in Sect. 1. We start by defining the steps required to extract contextual information from mobility traces collected in the form of WiFi fingerprints. There is a description of the dataset collected during the Sonar festival, as well as the challenges encountered during the study.

fridaya@uef.fi

3.1 Sonar Dataset Description

The Sonar dataset offers information about mobility exhibited by attendees to the Sonar festival that took place from the 18th to the 20th of June, 2015 in Barcelona, Spain. The Sonar festival is a multistage event with more than 100,000 attendees in two main venues, Sonar by Day and Sonar by Night. We only collected information in the day part from 12:00 to 22:00 in the four simultaneous stages. We also collected but did not analyze data in an art installation space (SonarPlanta) and a collocated technology congress, Sonar+D (see the plan of the location in Fig. 1). The Village stage (main stage) is surrounded by food trucks, merchandising stores, benches, and other non-music related attractions.

The dataset comprises a total of about six million WiFi events captured by sensors strategically distributed to cover the main halls of the festival, as shown in Fig. 1.

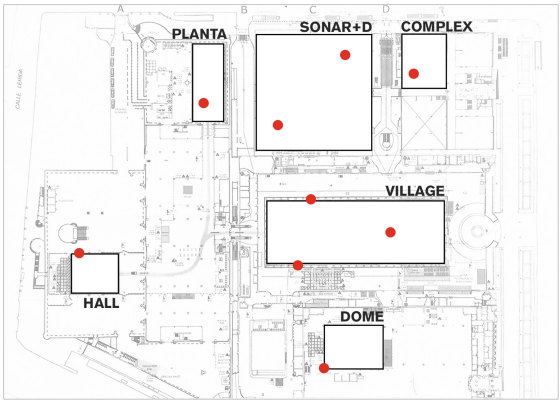


Fig. 1. Plan of the Sonar by Day location. The circles mark the position of the Raspberry Pi 2 computers strategically distributed to cover the main halls.

Each WiFi event collected (summarized in Table 1) contains enough information to locate and track users and capture their behavior during the festival.

Table 1. Summary of data collected for each WiFi event.

Data	Description	Example
MAC Address	Unique id	15015608320
Timestamp	Data recording time	2015-06-18 14:25:32
Hall	User location	Sonar Village
RSS	Received signal strength	−76

fridaya@uef.fi

In order to determine the location of a user at a given time, we associated them to the WiFi scanning device that exhibited the strongest signal strength. We detected a large number of spoof MAC addresses sent by modern smart phones [2] in an attempt to prevent privacy (recognized because their Organizationally Unique Identifier (OUI) does not correspond to any registered vendor [1]). For analysis we discarded all events with a spoof MAC address.

The event timestamps were grouped in time intervals of 5 min. About 46,000 unique users were identified. After filtering outlier behavior (detected by Tukey's method) we kept about 17,000 unique users who spent more than 65 min or less than 687 min in the music festival. The outliers were mostly users observed only for a few minutes or consistently in the same place for hours, perhaps people walking nearby the festival venue or devices used by the organization respectively.

3.2 Data Processing

Data was gathered using RPi2 computers with two WiFi antennas, one for communication with our servers and the other one in monitor mode, sniffing activity over the WiFi network and collecting their MAC addresses and signal strength. Measurements were taken by the RPi2 devices every 50 ms, with all detected devices stored in a local table. Every 30 s the table was grouped by MAC address (keeping the largest observed signal for each device), sent to our servers, and flushed to prepare for the next 30 s window. With this strategy we sought to minimize the noise introduced by packet collision in crowded rooms, which prevented us from detecting all devices on a single measurement.

Collection of the data was done using an Apache Tomcat server that provided an API to store data directly in a distributed Cassandra database, and to obtain live information of the status of all the RPi2 devices. After the festival, the data was dumped into an MySQL database and cleaned as described above. Furthermore, device trajectories that were detected simultaneously by multiple devices were smoothed out. The assignment was done using reasonable space/time constrains (identifying impossibly fast translations from one hall to the other) and the running average of the strength of the detected signal.

Data preparation was performed using Apache Spark [26]. This part of the process consisted of a number of descriptive analysis, outlier removal, arrange of data into suitable data structures (e.g. graphs, visits per user, etc.), as well as computation of implicit rating scores. Lastly, graph analysis (see Sect. 6) was performed using the *Louvain* method implemented in Gephi [3] with randomization and different *resolution limits*.

3.3 Data Challenges

The dataset contains a considerable amount of noise due, for example, to the proximity of sensors to the street detecting devices outside the festival. We also identified over one hundred thousand short lived (less than 20 min) spoof mac

fridaya@uef.fi

addresses, indicating that device protection strategies might include randomization of the fake address. Devices were also either turned off (or at least their WiFi connectivity) intermittently, leading to a loss of information in the form of sporadic or spotty signal. Lastly, the scanning devices frequently lost connection to the server or were switched off by external factors.

4 Understanding Global Audience Behavior

Public assistance to a festival which lasts 12 h for 3 consecutive days changes constantly, and data analysis of the public behavior has to address this diversity. A first approach to understanding audience behavior is measuring the audience size on each location of the festival, which can be directly obtained from the main dataset. In particular we look for metrics that describe the dynamics of the audience, and a clear picture of the impact of each show in the festival. Furthermore, we wish to compensate for artists playing in an unfavorable times or in smaller locations.

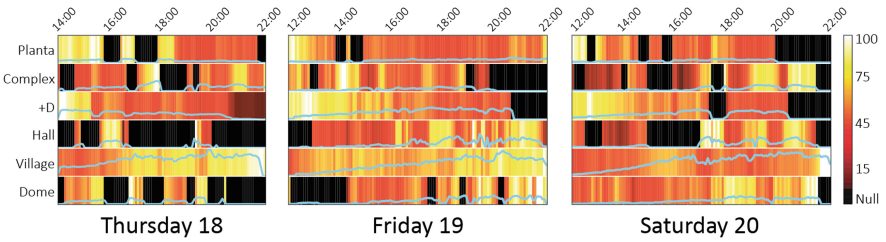


Fig. 2. Global impact of each show along the 3 days of the festival. The leftmost figure is the activity of day one of the festival, followed by days two and three respectively. In this measure, we compensate for the global audience and the maximum capacity of each venue. As a result, the color code reveals bands with a high impact as measured by the relative audience.

We propose to measure the impact of each show by normalizing the size of their audience by the total number of persons at Sonar at every moment. As venues are different, we also normalize each venue audience by the maximum observed on this venue throughout the festival. Thus, we normalize on a temporal axis globally, and locally for each venue. As a result, we can compare the impact that each show has at every moment on the festival program.

Figure 2 shows how some bands playing at different venues and at different times may have a similar impact. It is worth mentioning how these plots also reveal the different dynamics of each venue, as some of these may have a constant impact while other have a strong dependence on whoever is playing at each moment. Thus, our metric provides a fine grain view of the impact that each band has on the program of the festival, and is a useful tool to easily detect remarkable shows at each moment or venue.

fridaya@uef.fi

5 Computing Weighted Ratings

In this section we develop a novel metric to measure the Weighted User Rating (WUR) assigned to musical bands from the mobility patterns of individuals during the festival. The traditional methods to compute rating rely on the users assigning a value of how much he/she likes an artist. Instead, our approach takes into consideration the concerts a user visited, as well as the overall behavior of the individual during the festival.

If we define the user rating as a measurement of the degree of significance (i.e. importance) a person gives to an artist, we can suggest that the user rating be approximated as the amount of time a user spent in a given artists' show. In this case, the rating value assigned by user u to band b can be computed as

$$ur_{u,b} = visits_{u,b}, \tag{1}$$

where $visits_{u,b}$ is the amount of time a user u spent at the concert of artist or band b. The assumption in this simple metric suggests that users spend more time in concerts they like more, and thus the rating should be higher. This approach is very intuitive but at the same time it may be misleading as it does not take into account the amount of time a user spent at the festival as a whole, and neither their overall behavior.

An even better estimate would be to take the proportion of time a user invested in a concert relative to his/her own mobility patterns, which we measure as

$$wur_{u,b} = \frac{ur_{u,b}}{\sum_{b \in B} ur_{u,b}} = \frac{visits_{u,b}}{\sum_{b \in B} visits_{u,b}} \tag{2}$$

where $visits_{u,b}$ is the amount of time a user u spent at the concert of band b, computed in Eq. (1), and $\sum_{b \in B} visits_{u,b}$ is the amount of time the user spent at the festival where B bands played.

The rating value computed from Eq. (2) captures a more realistic approximation of how much a user u likes a band b, as it incorporates some form of comparison to how the user fared other bands, and the total amount of time a user stayed at the festival. For example, a person that attends only one artist show for one hour would give this artist ten times more rating than another person that also attended nine other one-hour concerts.

The rating estimated using Eq. (2) improves over the estimation from Eq. (1), but this approximation is still incomplete as it does not take into account a phenomenon we call *the village effect*, something that might be common in multistage festivals with common areas. It refers to the formation of groups of people that stay in the same place (primarily the main venue) for long periods of time, regardless of the performance taking place there. Figure 3 shows that bands that perform in the main hall (where we can clearly identify this phenomenon) tend to have higher ratings than those that performed in other scenarios.

In order to fix this bias, we add a penalization that punishes the opinion of those users who had little mobility during the festival. In other words, we seek to increase the impact of the rating value assigned by users who exhibit higher

fridaya@uef.fi

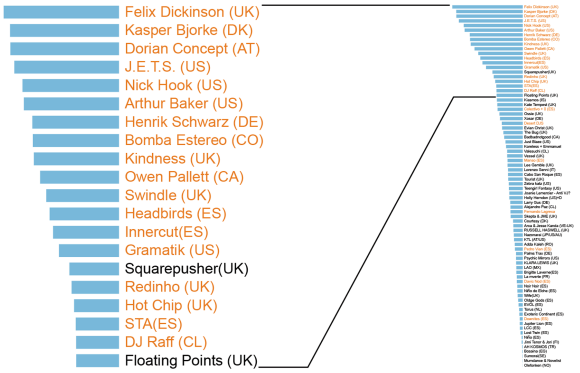


Fig. 3. Names highlighted in red are the bands performing in the main hall. Bands that perform in the main scenario tend to have higher ratings than those that performed in other scenarios. This phenomenon, which we call *village effect*, is common in multistage festivals.

mobility during the festival, as those are more likely to have attended a concert on purpose rather than just hang out at the same space, making their opinion more valuable. If we add a penalization factor to Eq. (2) we obtain

$$wurp_{u,b} = \frac{1}{\sum_{\tilde{b} \in r(b)} visits_{u,\tilde{b}}} \left(\frac{visits_{u,b}}{\sum_{b \in B} visits_{u,b}} \right) \tag{3}$$

where $\sum_{\tilde{b} \in r(b)} visits_{u,\tilde{b}}$ is the total amount of time a user u spent in the room r where band b performed. Users who overall spent more time in a single room will have a higher penalization than users who moved more often and thus spent little time in the same room throughout the event.

Equation (3) expands on the intuitive approach to computing user rating from implicit feedback. The new pieces of information reflect the need to understand the sometimes underestimated factors that impact the estimated rating value. In this case, and as it was noted, our work focuses on implicit feedback collected in the form of mobility patterns, but the proposed metric may be modified to match other types of sources of information including those that focus on listening habits (also called online streaming) [6,13]. As it will be shown in Sect. 7, Eq. (3) will be expanded even further to incorporate more contextual information in order to adapt it to understand the effect of loyalty over size of audience in the ranking of musical bands.

6 Discovering Similarity from Mobility Data

Another contribution of this work is the discovery of similarity and relatedness from musical preferences. In this case, we determine musical preference from the behavior of users during the festival using the metric explained in Sect. 5.

fridaya@uef.fi

In this section, we propose an algorithm to discover related bands using a network-based approach. The methodology to construct a network of musical bands and discover groups of related ones consists of four steps:

1. Create a graph where each node represents a musical band.
2. Create a link between a pair of bands that have assistants in common.
3. Update the weight of a link between a pair of bands using the Jaccard Index or Weighted Jaccard Index.
4. Apply community detection to identify groups of related bands.

The critical step of the methodology proposed is the updating of the link weights, as the communities discovered depend on the weights of the links between bands. Recall that the weight of links between two bands can be computed from the number of assistants in common or a combination of the users and the ratings assigned by each user to the band. In this regard, we defined the first experimental setup where we assume that the more visitors two musical bands have in common, the more similar they are. This approach is identical to other works [6,13,17,20] where there is an assumption that two songs are said to be highly similar if both songs appear in the play list of the same user or shared more characteristics in common, either content-based or context-based features. This is a strong assumption, as people certainly listen to a broad range of music styles. Technically, we are discovering only sets of bands that are preferred by the same people. After contrasting with contextual information about the particular bands at Sonar, we will show that these sets are similar either in musical style or in the type of show they offer.

6.1 Discovering Similarity from Assistants in Common

In this first approach we study the effect of incorporating only the number of assistants that two bands have in common to compute the similarity coefficient, that is assigned as the weight of link, between the musical bands. We compute the similarity coefficient (J_μ) between bands b_i and b_j using the Jaccard Similarity Coefficient, defined by Eq. (4):

$$J_\mu(\beta_i, \beta_j) = \frac{|\beta_i \cap \beta_j|}{|\beta_i \cup \beta_j|}, \tag{4}$$

where β_i is the set of all the assistants to the concert of musical band i, and β_j is the set of all the assistants to the concert of musical band j.

In this approach, we consider an individual to be part of the set of assistants to a concert of a musical band if the user was observed in the concert at least once. In order to segment musical bands into clusters, we used the Louvain method for community detection [5] to reveal the intrinsic network structure. Our assumption is that the communities identified reveal artist similarity, with similar bands belonging to the same cluster.

We observed many links between shows coming from individuals attending only for a few minutes and then moving on to another venue, which introduced

fridaya@uef.fi

some noise in the form of many more links than expected. We decided to keep only valuable attendance events by setting a minimum visit length threshold, affected by the average amount of time required for an individual to transfer from one hall to another. Our data suggested that 15 min was a good threshold to remove short term visit links from the graph.

We were able to identify three different groups of musical bands, shown in Fig. 4. However, the clusters obtained from this segmentation method exhibits groups separated by the day in which they performed in the festival. In other words, our direct use of Eq. (4) was strongly influenced by the time table and offered no information about preference of the audience.

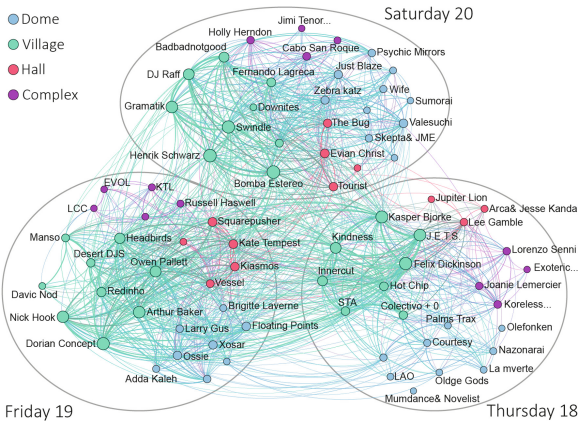


Fig. 4. The clusters obtained from the segmentation method exhibit bands separated by the day in which they performed in the festival.

The distribution of assistance to the festival according to the number of days they were observed is shown in Table 2, where it is clear that about 50% of the audience attended only one of the three days. We suspected that the large contribution of opinions of people attending only one day would induce a bias in the clustering analysis discussed above. We thus repeated the study with the same settings but removing, first, the users that attended the festival only one out of the three days, and then, the users that attended at most two days. However, the segmentation of similar bands remained similar to the festival schedule.

In order to improve on the capability of finding artist similarity we need to incorporate a new piece of information, as discussed in the next section.

6.2 Discovering Similar Bands from Assistants in Common and Ratings (Weighted Jaccard Index)

We set out to complement the physical presence data that establishes connections between artists with the value or rating (computed implicitly from mobility data)

fridaya@uef.fi

Table 2. Distribution of assistance by number of days. The festival lasted three days.

Threshold value	# of assistants	% of assistants
1 day	8368	≈50
2 days	4597	≈27
3 days	3856	≈23

that each connecting person assigns to those artists. For this, we incorporated not only the number of assistants in common to our new metric, but also the rating given by each assistant to the musical band (see Eq. (3)).

The experimental setup is similar to the one described in Sect. 6.1 except that we modify the metric to compute the similarity between a pair of musical bands, which is assigned to the weight of their link. The new metric adds new contextual information to the similarity coefficient, and is given by Eq. (5):

$$J\omega_{\mu}(\beta_i, \beta_j) = \frac{\sum_k min(\beta_{i,k}, \beta_{j,k})}{\sum_k max(\beta_{i,k}, \beta_{j,k})} \tag{5}$$

where $J\omega_{\mu}$ is the Weighted Jaccard Coefficient, which is defined as the ratio of the corresponding sums of weights, $\beta_{i,k}$ and $\beta_{j,k}$ are the weights (i.e. ratings) assigned by user k to each band in set β_i and β_j, respectively.

In general, Eq. (5) assigns a weight to a link between two musical bands based on the assistants in common, but integrates the rating each assistant assigns to the musical band. In Eq. (4), a band could have a generous weight even if the users exhibited dislike to a concert by entering and leaving right away. In this regard, Eq. (5) regulates the weight of the link by adding the rating given by each assistant in common, or in other words, the number of assistants weighted by the loyalty to the musical band, as we will see in Sect. 7.

As mentioned above we kept the visit-length threshold of 15 min and only attendees that were observed for the full three days of the festival (23% of detected users, or about four thousand attendees).

Under these considerations, we identified three clusters of artists that expose preferences of music of the assistants. The results are shown in Fig. 5.

We now want to check what insights this segmentation offers about different types of audiences that attended the music festival, and if it is related to the similarity between artists. The first group lies mostly within the bands that played at the Sonar Dome room, which is a space dedicated to electronic DJ type of show (that is, a strong similarity in musical style). The second group had a heavy weight on the Sonar Complex room, a seated hall with audiovisual shows of experimental and avant grade artists (possibly different musical styles but with a large component of novelty and innovation). This group was also detected in the main stages but at early times, when the lesser known bands performed (coinciding with the novelty component). Lastly, group number three was found in the larger stages Sonar Village and Sonar Hall, but predominantly in the evening where the main artists and bands performed (again, no a priori musical similarity but safer,

fridaya@uef.fi

more popular music). The implications are that group 1 preferred DJ style electronic music, group 2 sought new music, and group 3 (arriving late each day) was attracted to mainstream artists (at least in the frame of reference of the festival). This is one of the two main results of this paper.

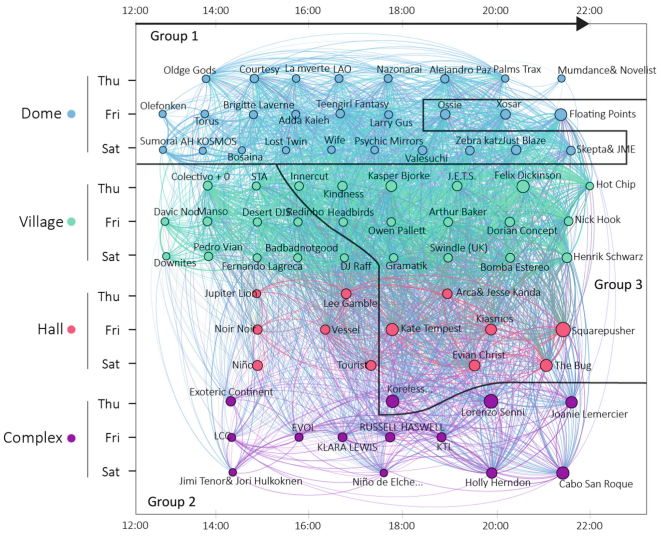


Fig. 5. Music bands separated by room (color), day (vertical axis), and performance time (horizontal), and linked by shared audience (width of lines). The three clusters detected here correspond to (1) acts in the Sonar Dome, with a homogeneous music style, (2) acts in Sonar Complex, where the most experimental music of the festival was showcased (and early shows in the main stages, with lesser known artists), and (3) bands in the main stages late in the evening, when the most well known bands played. This correlates well with three target audiences of the festival: Hardcore DJ music fans, searchers of novelty, and audiences of more mainstream artists.

The results obtained from this part of the research highlight an important insight that exhibits the relationship between the size of audience, analogous to the number of fans that have the same artist in their play-lists, and the loyalty of the audience. As we discuss below, we used domain experts to confirm that two bands have a higher degree of similarity when they have more visitors in common and the rating values assigned by those users is also higher. Furthermore, we will see in Sect. 7 how the size of the audience is not as relevant as the ratings provided by their audience.

6.3 Evaluation

Before describing the methodology followed to evaluate the clusters of bands, it is worth mentioning some intrinsic characteristics of the type of data we are using to uncover similarity of musical bands. Among other things we need to consider that

fridaya@uef.fi

1. the festival focuses on experimental music and thus some musical bands are not well known to the general public, making it harder to classify;
2. the weights of the links correspond to rankings that are computed based on the physical behavior of the people, also called *implicit feedback* (differing from the typical rating systems, i.e. *explicit feedback*, typically done when rating stream-radio like Yahoo Music, Spotify, etc.);
3. we have partial ground-truth, a problem that is common to most real-world datasets.

Evaluation of clustering methods on graphs, also called community detection algorithms, can be done following different criteria. The interpretation of results is highly dependent on the domain of the problem at hand and the particular objectives of the community detection algorithm. Yang et al. [24] discuss a number of scoring functions, i.e. functions that measure the quality of the communities that reveal the internal structure of the communities. As described in their work, scoring functions are naturally grouped in four different categories that measure different characteristics of the structure of the communities. The general categories include (1) internal connectivity, (2) external connectivity, (3) combination of internal and external connectivity, and (4) modularity. It is worth mentioning that community detection is commonly done by identifying segments based on the structure of the network followed by the evaluation of communities based on their function [24]. In this case, *function* refers to a common role, affiliation or attribute that two nodes share in common.

Several quantitative measures for validating the output of community detection algorithms exist. Chen et al. indicate that *modularity* is an effective measure of the strength of the community structure found by network-based clustering methods (also known as community detection algorithms) [7]. Although it is widely used, it relies strictly on the network structure. It is specifically designed for the purpose of measuring quality of a network partition into communities. In the absence of complete ground-truth, *modularity* allows us to estimate a partition of nodes as it is shown in [22]. In our case, we obtain a *modularity* value of 0.43 (consider that real-world networks typically range from 0.3 to 0.7 [18]). Steinhaeuser cautions against the use of modularity as the only criterion for the evaluation and comparison of community detection algorithms since it serves more as a descriptive measure of data than a 'true performance metric' [22]. Furthermore, Lee et al. [15], warns about the use of modularity as an evaluation criteria since it may lead to a "rough sketch of some parts of a network's modular organization." Following this advice, we seek to incorporate other metrics in order to obtain another perspective on the communities detected.

As explained by Steinhaeuser et al., a combination of metrics evaluating both the structure and the agreement with real communities provides a better performance metric. When ground-truth is available a different set of validation metrics is proposed. Consider for example Adjusted Rand Index (ARI), described in [22], which captures the extent to which the real and the computed partitions agree with one another. An $AdjustedRandIndex$ value of 1.0 indicates that the computed clusters and the ground-truth are identical and values close to 0.0

fridaya@uef.fi

indicate little agreement between computed clusters and ground-truth. In our work, we have partial ground-truth since, as it was previously explained, the festival focuses on experimental music (some of the bands are unknown to the general public and as a consequence hard to classify). The ground-truth was obtained from human experts, in this case, the organizers of the Sonar festival, with good level of inter-judge agreement. Human experts also evaluated the resulting clusters based on their knowledge of the domain. Using partial ground-truth, we obtained an *ARI* value of 0.625, which can be partially interpreted since it was computed on partial ground-truth. It is worth mentioning that only 40 out of the 78 artists that performed during the festival were classified. The rest of the musical bands were either experimental or simply harder to classify. As a consequence, an *ARI* score of 0.625 does not give us a full perspective of the correctness of the communities.

In order to improve the efficiency of our evaluation, we incorporate a metric of robustness that measures how many nodes remain together as we change the *resolution limit* parameter. We consider different resolution limits of the Louvain method (0.8, 0.9, 1.0, 1.1, 1.2) with randomization to produce a better decomposition. After running the community detection algorithm [5] to find similar musical bands with the different parameters, we found that the underlying structure remains unchanged for different resolution parameters, as shown in Fig. 6.

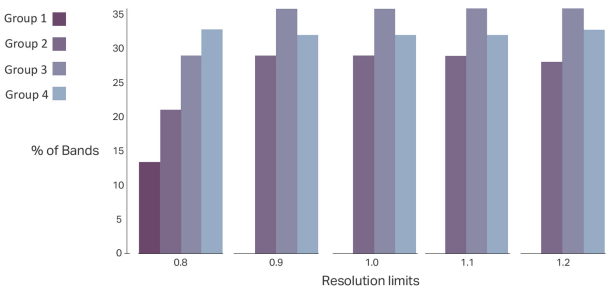


Fig. 6. The underlying network structure remains unchanged for different resolution limits of the Louvain method. Randomization was used to produce a better decomposition.

As we can see, the underlying structure of networks of bands is pretty robust as we modify the parameters. The variations are too small to affect the underlying structure of the clusters determined by the community detection algorithm. The nature of the Louvain method for community detection allows us to conclude that groups found by the clustering algorithm are in fact the same groups with different resolution parameters. Let us recall that the Louvain method for community detection works by merging partitions that maximize the modularity, thus, changing the resolution parameter value is like zooming into the network. Different resolution values result in partitions composed of merging one or more

fridaya@uef.fi

partitions from communities with lower resolutions. As a consequence, we can deduce that the structure of the communities (i.e. clusters) becomes stabilized, showing a more accurate segmentation of similar musical bands.

7 The Effect of Loyalty in the Ranking

In this section, we study the effect of loyalty vs size of audience to estimate the popularity of musical bands that performed during the festival. We start by the naive approach which suggests that more popular bands attract more people to their concerts (preferential attachment). We can take the average number of visits each band received as the overall popularity of the artist. However, this metric would make sense if all bands were allocated the same amount of performance time, at similarly popular times, and that all halls had similar capacity, assumptions that are not true in the kind of festival we studied.

As shown in Fig. 7, ranking by audience size allocates the bands that performed in bigger scenarios at the top of the ranking, as expected. We can modify the setup in order to include not only the size of the audience but also the rating (see Sect. 5) estimated assigned by each assistant.

Equation (3) defines the opinion measured in terms of the mobility behavior exhibited during the festival. We can use this metric to compute a ranking of musical bands or artists and improve on the ranking estimated using only the size of the audience to measure popularity. In order to combine the ratings of the audience, we propose a metric that assigns a significance value to the rating of each individual. Taking Eq. (3) as the basis, we expand it to incorporate a new factor that determines the significance of the opinion. This factor consists of size of the audience of band b, the amount of time allocated for the concert of band b, and the capacity of the hall in which band b performed. We incorporate what we call "general factor" in order to compensate for factors that affect the overall ranking. All this information is consolidated in Eq. (6).

$$wur_{u,b} = \frac{1}{visits_{u,r,b}} \left(\frac{visits_{u,b}}{visits_u} \right) log \left(\frac{visits_b}{l_b * c_{r,b}} \right) \tag{6}$$

where $log \left(\frac{visits_b}{l_a * c_{r,b}} \right)$ is the "general factor", $visits_b$ is the size of the audience of band b, l_b is the amount of time allocated for the concert of band b, and $c_{r,b}$ is the capacity of the room r where band b performed. The "general factor" compensates for other factors that are usually underestimated such as the size of the hall, time allocated, and other factors.

Equation (6), one of the main contributions of this paper, estimates the weighted rating (i.e. opinion) assigned by user u to musical band b. The total ranking points of band b is estimated as the sum of the weighted ratings computed for all the users that attended the performance of band b, as it is described in Eq. (7).

$$trp_b = \sum_{u \in U} wur_{u,b}, \tag{7}$$

fridaya@uef.fi

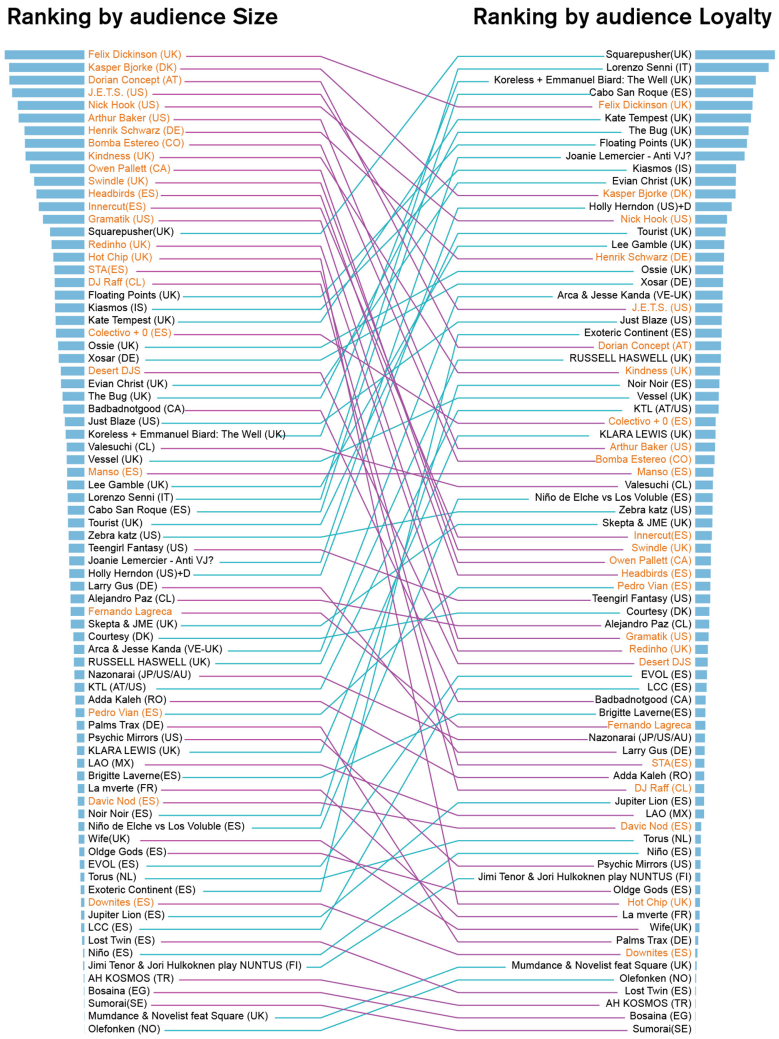


Fig. 7. Audience Size vs Audience Loyalty. In general, the ranking by audience size (left) allocates the bands that performed in bigger scenarios at the top of the ranking. In contrast, our loyalty ranking (right) brings up artists that, even though they played in smaller rooms, managed to attract the audience more consistently. Independently, we confirmed that the organization of the festival had considered many of these smaller shows a success.

where trp_b refers to the total ranking points and $wur_{u,b}$ is the weighted user rating computed in Eq. (6).

The rankings computed using Eq. (7) suggest that bands that were able to retain users for longer periods of time (ranking by audience loyalty) are placed at the top of the ranking (see Fig. 7). On the other hand, ranking by audience

fridaya@uef.fi

size favors those who performed in bigger halls, as it is shown in Fig. 7. Although some of the highly ranked artists in the ranking by audience size exhibited a lower popularity in the ranking by audience loyalty, artists such as Felix Dickinson was able to stay at the top part of the ranking due to the loyalty of his audience.

8 Conclusion and Future Work

We showed that we can leverage *implicit feedback* in the form of physical contextual information extracted from mobility traces to infer music preferences of an audience. We proposed a metric to estimate ratings assigned by attendees of a multistage music festival to the musical bands. These ratings, as well as other useful information such as time tables of the festival, are combined to give insights about audience mobility behavior during the festival, leading to music similarity estimation, and thus allowing us to cluster musical bands that appeal to defined audiences. Furthermore, we used this rating to define loyalty or engagement with the artists, and showed that rankings based on audience size may lead to an incomplete understanding of the success of a live show.

Our hope is that this article will awaken scientific curiosity about music similarity, music ranking, and music recommendation from human physical behavior. The results obtained in this work can be extended to other data sets collected from other sources of information and other settings such as technology fairs, conferences, and gatherings of people with multiple points of activity. As part of the future work, we seek to investigate the effect of combining both *implicit* and *explicit feedback*, and how we could overcome the cold-start problem in order to deploy a real-time recommender system during a multi-stage festival and provide useful suggestions to the audience based on their current behavior.

Acknowledgment. The authors would like to thank Advanced Music SL and Sonar Festival for their collaboration in this project, Barcelona Supercomputing Center, and CONACyT México for funding.

References

1. Mac address vendors. IEEE Public OUI list. http://standards-oui.ieee.org/oui/oui.txt
2. Analysis of iOS 8 MAC Randomization on Locationing. Technical report, Zebra Technologies, June 2015
3. Bastian, M., Heymann, S., Jacomy, M.: Gephi an open source software for exploring and manipulating networks (2009)
4. Benevenuto, F., Rodrigues, T., Cha, M., Almeida, V.: Characterizing user behavior in online social networks. In: Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference, IMC 2009, pp. 49–62. ACM, New York (2009)
5. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech. Theor. Exp. **2008**(10), P10008 (2008)

fridaya@uef.fi

6. Celma, Ò., Cano, P.: From hits to niches?: or how popular artists can bias music recommendation and discovery. In: Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition, NETFLIX 2008, pp. 5:1–5:8. ACM, New York (2008)
7. Chen, M., Nguyen, T., Szymanski, B.K.: On measuring the quality of a network community structure. In: 2013 International Conference on Social Computing (SocialCom), pp. 122–127, September 2013
8. Du, W., Lin, H., Sun, J., Yu, B., Yang, H.: Content-based music similarity computation with relevant component analysis. In: 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pp. 1043–1047, October 2016
9. Eagle, N., (Sandy) Pentland, A., Lazer, D.: Inferring social network structure using mobile phone data. In: Proceedings of National Academy of Sciences, p. 2009 (2009)
10. Hanna, P., Robine, M., Rocher, T.: An alignment based system for chord sequence retrieval. In: Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL 2009, pp. 101–104. ACM, New York (2009)
11. Hanna, P., Rocher, T., Robine, M.: A robust retrieval system of polyphonic music based on chord progression similarity. In: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009, pp. 768–769. ACM, New York (2009)
12. Hu, Y., Koren, Y., Volinsky, C.: Collaborative filtering for implicit feedback datasets. In: 2008 Eighth IEEE International Conference on Data Mining, pp. 263–272, December 2008
13. Knees, P., Schedl, M.: A survey of music similarity and recommendation from music context data. ACM Trans. Multimedia Comput. Commun. Appl. **10**(1), 2:1–2:21 (2013)
14. Kossinets, G., Watts, D.J.: Empirical analysis of an evolving social network. Science **311**(5757), 88–90 (2006)
15. Lee, C., Cunningham, P.: Community detection: effective evaluation on large social networks. J. Complex Netw. **2**(1), 19–37 (2014)
16. Li, T., Ogihara, M.: Music artist style identification by semi-supervised learning from both lyrics and content. In: Proceedings of the 12th Annual ACM International Conference on Multimedia, MULTIMEDIA 2004, pp. 364–367. ACM, New York (2004)
17. Lin, N., Tsai, P.C., Chen, Y.A., Chen, H.H.: Music recommendation based on artist novelty and similarity. In: 2014 IEEE 16th International Workshop on Multimedia Signal Processing (MMSP), pp. 1–6, September 2014
18. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. **E69**(026113) (2004)
19. Panteli, M., Bittner, R., Bello, J.P., Dixon, S.: Towards the characterization of singing styles in world music. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 636–640, March 2017
20. Schedl, M., Hauger, D.: Mining microblogs to infer music artist similarity and cultural listening patterns. In: Proceedings of the 21st International Conference on World Wide Web, WWW 2012 Companion. pp. 877–886. ACM, New York (2012)
21. Staab, S., Domingos, P., Mika, P., Golbeck, J., Ding, L., Finin, T., Joshi, A., Nowak, A., Vallacher, R.R.: Social networks applied. IEEE Intell. Syst. **20**(1), 80–93 (2005)
22. Steinhaeuser, K., Chawla, N.V.: Identifying and evaluating community structure in complex networks. Pattern Recogn. Lett. **31**(5), 413–421 (2010)

fridaya@uef.fi

23. van Gennip, Y., Hunter, B., Ahn, R., Elliott, P., Luh, K., Halvorson, M., Reid, S., Valasik, M., Wo, J., Tita, G.E., Bertozzi, A.L., Jeffrey Brantingham, P.: Community detection using spectral clustering on sparse geosocial data. *CoRR*, abs/1206.4969 (2012)
24. Yang, J., Leskovec, J.: Defining and evaluating network communities based on ground-truth. In: Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics, MDS 2012, pp. 3:1–3:8. ACM, New York (2012)
25. You, S.D., Shih, H.: Subjective evaluation of music similarity system based on onsets. In: 2018 IEEE International Conference on Applied System Invention (ICASI), pp. 378–380, April 2018
26. Zaharia, M., Xin, R.S., Wendell, P., Das, T., Armbrust, M., Dave, A., Meng, X., Rosen, J., Venkataraman, S., Franklin, M.J., Ghodsi, A., Gonzalez, J., Shenker, S., Stoica, I.: Apache spark: a unified engine for big data processing. Commun. ACM **59**(11), 56–65 (2016)
27. Zhu, X., Shi, Y.Y., Kim, H.G., Eom, K.W.: An integrated music recommendation system. IEEE Trans. on Consum. Electron. **52**(3), 917–925 (2006)

fridaya@uef.fi

Characterization and Recognition of Proper Tagged Probe Interval Graphs

Sanchita Paul[1], Shamik Ghosh[2(✉)], Sourav Chakraborty[3], and Malay Sen[4]

[1] Department of Mathematics, Jadavpur University, Kolkata 700032, India
sanchitajumath@gmail.com
[2] Department of Mathematics, Jadavpur University, Kolkata 700032, India
ghoshshamik@yahoo.com
[3] Indian Statistical Institute, Kolkata 700108, India
sourav@isical.ac.in
[4] Department of Mathematics, North Bengal University,
Siliguri 734430, West Bengal, India
senmalay10@gmail.com

Abstract. Interval graphs were used in the study of genomics by the famous molecular biologist Benzer. Later on probe interval graphs were introduced by Zhang as a generalization of interval graphs for the study of cosmid contig mapping of DNA. A tagged probe interval graph (briefly, TPIG) is motivated by similar applications to genomics, where the set of vertices is partitioned into two sets, namely, probes and nonprobes and there is an interval on the real line corresponding to each vertex. The graph has an edge between two probe vertices if their corresponding intervals intersect, has an edge between a probe vertex and a nonprobe vertex if the interval corresponding to a nonprobe vertex contains at least one end point of the interval corresponding to a probe vertex and the set of nonprobe vertices is an independent set. This class of graphs have been defined nearly two decades ago, but till today there is no known recognition algorithm for it. In this paper, we consider a natural subclass of TPIG, namely, the class of *proper tagged probe interval graphs* (in short PTPIG). We present characterization and a linear time recognition algorithm for PTPIG. To obtain this characterization theorem we introduce a new concept called *canonical sequence* for proper interval graphs, which, we belief, has an independent interest in the study of proper interval graphs. Also to obtain the recognition algorithm for PTPIG, we introduce and solve a variation of consecutive 1's problem, namely, *oriented-consecutive* 1*'s problem* and some variations of PQ-tree algorithm.

Keywords: Interval graph · Proper interval graph · Probe interval graph · Probe proper interval graph · Tagged probe interval graph · Consecutive 1's property · PQ-tree algorithm

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 62–75, 2019.
https://doi.org/10.1007/978-3-030-22868-2_5

fridaya@uef.fi

1 Introduction

A graph $G = (V, E)$ is an *interval graph* if one can map each vertex into an interval on the real line so that any two vertices are adjacent if and only if their corresponding intervals intersect. Such a mapping of vertices into an interval on the real line is called an *interval representation of* G. The study of interval graphs was motivated by the study of the famous molecular biologist Benzer [1] in 1959. Since then interval graphs has been widely used in molecular biology and genetics, particularly for DNA sequencing. Different variations of interval graphs have been used to model different scenarios arising in the area of DNA sequencing. Literature on the applications of different variations of interval graphs can be found in [4,11,17].

In an attempt to aid a problem called *cosmid contig mapping*, a particular component of the physical mapping of DNA, in 1998 Sheng, Wang and Zhang [25] defined a new class of graphs called *tagged probe interval graphs* (briefly, TPIG) which is a generalization of interval graphs. Since then one of the main open problems in this area has been "Given a graph G, recognizing whether G is a *tagged probe interval graphs*".

Definition 1.1. A graph $G = (V, E)$ is a *tagged probe interval graph* if the vertex set V can be partitioned into two disjoint sets P (called "probe vertices") and N (called "nonprobe vertices") and one can map each vertex into an interval on the real line (vertex $x \in V$ mapped to $I_x = [\ell_x, r_x]$) such that all the following conditions hold:

1. N is an independent set in G, i.e., there is no edge between nonprobe vertices.
2. If $x, y \in P$, then there is an edge between x and y if and only if $I_x \cap I_y \neq \emptyset$, or in other words the mapping is an interval representation of the subgraph of G induced by P.
3. If $x \in P$ and $y \in N$, then there is an edge between x and y if and only if the interval corresponding to the nonprobe vertex contains at least one end point of the interval corresponding to the probe vertex, i.e., either $\ell_x \in I_y$ or $r_x \in I_y$.

We call the collection $\{I_x \mid x \in V\}$ a *TPIG representation* of G. If the partition of the vertex set V into probe and nonprobe vertices is given, then we denote the graph as $G = (P, N, E)$.

Problem 1. Given a graph $G = (P, N, E)$, give a linear time algorithm for checking if G is a *tagged probe interval graph.*

Tagged probe inteval graphs have been defined nearly two decades ago and its importance in the context of molecular biology has been emphasized over and over again [24–27]. Yet untill this paper there was no known algorithm for tagged probe interval graphs or any natural subclass of tagged probe interval graphs, excepting probe proper interval graphs.

A natural and well studied subclass of interval graphs are the *proper interval graphs*. A *proper interval graph* G is an interval graph in which there is an

fridaya@uef.fi

interval representation of G such that no interval contains another properly. Such an interval representation is called a *proper interval representation of* G. Proper interval graphs is an extremely rich class of graphs and we have a number of different characterizations of proper interval graphs.

In this paper we study a natural special case of tagged probe interval graphs which we call *proper tagged probe interval graph* (in short, *PTPIG*). The only extra condition that a PTPIG should satisfy over TPIG is that the mapping of the vertices into intervals on the real line that gives a TPIG representation of G should be a **proper interval representation** of the subgraph of G induced by P.

In this paper, we present a linear time (linear in $(|V| + |E|)$ recognition algorithm for PTPIG. The backbone of our recognition algorithm is the characterization of proper tagged probe interval graphs that we obtain in Theorem 3.3 in Sect. 3. To obtain this characterization theorem we introduce (in Sect. 2) a new concept called **"canonical sequence"** for proper interval graphs, which we believe would be of independent interest in the study of proper interval graphs. The concept of canonical sequence for proper interval graphs can be used to solve other problems related to proper interval graphs, for example testing isomorphism of proper interval graphs.

1.1 Organization of the Paper

We present the characterization of Proper Tagged Probe Interval Graphs (PTPIG) in Sect. 3. The main combinatorial object we need for the characterization is the Canonical Sequence of Proper Interval Graphs that is presented in Sect. 2. We present our recognition algorithm in Sect. 4. In the road to obtain the linear time recognition algorithm for PTPIG, we face a number of algorithmic challenges that led us to solve a number of sub-problems that can be of independent interest. One particular problem worth mentioning is a generalization of the well known consecutive 1's problem, which we call the *oriented-consecutive 1's problem*. It is a very important sub-routine in our recognition algorithm.

Due to lack of space we are unable to provide the technical proofs of various lemmas and theorems here. We do give a brief idea of the proofs wherever possible. However the full details are available in the arXiv version of our paper [20].

1.2 Notations

Suppose a graph G is a PTPIG (or TPIG), then we will assume that the vertex set is partitioned into two sets P (for probe vertices) and N (for nonprobe vertices). To indicate that the partition is known to us, we will sometimes denote G by $G = (P, N, E)$, where E is the edge set. We will denote by G_P the subgraph of G that is induced by the vertex set P. We will assume that there are p probe vertices $\{u_1, \ldots, u_p\}$ and q nonprobe vertices $\{w_1, \ldots, w_q\}$. To be consistent in our notation we will use i or i' or $i_1, i_2, \ldots$ as indices for probe vertices and use j or j' or $j_1, j_2, \ldots$ as indices for nonprobe vertices.

fridaya@uef.fi

Let $G = (V, E)$ be a graph and $v \in V$. Then by the *closed neighborhood* of v in G we mean the set $N[v] = \{u \in V \mid u \text{ is adjacent to } v\} \cup \{v\}$. A graph G is called *reduced* if no two vertices have the same closed neighbourhood. If the graph is not reduced then we define an equivalence relation on the vertex set V such that v_i and v_j are equivalent if and only if v_i and v_j have the same (closed) neighbors in V. Each equivalence class under this relation is called a *block* of G. For any vertex $v \in V$ we denote the equivalence class containing v by $B(v)$. So, the collection of blocks is a partition of V. The reduced graph of G (denoted $\widetilde{G} = (\widetilde{V}, \widetilde{E})$) is the graph obtained by merging all the vertices that are in the same equivalence class.

If M is a $(0, 1)$-matrix, then we say M satisfies the *consecutive* 1*'s property* if in each column 1's appear consecutively. [15,18] By $A(G)$ we denote the *augmented adjacency matrix* of the graph G, in which the diagonal entries are 1 and non-diagonal entries are same as the adjacency matrix of G.

1.3 Background Materials

PQ-Trees: In the past few decades many variations of interval graphs has been studied mostly in context of modeling different scenario arising from molecular biology and DNA sequencing. Understanding the structure and properties of these classes of graphs and designing efficient recognition algorithms are the central problems in this field. Many times this studies have led to nice combinatorial problems and development of important data structures.

For example, the original linear time recognition algorithm for interval graphs by Booth and Lueker [3] in 1976 is based on their complex PQ tree data structure (also see [2]). Habib et al. [12] in 2000 showed how to solve the problem more simply using lexicographic breadth-first search, based on the fact that a graph is an interval graph if and only if it is chordal and its complement is a comparability graph. A similar approach using a 6-sweep LexBFS algorithm is described in Corneil, Olariu and Stewart [7] in 2010.

In this paper we will be using the data structure of PQ-trees quite extensively. PQ-trees are not only used to check whether a given matrix satisfy the consecutive 1's property, they also store all the possible permutations such that if one permuted the rows using the permutation, the matrix would satisfy the consecutive 1's property. We generalize the problem of checking consecutive 1's property to Oriented-consecutive 1's problem and used the PQ-tree representation to solve this problem also.

Proper Interval Graphs: For proper interval graphs we have a number of equivalent characterizations. Recall that a proper interval graph G is an interval graph in which there is an interval representation of G such that no interval contains another properly and such an interval representation is called a *proper interval representation* of G. It is important to note that a proper interval graph G may have an interval representation which is not proper. Linear-time recognition algorithms for proper interval graphs are obtained in [8,9,13,14]. A *unit interval graph* is an interval graph in which there is an interval representation of

fridaya@uef.fi

Table 1. Characterizations of proper interval graphs: equivalent conditions on an interval graph $G = (V, E)$.

Index	Properties	References
1	G is a Proper Interval Graph	[10,21–23]
2	G is a Unit Interval Graph	
3	G is *claw-free*, i.e., G does not contain $K_{1,3}$ as an induced subgraph	
4	For all $v \in V$, elements of $N[v] = \{u \in V \mid uv \in E\} \cup \{v\}$ are consecutive for some ordering of V (closed neighborhood condition)	[5–7,10]
5	There is an ordering $v_1, v_2, \cdots, v_n$ of V such that G has a proper interval graph representation $\{I_{v_i} = [a_i, b_i] \mid i = 1, 2, \cdots, n\}$ where $a_1 < a_2 < \cdots < a_n$ and $b_1 < b_2 < \cdots < b_n$	
6	There is an ordering of V such that the augmented adjacency matrix $A(G)$ of G satisfies the consecutive 1's property	
7	*A straight enumeration of G is a linear ordering of blocks(vertices having same closed neighborhood) in G, such that for every block, the block and its neighboring blocks are consecutive in the ordering* G has a straight enumeration (which is unique up to reversal, if G is connected)	[8,13,14,19]
8	*The reduced graph $\widetilde{G}$ is obtained from G by merging vertices having same closed neighborhood* $G(n, r)$ *is a graph with n vertices $x_1, x_2, \ldots, x_n$ such thatx_i is adjacent to x_j if and only if $0 < \lvert i - j \rvert \leqslant r$, where r is a positive integer* $\widetilde{G}$ is an induced subgraph of $G(n, r)$ for some positive integers n, r with $n > r$	[16]

G such that all intervals have the same length. Interestingly, these two concepts are equivalent. Another equivalence is that an interval graph is a proper interval graph if and only if it does not contain $K_{1,3}$ as an induced subgraph. Apart from these, there are several characterizations of proper interval graphs (see Table 1). Among them we repeatedly use the following equivalent conditions in the rest of the paper:

Theorem 1.2. *Let $G = (V, E)$ be an interval graph, then the following are equivalent:*

1. *G is a proper interval graph.*
2. *There is an ordering of V such that for all $v \in V$, elements of $N[v]$ are consecutive (the closed neighborhood condition).*
3. *There is an ordering of V such that the augmented adjacency matrix $A(G)$ of G satisfies the consecutive 1's property.*

fridaya@uef.fi

4. *There is an ordering* $\{v_1, \ldots, v_n\}$ *of* V *such that* G *has a proper interval representation, say* $\{I_{v_i} = [a_i, b_i] \mid i = 1, 2, \ldots, n\}$ *where* $a_i \neq b_j$ *for all* $i, j \in \{1, 2, \ldots, n\}$ *and* $a_1 < a_2 < \cdots < a_n$ *and* $b_1 < b_2 < \cdots < b_n$.

Remark 1.3. We note that in a proper interval graph $G = (V, E)$, the ordering of V that satisfies any one of the conditions (2), (3) and (4) in the above theorem also satisfies the other conditions.

2 Canonical Sequence of Proper Interval Graphs

Let G be a proper interval graph. Then there is an ordering of V that satisfies conditions (2), (3) and (4) of Theorem 1.2. Henceforth we call such an ordering, a *natural* or *canonical* ordering of V. But this canonical ordering is not unique. A proper interval graph may have more than one canonical orderings of its vertices. Interestingly, it follows from Corollary 2.5 of [8] (also see [19]) that the canonical ordering is unique up to reversal for a connected reduced proper interval graph.

Definition 2.1. Let $G = (V, E)$ be a proper interval graph. Let $\{v_1, v_2, \ldots, v_n\}$ be a canonical ordering of the set V with the interval representation be

$$\{I_{v_i} = [a_i, b_i] \mid i = 1, 2, \ldots, n\}$$

where $a_i \neq b_j$ for all $i, j \in \{1, 2, \ldots, n\}$, $a_1 < a_2 < \cdots < a_n$ and $b_1 < b_2 < \cdots < b_n$. Now we combine all a_i and b_i $(i = 1, 2, \ldots, n)$ in an increasing sequence which we call the *interval canonical sequence* with respect to the canonical ordering of vertices of G and is denoted by $\mathcal{I}_G$.

Now if we replace a_i or b_i by i for all $i = 1, 2, \ldots, n$ in $\mathcal{I}_G$, then we obtain a sequence of integers belonging to $\{1, 2, \ldots, n\}$ each occurring twice. We call such a sequence a *canonical sequence* of G with respect to the canonical ordering of vertices of G and is denoted by $\mathcal{S}_G$. Moreover, if we replace i by v_i for all $i = 1, 2, \ldots, n$ in $\mathcal{S}_G$, then the resulting sequence is called a *vertex canonical sequence* of G (corresponding to the canonical sequence $\mathcal{S}_G$) and is denoted by $\mathcal{V}_G$.

Note that $\mathcal{S}_G$ and its corresponding $\mathcal{V}_G$ and $\mathcal{I}_G$ can all be obtained uniquely from each other. By abuse of notations, sometimes we will use the term canonical sequence to mean any of these.

2.1 Structure of the Canonical Sequence for Proper Interval Graphs

If a graph G is a connected reduced proper interval graph then the following lemma states that the canonical sequence for G is unique up to reversal.

Lemma 2.2. *Let* $G = (V, E)$ *be a proper interval graph and* $V = \{v_1, v_2, \ldots, v_n\}$ *be a canonical ordering of vertices of* G. *Then the canonical sequence* $\mathcal{S}_G$ *is independent of proper interval representations that satisfy the given canonical ordering. Moreover* $\mathcal{S}_G$ *is unique up to reversal for connected reduced proper interval graphs.*

fridaya@uef.fi

Now there is an alternative way to get the canonical sequence directly from the augmented adjacency matrix. Let $G = (V, E)$ be a proper interval graph with $V = \{v_i \mid i = 1, 2, \ldots, n\}$ and $A(G)$ be the augmented adjacency matrix of G with consecutive 1's property. We partition positions of $A(G)$ into two sets (L, U) by drawing a polygonal path from the upper left corner to the lower right corner such that the set L [resp. U] is closed under leftward or downward [respectively, rightward or upward] movement (called a *stair partition*) and U contains precisely all the zeros right to the principal diagonal of $A(G)$. This is possible due to the consecutive 1's property of $A(G)$. Now we obtain a sequence of positive integers belonging to $\{1, 2, \ldots, n\}$, each occurs exactly twice, by writing the row or column numbers as they appear along the stair. We call this sequence, the *stair sequence* of $A(G)$ and note that it is same as the canonical sequence of G with respect to the given canonical ordering of vertices of G.

Proposition 2.3. *Let $G = (V, E)$ be a proper interval graph with a canonical ordering $V = \{v_1, v_2, \ldots, v_n\}$ of vertices of G. Let $A(G)$ be the augmented adjacency matrix of G arranging vertices in the same order as in the canonical ordering. Then the canonical sequence $\mathcal{S}_G$ of G is the same as the stair sequence of $A(G)$.*

Corollary 2.4. *If G is a connected proper interval graph then $\mathcal{S}_G$ is unique up to reversal.*

Remark 2.5. Let G be a connected proper interval graph which is not reduced and $\widetilde{G}$ be the reduced graph of G. Then the graph $\widetilde{G}$ has a unique (upto reversal) canonical ordering of vertices, say, $b_1, \ldots, b_t$ (corresponding to the blocks $B_1, \ldots, B_t$) as it is connected and reduced. Now the canonical orderings of the vertices of G are obtained from this ordering (and its reversal) by all possible permutation of the vertices of G within each block. In all such cases $\mathcal{S}_G$ will remain same up to reversal.

3 Structure of PTPIG

Let us recall the definition of proper tagged probe interval graph.

Definition 3.1. A tagged probe interval graph $G = (P, N, E)$ is a *proper tagged probe interval graph* (PTPIG) if G has a TPIG representation $\{I_x \mid x \in P \cup N\}$ such that $\{I_p \mid p \in P\}$ is a proper interval representation of the graph G_P. We call such an interval representation a *PTPIG representation* of G.

It is interesting to note that there are examples of $TPIG$, G for which G_P is a proper interval graph but G is not a $PTPIG$. For example, the graph G_b (see Fig. 1) in [25] is a TPIG in which $(G_b)_P$ consists of a path of length 4 along with 2 isolated vertices which is a proper interval graph. But G_b has no TPIG representation with a proper interval representation of $(G_b)_P$.

fridaya@uef.fi

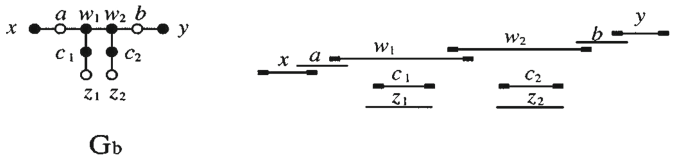


Fig. 1. The graph G_b and its TPIG representation [25]

Now let us consider a graph $G = (V, E)$, in general, with an independent set N and $P = V \setminus N$ such that the subgraph G_P of G induced by P is a proper interval graph. Let us order the vertices of P in a canonical ordering. Now the adjacency matrix of G looks like the following:

	P	N
P	$A(P)$	$A(P, N)$
N	$A(P, N)^T$	$\mathbf{0}$

Note that the (augmented) adjacency matrix $A(P)$ of G_P satisfies the consecutive 1's property and the $P \times N$ submatrix $A(P, N)$ of the adjacency matrix of G represents edges between probe vertices and nonprobe vertices. In the following lemma we obtain a necessary condition for a PTPIG.

Lemma 3.2. *Let $G = (P, N, E)$ be a PTPIG. Then for any canonical ordering of the vertices belonging to P each column of $A(P, N)$ can not have more than two consecutive stretches of 1's.*

Unfortunately the condition in the above lemma is not sufficient. For convenience, we say an interval $I_p = [a, b]$ *contains strongly* an interval $I_n = [c, d]$ if $a < c \leqslant d < b$, where $p \in P$ and $n \in N$.[1] The following is a characterization theorem for a PTPIG and is our main Theorem. For convenience, henceforth, a continuous stretch (subsequence) in a canonical sequence will be called a *substring*.

Theorem 3.3. *Let $G = (V, E)$ be a graph with an independent set N and $P = V \setminus N$ such that G_P, the subgraph induced by P is a proper interval graph. Then G is a proper tagged probe interval graph with probes P and nonprobes N if and only if there is a canonical ordering of vertices belonging to P such that the following condition holds:*

(A) *for every nonprobe vertex $w \in N$, there is a substring in the canonical sequence with respect to the canonical ordering such that all the vertices in the substring are neighbors of w and all the neighbors of w are present at least once in the substring.*

[1] In [27], Sheng et al. used the term "contains properly" in this case. Here we consider a different term in order to avoid confusion with the definition of proper interval graph. Note that if $a \leqslant c \leqslant d < b$ or $a < c \leqslant d \leqslant b$, then also I_p contains I_n properly, but not strongly.

fridaya@uef.fi

Such a substring will be called a perfect substring *(cf. Definition 3.5)*.

Idea of the Proof: As G_P is a proper interval graph, there exist a canonical ordering of vertices of P satisfying conditions of Theorem 1.2. Hence one can get a canonical sequence S_{G_P} from this ordering which is basically a combined increasing sequence of the endpoints of the probe intervals. From the definition of PTPIG we know that adjacency between a probe and a nonprobe vertex happens when interval corresponding to the nonprobe vertex intersects interval corresponding to the probe vertex to one of its ends. Hence the end points of the neighbours of a nonprobe vertex $w \in N$ must occur consecutively in S_{G_P}. As all the vertices which are neighbours of $w \in N$ occur consecutively as a substring in S_{G_P}, one can able to construct interval corresponding to w by taking first and last positions of the substring as its endpoint. Note that each probe vertex occurs twice in S_{G_P}. Hence one can assign the intervals corresponding to the probe vertices by taking their first and second occurrence positions in S_{G_P} as their end points.

Remark 3.4. If G is a PTPIG such that G_P is connected and reduced, then there is a unique (up to reversal) canonical ordering of vertices belonging to P, as we mentioned at the beginning of Sect. 2. Thus the corresponding canonical sequence is also unique up to reversal. Also if condition (A) holds for a canonical sequence, it also holds for its reversal. Thus in this case condition (A) holds for **any** canonical ordering of vertices belonging to P.

We conclude this section with some nice structural insight of the substrings described in condition (A) of Theorem 3.3.

Definition 3.5. Let $G = (V, E)$ be a graph with an independent set N and $P = V \setminus N$ such that G_P, the subgraph induced by P is a proper interval graph. Let $\mathcal{S}_{G_P}$ be a canonical sequence of G_P. Let $w \in N$. If there exists a substring in $\mathcal{S}_{G_P}$ which contains all the neighbors of w and all the vertices in the substring are neighbors of w then we call the substring a *perfect substring of w in G*. If the canonical sequence $\mathcal{S}_{G_P}$ contains a *perfect substring of w in $\mathcal{S}_{G_P}$* for all $w \in N$, we call it a *perfect canonical sequence for G*.

Proposition 3.6. *Let $G = (P, N, E)$ be a PTPIG such that G_P is a connected reduced proper interval graph and $\mathcal{S}_{G_P}$ be a canonical sequence of G_P. Then for any nonprobe vertex $w \in N$, there cannot exist more than one disjoint perfect substrings of w in $\mathcal{S}_{G_P}$, unless the substring consists of a single element.*

In fact, we can go one step more in understanding the structure of a PTPIG. If G is a PTPIG, not only there cannot be two disjoint perfect substrings (of length more than 1) for any nonprobe vertex in any canonical sequence but also any two perfect substrings for the same vertex must intersect at atleast two places, except two trivial cases.

Lemma 3.7. *Let $G = (P, N, E)$ be a PTPIG such that G_P is a connected reduced proper interval graph with a canonical ordering of vertices*

fridaya@uef.fi

$\{u_1, u_2, \ldots, u_p\}$ and let $\mathcal{V}_{G_P}$ be the corresponding vertex canonical sequence of G_P. Let $w \in N$ be such that w has at least two neighbors in P and T_1 and T_2 be two perfect substrings for w in $\mathcal{V}_{G_P}$ intersecting in exactly one place. Then one of the following holds:

1. *$\mathcal{V}_{G_P}$ begins with $u_1u_2u_1$ and only u_1 and u_2 are neighbors of w.*
2. *$\mathcal{V}_{G_P}$ ends with $u_pu_{p-1}u_p$ and only u_{p-1} and u_p are neighbors of w.*

4 Recognition Algorithm

In this section, we present a linear time recognition algorithm for PTPIG. That is, given a graph $G = (V, E)$, and a partition of the vertex set into N and $P = V \setminus N$ we can check, in time $O(|V| + |E|)$, if the graph $G(P, N, E)$ is a PTPIG. Now $G = (P, N, E)$ is a PTPIG if and only if it a TPIG, i.e., it satisfies the three conditions in Definition 1.1 and G_P is a proper interval graph for a TPIG representation of G. Note that it is easy to check in linear time if the graph G satisfies the first condition, namely, if N is an independent set in the graph. Now for testing if the graph satisfies the other two properties we will use the characterization we obtained in Theorem 3.3.

We will use the recognition algorithm for proper interval graph $H = (V', E')$ given by Booth and Lueker [3] as a blackbox that runs in $O(|V'| + |E'|)$. The main idea of their algorithm is that H is a proper interval graph if and only if the adjacency matrix of the graph satisfies the consecutive 1's property. In other words, H is a proper interval graph if and only if there is an ordering of the vertices of H such that for any vertices v in H, the neighbors of v must be consecutive in that ordering. So for every vertex v in H they consider restrictions, on the ordering of the vertices, of the form "all vertices in the neighborhood of v must be consecutive". This is done using the data structure of PQ-trees. The PQ-tree helps to store all the possible orderings that respect all these kind of restrictions. It is important to note that all the orderings that satisfy all the restrictions are precisely all the canonical orderings of vertices of H.

The main idea of our recognition algorithm is that if the graph $G = (P, N, E)$ is PTPIG then, from Condition **(A)** in Theorem 3.3, we can obtain a series of restrictions on the ordering of vertices that also can be "stored" using the PQ-tree data structure. These restrictions are on and above the restrictions that we need to ensure the graph G_P is a proper interval graph. If finally there exists any ordering of the vertices that satisfy all the restrictions, then that ordering will be a canonical ordering that satisfies the condition **(A)** in Theorem 3.3. So the main challenge is to identify all the extra restrictions on the ordering and how to store them in the PQ-tree.

Once we have verified that the N is an independent set and the graph G_P is a proper interval graph and we have stored all the possible canonical ordering of the vertices of the subgraph $G_P = (P, E_1)$ in a PQ-tree (in $O(|P| + |E_1|)$ time), we proceed to find the extra restrictions that is necessary to be applied on the orderings. We present our algorithm in 3 steps - each step handling a class of graphs that is a generalization of the class of graphs handled in the previous one.

fridaya@uef.fi

- **STEP I:** First we consider the case when G_P is a connected reduced proper interval graph.
- **STEP II:** Next we consider the case when G_P is a connected proper interval graph, but not necessarily reduced.
- **STEP III:** Finally we consider case when the graph G_P is a proper interval graph, but may not be connected or reduced.

4.1 Step I: The Graph G_P is a Connected Reduced Proper Interval Graph

By Lemma 2.2, there is a unique (up to reversal) canonical ordering of the vertices of G_P. By Theorem 3.3, we know that the graph G is PTPIG if and only if the following condition is satisfied:

Condition (A1): For all $1 \leq j \leq q$, there is a substring in $\mathcal{S}_{G_P}$ where only the neighbors of w_j appear and all the neighbors of w_j appear at least once.

In this case, when the graph G_P is connected reduced proper interval graph, since there is a unique canonical ordering of the vertices, all we have to do is to check if the corresponding canonical sequence satisfies the Condition (A1). So the rest of the algorithm in this case is to check if the property is satisfied.

Idea of the Algorithm: Since we know the canonical sequence $\mathcal{S}_{G_P}$ (or obtain by using known algorithms described before in $O(|P| + |E_1|)$ time, where E_1 is the set of edges between probe vertices), we can have two look up tables L and R such that for any vertex $v_i \in P$, the $L(v_i)$ and $R(v_i)$ has the index of the first and the second appearance of v_i in $\mathcal{S}_{G_P}$ respectively. We can obtain the look up tables in time $O(|P|)$ steps.

Also by $\mathcal{S}_{G_P}[k_1, k_2]$ (where $1 \leq k_1 \leq k_2 \leq 2p$) we will denote the substring of the canonical sequence sequence $\mathcal{S}_{G_P}$ that start at the k_1^{th} position and ends at the k_2^{th} position in $\mathcal{S}_{G_P}$.

To check the Condition (A1), we will go over all the $w_j \in N$. For $j \in \{1, 2, \ldots, q\}$, let $L(A_j[1]) = \ell_j$ and $R(A_j[1]) = r_j$. Now since all the neighbors of w_j have to be in a substring, there must be a substring of length at least d_j and at most $2d_j$ (as each number appears twice) in $\mathcal{S}_{G_P}[\ell_j - 2d_j, \ell_j + 2d_j]$ or $\mathcal{S}_{G_P}[r_j - 2d_j, r_j + 2d_j]$ which contains only and all the neighbors of w_j. We can identify all such possible substrings by first marking the positions in $\mathcal{S}_{G_P}[\ell_j - 2d_j, \ell_j + 2d_j]$ and $\mathcal{S}_{G_P}[r_j - 2d_j, r_j + 2d_j]$ those are neighbors of w_j and then by doing a double pass, we find all the possible substrings of length greater than or equal to d_j in $\mathcal{S}_{G_P}[\ell_j - 2d_j, \ell_j + 2d_j]$ and $\mathcal{S}_{G_P}[r_j - 2d_j, r_j + 2d_j]$ that contains only neighbors of w_j.

Going through this way one can correctly decide whether G is a PTPIG with probes P and nonprobes N in time $O(|P| + |N| + |E_2|)$, where E_2 is the set of edges between probes P and nonprobes N when G_P is connected reduced proper interval graph. As obtaining $\mathcal{S}_{G_P}$ requires $O(|P| + |E_1|)$ time, the total recognition time is $O(|P| + |N| + |E_1| + |E_2|) = O(|V| + |E|)$.

fridaya@uef.fi

4.2 Step II: The Graph G_P is a Connected (Not Necessarily Reduced) Proper Interval Graph

In this case, that is when the graph G_P is not reduced, we cannot say that there exists a unique canonical ordering of the vertices of G_P. By Theorem 3.3, all we can say is that among the set of canonical orderings of the vertices of G_P, is there an ordering such that the corresponding canonical sequence satisfies Condition (A) of Theorem 3.3. As mentioned before we will assume that we have all the possible canonical ordering of the vertices of G_P stored in a PQ-tree. Now we will impose more constraints on the orderings so that the required condition is satisfied.

Let $\widetilde{G_P}$ be the reduced graph of G_P. By Remark 2.5, $\widetilde{G_P}$ has a unique (upto reversal) canonical ordering of vertices, say, $b_1, \ldots, b_t$ (corresponding to the blocks $B_1, \ldots, B_t$ of the vertices of G_P) and the canonical orderings of the vertices of G_P are obtained by all possible permutations of the vertices of G within each block.

For G to be a PTPIG we will have to find a suitable canonical ordering of the vertices or in other words, by Remark 2.5, we need to find suitable ordering of vertices in each block such that the Condition (A) from Theorem 3.3 is satisfied. Using the structure of the canonical sequence of $\widetilde{G_P}$ we identity different cases and in each case we can identify the various restrictions on the ordering of the vertices in each block that are necessary and sufficient.

Oriented-Consecutive 1's Problem: We introduce a generalization of the consecutive 1's problem called the *Oriented-consecutive 1's problem* and reduce the problem of checking if there is orderings of the vertices which satisfy all the restrictions to this problem. We can solve the oriented-consecutive 1's problem using the PQ-tree.

4.3 Step III: The Graph G_P is a Proper Interval Graph (Not Necessarily Connected or Reduced)

Finally, we consider the graph $G = (V, E)$ with an independent set N (non-probes) and $P = V \setminus N$ (probes) such that G_P is a proper interval graph, which may not be connected. Let the connected components of G_P be $G_1, \ldots, G_r$. Now, G is a PTPIG if and only if the following condition is satisfied:

Condition (C1): There exists permutation $\pi : \{1, \ldots, r\} \to \{1, \ldots, r\}$ and canonical sequences $\mathcal{S}_{G_1}, \ldots, \mathcal{S}_{G_r}$ of $G_1, \ldots, G_r$ such that the canonical sequence $\mathcal{S}_{G_P}$ of G_P obtained by concatenation of the canonical sequences of $G_{\pi(1)}, \ldots, G_{\pi(r)}$ (that is, $\mathcal{S}_{G_P} = \mathcal{S}_{G_{\pi(1)}} \ldots \mathcal{S}_{G_{\pi(r)}}$) has the property that for all $w \in N$ there exists a *perfect substring* of w in $\mathcal{S}_{G_P}$ (that is, there exists a substring of $\mathcal{S}_{G_P}$ where only the vertices of w appear and all the neighbors of w appears at least once).

Firstly using previous steps we store all the possible canonical orderings of the vertices in each component so that the graphs induced by $G_k \cup N$ is a PTPIG, for each k. As usual we will store the restrictions using the PQ-tree. Next we

fridaya@uef.fi

will have to add some more restrictions on the canonical ordering of the vertices in each of the connected components which are necessary for the graph G to be a PTPIG. These restriction will be be stored in the same PQ-tree. At last we check if there exists an ordering of the components such that the Condition (C1) is satisfied.

5 Conclusion

The study of interval graphs was spearheaded by Benzer [1] in his studies in the field of molecular biology. In [28], Zhang introduced a generalization of interval graphs called probe interval graphs (PIG) in an attempt to aid a problem called cosmid contig mapping. In order to obtain a better model another generalization of interval graphs were introduced that capture the property of overlap information, namely tagged probe interval graphs (TPIG) by Sheng, Wang and Zhang in [25]. Still there is no recognition algorithm for TPIG, in general.

In this paper, we characterize and obtain linear time recognition algorithm for a special class of TPIG, namely proper tagged probe interval graphs (PTPIG). The problem of obtaining a recognition algorithm for TPIG, in general is challenging and open till date. It is well known that an interval graph is a proper interval graph if and only if it does not contain $K_{1,3}$ as an induced subgraph of it. Similar forbidden subgraph characterization for PTPIG is another interesting problem.

References

1. Benzer, S.: On the topology of the genetic fine structure. Proc. Nat. Acad. Sci. USA **45**, 1607–1620 (1959)
2. Booth, K.S., Lueker, G.S.: Linear algorithm to recognize interval grapha and test for the consecutive ones property. In: Proceedings of the 7th ACM Symposium Theory of Computing, pp. 255–265 (1975)
3. Booth, K.S., Lueker, G.S.: Testing for the consecutive ones property, interval graphs and graph planarity using PQ-tree algorithms. J. Comput. Syst. Sci. **13**, 335–379 (1976)
4. Brown, D.E.: Variations on interval graphs. Ph.D. Thesis, University of Colorado at Denver, USA (2004)
5. Corneil, D.G.: A simple 3-sweep LBFS algorithm for the recognition of unit interval graphs. Discrete Appl. Math. **138**, 371–379 (2004)
6. Corneil, D.G., Olariu, S., Stewart, L.: The LBFS structure and recognition of interval graphs, in preparation; extended abstract appeared as the ultimate interval graph recognition algorithm? In: Proceedings of SODA 98, Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, San Francisco, CA, USA, pp. 175–180 (1998)
7. Corneil, D.G., Olariu, S., Stewart, L.: The LBFS structure and recognition of interval graphs. SIAM J. Discrete Math. **23**, 1905–1953 (2010)
8. Deng, X., Hell, P., Huang, J.: Linear time representation algorithms for proper circular arc graphs and proper interval graphs. SIAM J. Comput. **25**, 390–403 (1996)

fridaya@uef.fi

9. de Figueiredo, C.M.H., Meidanis, J., de Mello, C.P.: A linear-time algorithm for proper interval graph recognition. Inform. Process. Lett. **56**, 179–184 (1995)
10. Golumbic, M.C.: Algorithmic graph theory and perfect graphs. Ann. Discrete Math. **57** (2004)
11. Golumbic, M.C., Trenk, A.: Tolerence Graphs. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (2004)
12. Habib, M., McConnel, R., Paul, C., Viennot, L.: Lex-BFS and partition refinement, with applications to transitive orientation, interval graph recognition, and consecutive ones testing. Theor. Comput. Sci. **234**, 59–84 (2000)
13. Hell, P., Huang, J.: Certifying LexBFS recognition algorithms for proper interval graphs and proper interval bigraphs. SIAM J. Discrete Math. **18**, 554–570 (2005)
14. Hell, P., Shamir, R., Sharan, R.: A fully dynamic algorithm for recognizing and representing proper interval graphs. SIAM J. Comput. **31**, 289–305 (2001)
15. Hsu, W.-L.: A simple test for the consecutive ones property. J. Algorithms **43**, 1–16 (2002)
16. Malik, D.S., Sen, M.K., Ghosh, S.: Introduction to Graph Theory. Cengage Learning, New York (2014)
17. McKee, T.A., McMorris, F.R.: Topics in Intersection Graph Theory. Siam Monographs on Discrete Mathematics and Applications. Siam, Philadelphia (1999)
18. Meidanis, J., Porto, O., Telles, G.P.: On the consecutive ones property. Discrete Appl. Math. **88**, 325–354 (1998)
19. Nussbaum, Y.: Recognition of probe proper interval graphs. Discrete Appl. Math. **167**, 228–238 (2014)
20. Paul, S., Ghosh, S., Chakraborty, S., Sen, M.: Characterization and recognition of proper tagged probe interval graphs. arxiv:1607.02922 [math.CO] (2016)
21. Roberts, F.S.: Representations of indifference relations, Ph.D. Thesis, Stanford Univeristy (1968)
22. Roberts, F.S.: Indifference graphs. In: Harary, F. (ed.) Proof Techniques in Graph Theory, pp. 139–146. Academic Press, New York (1969)
23. Roberts, F.S.: On the compatibility between a graph and a simple order. J. Combin. Theory Ser. B **11**, 28–38 (1971)
24. Sheng, L.: Cycle free probe intearval graphs. Congr. Numer. **140**, 33–42 (1999)
25. Sheng, L., Wang, C., Zhang, P.: Tagged probe interval graphs. J. Combin. Optim. **5**, 133–142 (2001)
26. Sheng, L., Wang, C., Zhang, P.: On the perfectness of tagged probe interval graphs. In: Wang, C., Zhang, P. (eds.) Discrete Mathematical Problems with Medical Applications, pp. 159–163. American Mathematical Society, Providence (2000)
27. Sheng, L., Wang, C., Zhang, P.: Tagged probe interval graphs. DIMACS Technical Report 98-12 (1998)
28. Zhang, P.: Probe interval graphs and their application to physical mapping of DNA. Manuscript (1994)

fridaya@uef.fi

Evolvable Media Repositories: An Evolutionary System to Retrieve and Ever-Renovate Related Media Web Content

Marinos Koutsomichalis[1(✉)] and Björn Gambäck[2]

[1] Department of Multimedia and Graphic Arts, Cyprus University of Technology, P.O. Box 50329, 3036 Limassol, Cyprus
m.koutsomichalis@cut.ac.cy

[2] Department of Computer Science, Norwegian University of Science and Technology, Sem Saelands vei 9, 7034 Trondheim, Norway
gamback@ntnu.no

Abstract. The paper tackles the question of evolvable media repositories, i.e., local pools of media files that are retrieved over the Internet and that are ever-renovated with new, related files in an evolutionary fashion. The herein proposed method encodes genotypic space by virtue of simple undirected graphs of natural language tokens that represent web queries without employing fitness functions or other evaluation/selection schemata. Once a first population is seeded, a series of modular crawlers query the particular World Wide Web repositories of interest for both media content and assorted meta-data. Then, a series of attached intelligent comprehenders analyse the retrieved content in order to eventually generate new genetic representations, and the cycle is repeated. Such a method is generic, scalable and modular, and can be made fit the purposes of a wide array of applications in all sorts of disparate contextual and functional scenarios. The paper features a formal description of the method, gives implementation guidelines, and presents example usages.

Keywords: Genetic algorithms · Database management · Multimedia information systems · Natural language processing

1 Introduction

The concept of *evolvable media repositories* designates local pools of media files that are retrieved over the Internet and that can be regularly renovated with new files that are related to the former in some desired way. A natural approach to

M. Koutsomichalis—Work carried out when the first author was at the Norwegian University of Science and Technology supported by an ERCIM Alain Bensoussan Fellowship.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 76–92, 2019.
https://doi.org/10.1007/978-3-030-22868-2_6

fridaya@uef.fi

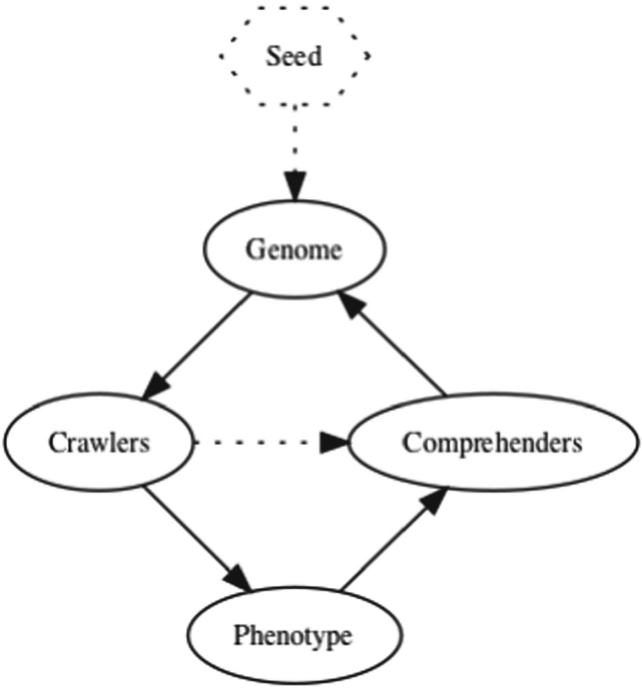


Fig. 1. The evolution cycle

the question of evolvable media repositories is to employ some evolutionary algorithm. Still, unless very specific criteria that are extrinsic to the content have to be satisfied, traditional 'survival of the fittest' approaches are of no real use here. It is both possible and often plausible to define fitness functions regarding how media perform with respect to some broader context, e.g., user preferences or some arbitrary aesthetic measure. The notion of fitness, however, is largely artificial when the retrieved media files are expected to intrinsically relate to the ones they replace in arbitrary respects, that may as well be unknown. Accordingly, the method we propose does not employ the traditional evaluation/selection, which is typically found in most evolutionary systems.

In a nutshell, the cyclic method consists of the stages shown in Fig. 1:

(i) seed a graph of natural language tokens,
(ii) employ a series of modular crawlers to query the particular Internet repositories of interest for relevant content and assorted meta-data, and
(iii) analyse the retrieved media by virtue of a series of modularly attached 'comprehenders' to generate a new graph of tokens.

A method such as the one proposed here has a series of advantages: it is generic, scalable and modular, and can be tailored to fit the needs of a wide array of applications with a demand for evolvable media repositories.

fridaya@uef.fi

Hence, it is possible (and often trivial) to extend, specialise or repurpose an implementation so that it accounts for the particular kinds of media, repositories, and (most importantly) the semantic or content-dependent associations that are of interest. It is also possible to implement systems that rely exclusively on content-dependent features, on natural language descriptions and meta-data, or any hybrid combination thereof.

The advocated genomic representation is, up to a certain extent, language agnostic, making it both theoretically possible and pragmatically feasible to encode arbitrary lexical, semantic, and symbolic information in any natural or artificial language.

The need for such a system arose concretely in the context of a media-art project concerning the real-time production of algorithmic audio, image, video, poetry, and 3D data in a generative fashion, and by virtue of manipulating and re-appropriating files retrieved from various media repositories such as YouTube and Flickr. As accounted for in Koutsomichalis and Gambäck [17], we have already employed contingent incarnations of this system to generate algorithmic audio mashups and synthetic soundscapes in real time. This particular case aside, the need for an evolvable media repository or some similar system often arises in various functional, or other contexts.

The remainder of this paper elaborates on this method and its implications. The next section introduces the background of the work and some relevant research literature. A formal generalisation of the method is presented in Sect. 3 and subsequently its implementation is discussed in Sect. 4. Section 5 then reports some results, followed by concluding remarks and directions for future work in Sect. 6.

2 Background and Related Work

Evolutionary algorithms (EAs) are typically thought of as cases of population based meta-heuristic optimisation systems. That is, they are typically meant as the means to arrive at individual (or populations of) phenotypes that represent solutions to some problem that is hard, or impossible, to solve otherwise. Evaluation and selection are almost unanimously accepted as essential steps in every EA system of sort. Still, there is more to evolution than just Darwinian or Lamarckian processes of selection and mutation [19,29], and not all cultural or natural life phenomena can be understood, or described, in terms of meta-heuristic optimisation or problem solving.

As a point of fact, there are certain scenarios where fitness functions and data analysis procedures are neither strictly relevant nor easily definable. In subareas specific to digital arts and computer music, for instance, mutation for the sake of variance, or in order to generate new and original content for the sake of it, is a very pragmatic need that cannot (or should not) always be thought of in terms of 'selection of the fittest'. In practice, nevertheless, the vast majority of art-oriented EAs rely on fitness functions of some sort, as shown, for instance, in the survey by Johnson [16] of the 2003–2011 EvoMUSART conferences.

fridaya@uef.fi

The evaluation/selection schemata that are most often encountered involve arbitrary aesthetic measures, human interaction, and corpora of real-life examples. It is debatable to what extent such solutions succeed in generating genuine artistic value in real-life contexts and are not simply justified as technological fetishes or extra-ordinariness [5,22]. Indeed, aesthetic measures often prove implicit and ill-defined, interactive systems biased or impractical, and the use of example corpora typically results in systems that imitate, rather than genuinely generate, creative behaviour.

While the question of 'artistic value' is a hard one to define in the first place, it cannot be taken for granted that it can be deterministically evaluated by either humans or machines. Accordingly, several approaches eschew or undermine the idea of fitness altogether. Biles [2], for instance, accounts for an alternative EA approach to jazz melody composition that pivots on an intelligent crossover operator. Furthermore, there are many documented cases employing 'endogenous' fitness functions, that is, functions that are implicitly defined and operate in the context of some local artificial environment rather than with respect to how the eventual artefacts evaluate against some objective or universal criteria. Such schemata do sustain a 'survival of the fittest' kind of architecture, but at the same time undermine or cast vague the very idea of evaluation to the extent that the system in question may no longer be thought of as evolving towards the 'fittest' (i.e., 'better' in any subjective sense) works of art. In Bird et al. [3], for example, a drawing robot is provided with a fitness function not meant to evaluate the resulting drawing but rather to reward local behaviour with respect to pen position and a line detector. Endogenous fitness definitions are also the norm in the case of ecosystemic evolution models [9], where an individual's fitness is not strictly defined, but instead implicitly emerges through its interactions with an environment and other individuals. Bown and McCormack [5] have accounted for several example art-oriented studies that pivot on ecosystemic evolution.

Another broad array of approaches that are generally encountered in many EA contexts and subfields is that of replacing fitness functions with intelligent critics or parallel co-evolution processes [2,8,25]. Recent advances in evolutionary and unsupervised learning (e.g., in subareas specific to Generative Adversarial Networks [15] or NeuroEvolution of Augmenting Topologies [30]) add novel understandings on what evaluation may stand for in the context of EA systems.

As far as the retrieval of media content over the Internet is concerned, text-based and meta-data driven research approaches are legion [10,24], as are more sophisticated content-based and hybrid approaches [11,14,20,21,23,32], used in various contexts and regarding media of various sorts. Notably though, pure content-dependent approaches are not easy (nor always possible) to implement when dealing with very big remote repositories that are dynamically updated by peers in real time. More importantly—and despite some promising breakthroughs involving deep learning [34]—they typically suffer from 'semantic gap' [28] related problems, so that contextual similarity with respect to higher level semantic features cannot be guaranteed.

fridaya@uef.fi

Albeit rather few, some projects have aimed to combine evolutionary and genetic approaches for multimedia content retrieval. Lai and Chen [18] and Cho and Lee [7] account for content-based image information retrieval pipelines featuring interactive genetic algorithms. Cho [6] extended this approach to musical information. Zacharis and Panayiotopoulos [35] presented *Webnaut*, a system utilising a genetic algorithm to collect and recommend Web pages. da Silva Torres et al. [27] described an image retrieval system employing genetic programming.

The above approaches, however, are largely irrelevant to the question of evolvable media repositories, both in their scope, and as far as methodology is concerned. Evolvable media repositories, the way we envisage them, call for a non-interactive approach and for a varying degree of granularity as far as what exact kinds of content-dependent features, meta-data descriptors and natural-language semantics should be taken into consideration. To the best of our knowledge, such a question is not explicitly addressed elsewhere.

3 Method

To formally describe the method illustrated in Fig. 1, let U denote the Unicode character set and $U^* = \bigcup_{i\in\mathbb{Z}^+} U_i$ the set of all finite words, and sequences of words, of arbitrary length over U. Natural language tokens, arbitrary named entities, numbers, computer code and, in general, any finite text string, as long as it is representable with the Unicode charset, is an element of U^*. Similarly, let B be the binary alphabet $\{0, 1\}$ and $B^+ = \bigcup_{i\in\mathbb{Z}^+} B_i$ the set of all non-empty finite permutations over B. Media content and, in general, arbitrary information of all sorts, as long as it can be digitally encoded and computationally represented, is an element of B^+. Then, the powerset $\mathcal{P}B^+$ denotes the set of all unique (non-ordered) combinations of one or more sequences over B^+, that is, the set of all possible pools comprising one or more digital files.

Accordingly, the entire Internet can be thought of as a subset of $\mathcal{P}B^+$ with a finite, albeit astronomically big, cardinality. Given that the World Wide Web is ever-contingent and constantly dynamically permuted by both human and machine agents, this is better expressed by means of a set-valued function of time $f\colon \mathbb{R}_{\geq 0} \to \mathcal{P}B^+$. Note, however, that while at any given time t all digital content on the Internet is $W = f(t)$, it is often more relevant to consider W as the union of all possible values of f in a period of time. Any computational process necessarily takes time during which the relevant Internet contents may or may not have changed, so that it is possible that a function mapping some domain to W (e.g., performing a series of web queries) returns some result A with $A \notin f(t_k)$ but $A \in \int_{t_0}^{t_p} f(t)$, where t_0, t_k mark the times the process started and finished, respectively, and $t_0 \leq t_p < t_k$. Whenever such rigor is unnecessary, of course, one may simply assume that $\Delta t = 0$. Hence W can be defined as:

$$W = \begin{cases} \int_{\Delta t} f(t)dt & : \Delta t > 0 \\ f(t) & : \Delta t = 0 \end{cases} \tag{1}$$

fridaya@uef.fi

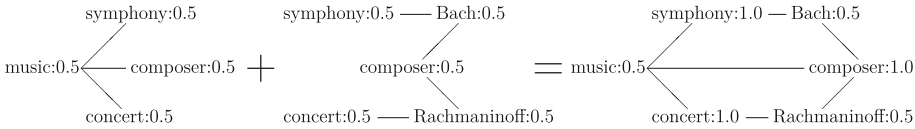


Fig. 2. Merging genomes results in new emergent associations

As mentioned in the previous section, we encode genetic material employing simple undirected graphs of (natural) language tokens. In detail, individual genomes are represented as graphs comprising arbitrary—albeit, in any practical implementation somehow limited in size—vertices of textual information. Each genome can be represented as an undirected graph (V, E), V being a set of vertices $\{v_1, v_2, \ldots, v_n\}$ with $n \in \mathbb{Z}^+$, $v_n \in U^* \wedge v_n \neq \emptyset$, and E being a (possibly empty) set of pairs $\{v_p, v_k\}$ for some $v_p, v_k \in V$ and $v_p \neq v_k$. Then, the set G^+, comprising all possible non-empty genomes, is formally defined as:

$$\mathrm{G}^+ = \Big(V, E \colon V \in U^* \wedge V \neq \emptyset, \quad E \equiv \big\{\{v_p, v_k\} \vee \emptyset \colon v_p, v_k \in V \wedge v_p \neq v_k\big\}\Big) \tag{2}$$

Note that an entire genomic population $\mathrm{G}_1, \mathrm{G}_2, \ldots, \mathrm{G}_n$ $(n \in \mathbb{Z}^+ \wedge G_n \in \mathrm{G}^+)$ can also be represented in terms of a single merger graph $P = \bigcup_{i=1}^{n} G_i$ comprising unique instances of all vertices and edges that are present in its integrals. Such a merger graph not only preserves all existent associations (edges), but also forges new, emergent, ones. This is illustrated in Fig. 2 where 'Rachmaninoff' and 'Bach' are inferred as being (indirectly) associated to 'music'.

The vertices in a genome need not be necessarily 'meaningful' natural language tokens, the various known natural language being just a subset of U^*. It is exactly because of the absence of such a restriction that the genome may properly encode and hierarchically represent arbitrary semantic, lexical and symbolic relationships that may arise naturally in different contexts and irrespective of what particular natural languages, micro-culture specific jargons, and even machine encodings might be at play. In this sense, genetic material may also encode artilects, named entities, ISBN numbers, Integrated Circuit product numbers, user IDs, indices of files, and in general anything that may make sense in some given context. The rationale behind the decision to represent the genome of individuals and of entire populations in the same way lies in the very fact that genetic information is nothing but the queries that will generate the actual population. Consider, for instance, that YouTube comprises videos associated with para-texts in all sorts of natural language, while often employing cryptic linguistic idioms and esoteric micro-cultural codes. It is both theoretically and practically possible that a genome represents associations occurring both in the context of any of these languages/idioms and among all of them. More importantly, such a genome may also encode associative relationships among the latter.

At each iteration of the algorithm, a series of set-valued 'crawlers' $\lambda_1, \lambda_2, \ldots, \lambda_n$ map genetic information to an associated phenotype—that is, to a pool of retrieved media files. Subsequently, a series of 'comprehenders' $\phi_1, \phi_2, \ldots, \phi_k$ process the phenotype, generating a series of genomes $\mathrm{G}_1, \mathrm{G}_2, \ldots, \mathrm{G}_n$ that are, eventually, merged to genetically represent a population which will take flesh on the next implementation of the algorithm. Note that $k, n \in \mathbb{Z}^+$ need not be equal. Note also that while it is often possible that associated meta-data information are directly encoded as genomes—as

fridaya@uef.fi

Algorithm 1. Evolution cycle in pseudo-code

$G \leftarrow \langle$genome for the initial population$\rangle$
$\lambda_0, \lambda_1, \ldots, \lambda_n \leftarrow$ crawlers for the repositories of interest
$\phi_0, \phi_1, \ldots, \phi_k \leftarrow$ comprehenders for the content of interest
loop
  $P \leftarrow []$
  for $i = 0$ to $n - 1$ **do**
    $P \xleftarrow{append} \lambda_i(G)$ ▷ may result in $0, 1$, or > 0 media files
  end for
  $l \leftarrow |P|$ ▷ the size of P
  $G^{'} \leftarrow []$
  for $i = 0$ to $k - 1$ **do**
    $G^{'} \xleftarrow{append} \phi_k(P[i])$ ▷ ϕ_k compatible with $P[i]$'s type
  end for
  $G \leftarrow G \bigcup G^{'}$
end loop

the dashed line in Fig. 1 indicates—we treat them here as elements of the phenotypes in all cases, both for reasons of conceptual clarity and because it is often desired that they are somehow comprehended, too, before they are encoded.

A crawler λ can be formally described as a function mapping a genotype in G^+ to a subset of W, that is, $\lambda\colon \mathrm{G}^+ \rightarrow W$. According to Eq. (1), λ may or may not return the same set of files for the very same genetic material, since its output domain is ever-contingent and dependent on time. At each iteration of the algorithm, the crawlers produce a series of sets $M_1, M_2, \ldots, M_n$ that are all elements of W. The union P of all the output sets constitutes the phenotype of some genome G and represents an evolved generation comprising $|P|$ media files. P is formally defined as:

$$P = \bigcup_{i=1}^{n} M_i = \bigcup_{i=1}^{n} \lambda_i(\mathrm{G}) \tag{3}$$

Given Eq. (1) and the definition of λ, it follows that $P \in W \equiv \int_{\Delta t} f(t)dt$.

Genome mutation occurs with respect to the aggregate effect of a series of 'comprehender' functions that map patterns from phenotypic to genotypic space. A comprehender function can be defined as $\phi_\Lambda\colon B^+ \cap \Lambda^+ \rightarrow G^+$. Here, Λ^+ denotes all media content that satisfy a set of conditions (e.g., that they are audio or image files) and ϕ_Λ some comprehender designed to *understand* Λ kind of content. Then, the new, mutated, genome can be defined as the union of all $\phi_\Lambda(P_i)$ for all $P_i \in P$.

$$\mathrm{G}^{'} = \bigcup \phi_\Lambda(P) \tag{4}$$

Combining Eqs. 1–4, the method, in its entirety, can be formalised as:

$$\mathrm{G_n} = \begin{cases} \bigcup \phi_\Lambda(\bigcup_{i=1}^{n} \lambda_i(\mathrm{G_{n-1}})) & : n \in \mathbb{N}_{>1} \\ \langle S \rangle & : n = 1 \end{cases} \tag{5}$$

where $\langle S \rangle$ denotes the 'seed', i.e., the genetic material of the initial population. Algorithm 1 presents a minimal description of Eq. (5) in pseudo-code.

fridaya@uef.fi

4 Implementation

The proposed method is meant to be generic and modular, so that it is of potential use to disparate functional and research contexts. For any actual implementation, a number of design decisions should be taken with respect to the particular requirements at play:

1. Should the genotypic space G^* be somehow specialised (e.g., by means of additional linguistic constraints or with respect to its size and structure)?
2. What kinds of media content should the program deliver and what particular media repositories should be accessed?
3. What kinds of meta-data should the program consider?
4. What comprehenders are needed?
5. Are there any special rules regarding the manipulation of genetic material?
6. How should network/API failures be dealt with?
7. Should there be any sanity checks, normalisation tasks or other media-file manipulations?
8. What is the desired lifespan of each individual in a phenotype?
9. How exactly should the phenotype be locally stored and administrated?
10. How to safe-proof the system from possibly malicious content?

There are no right answers, of course; it all depends on the context. The evolvable media repository described below has been implemented in the context of a real-life artistic project—still a work in progress—and follows the requirements given by the project. The project concerns a phenotype of all audio, image, video, text, 3D data, and associated meta-data and para-texts, the production of several continuous media streams in parallel and, most importantly, an eventual system expected to run unattended for prolonged periods of time. The last requirement suggests that retrieved media should remain local for at least as long as there are enough new instances to replace them without affecting performance. It also implies that in case of some failure (e.g., of the network), the evolution cycle should not seize up, but silently keep on until the failure is addressed.

Another idiosyncrasy of the implementation is that genomes are represented as graphs of *weighted* vertices, so that vertices may be ranked with respect to both context and number of occurrences, and so that they can be constrained to their most important integrals at each iteration of the algorithm. As noted in the Introduction, we have already successfully employed experimental versions of this implementation in order to generate algorithmic mashups and synthetic soundscapes in real time.

Figure 3 illustrates the most important system components and their cross-interactions. There are crawlers for YouTube (audio, video), Flickr (image), Shootka (speech), FreeSound (audio), SoundCloud (music), MLDb (lyrics), Thingiverse (3D data), Wikipedia (text), ConceptNet (lemmas, phrases), and WordNet (lemmas), retrieving content as well as associated tags, text descriptions and user comments when available.

All crawlers inherit from an `AbstractCrawler` class and are required to implement an iterator over the results, as well as methods to query for, retrieve, and post-process content, and retrieve meta-data and associated para-texts. Implementing such methods is of varying difficulty, depending on the particular APIs and the authorisation specifics at play.

fridaya@uef.fi

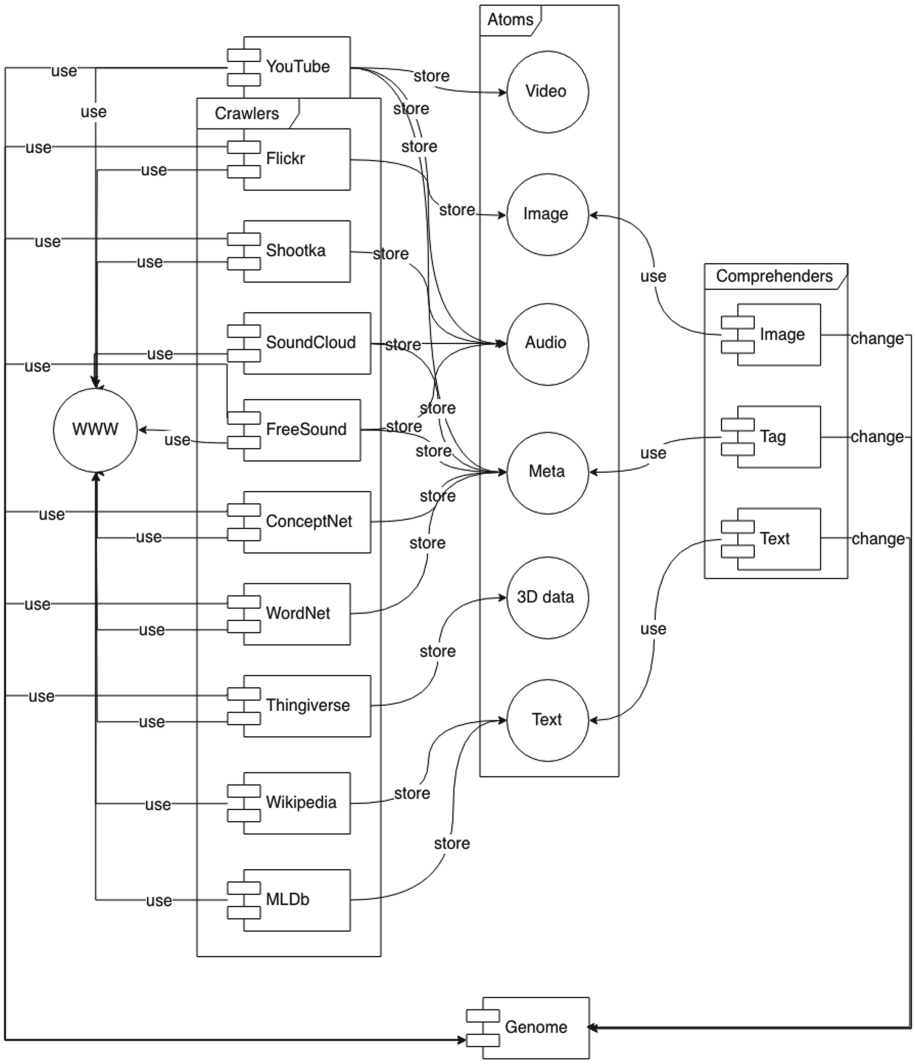


Fig. 3. Overview of a realised implementation

The system currently features comprehenders for text, keywords (tags), and images only. Attaching more sophisticated comprehenders performing, e.g., music and video understanding is to be considered for a future version. All comprehenders inherit from an `AbstractComprehender` class and are required to implement methods to prepare and understand data, and to take care of any cleaning tasks. Their eventual output is a graph of weighted string tokens, which is typically completely interconnected, that is:

$$G = \left(V, \binom{V}{2}\right) \tag{6}$$

fridaya@uef.fi

The `TagComprehender` is the most trivial of all. It essentially generates a graph comprising all available tags as individual entries having a nominal weight of 0.6. The `TextComprehender` is more sophisticated, relying on the Rapid Automatic Keyword Extraction [26] algorithm for language understanding, and resulting in a series of ranked word sequences that summarises the gist of the original text. In this case the output comprises the resulting word sequences with their extracted weights as long as they satisfy a minimum score. The Image comprehender draws on the Inception-v3 [31] network trained on the ImageNet [12] LSVRC-2012 challenge data set. Here, too, the comprehender returns a graph with all identified properties that surpass a minimum score along with all their associated weights.

Following Eq. (4), the new genotype is computed as the union of all individual genomes that have been generated for each file in the phenotype. Given that all constituent graphs have been fully interconnected, their union comprises string queries that are connected with

(i) all other queries generated by the same comprehender for the same file, and
(ii) all other queries that any other comprehender has associated as relevant to this particular query encountered in the context of some other file.

Iteration over some genotype G is addressed employing a `GraphIterator` object, the constructor of which generates an ordered sequence of queries Σ:

$$\Sigma = \{u_n : \forall u_n \in G \wedge ([u_n - u_{n-1} \wedge w(u_n) \geq w(u_{n-1})] \vee w(u_n) > w(u_p) \forall p < n)\} \quad (7)$$

where $w\colon U^* \rightarrow Q_{[0,1]}$ is a function returning the weight of each vertex u_n, while $u_n - u_{n-1}$ denotes that the two vertices are adjacent. In simpler words, Eq. (7) suggests that the order of vertices in Σ is such that every next vertex u_n is either the next highest ranking vertex that is connected to u_{n-1} or the next highest ranking vertex that has not already appeared.

Algorithm 2 (below) shows how this order is generated. Crawlers respect the order and draw as many queries as needed until there are either no entries left or the maximum threshold of files to download (a user-specified number) is met. More specifically, crawlers attempt to download as many media files as possible using each query in turn and only draw on the queries following if the threshold is not met.

Algorithm 2. Iteration over a genome in pseudo-code

```
G ← Genome
S ← []
while vertices left in G do
    N ←higher ranking node in G
    S ←append— N
    remove N from G
    C ← all vertices connected to N                ▷ possibly also indirectly
    while vertices left in C do
        D ←higher ranking node in C
        S ←append— D
        remove D from G
    end while
end while
```

fridaya@uef.fi

5 Results

It is important to emphasise that the kind of results one may expect can be, up to a certain extent, engineered to match the variety, veracity, and degree of randomness desired. For instance, by means of bespoke comprehenders it is possible to end up with long, short, or even single-word tokens. Longer ones typically result in more relevant media files, but also suffer many failures, while single-word ones most likely succeed all the time, but often produce irrelevant or even haphazard results.

More than simply a question of quantity, this is a question of quality, too. Public user comments, for instance, often tend to regard all sorts of irrelevant topics, and while exploring them might be desired in some particular context, it may introduce all sorts of problems in others.

Furthermore, by means of carefully selected and designed crawlers one may not simply constrain the system to search for particular kinds of content in those repositories that may deliver it, but also refine the ways in which queries themselves are manipulated and

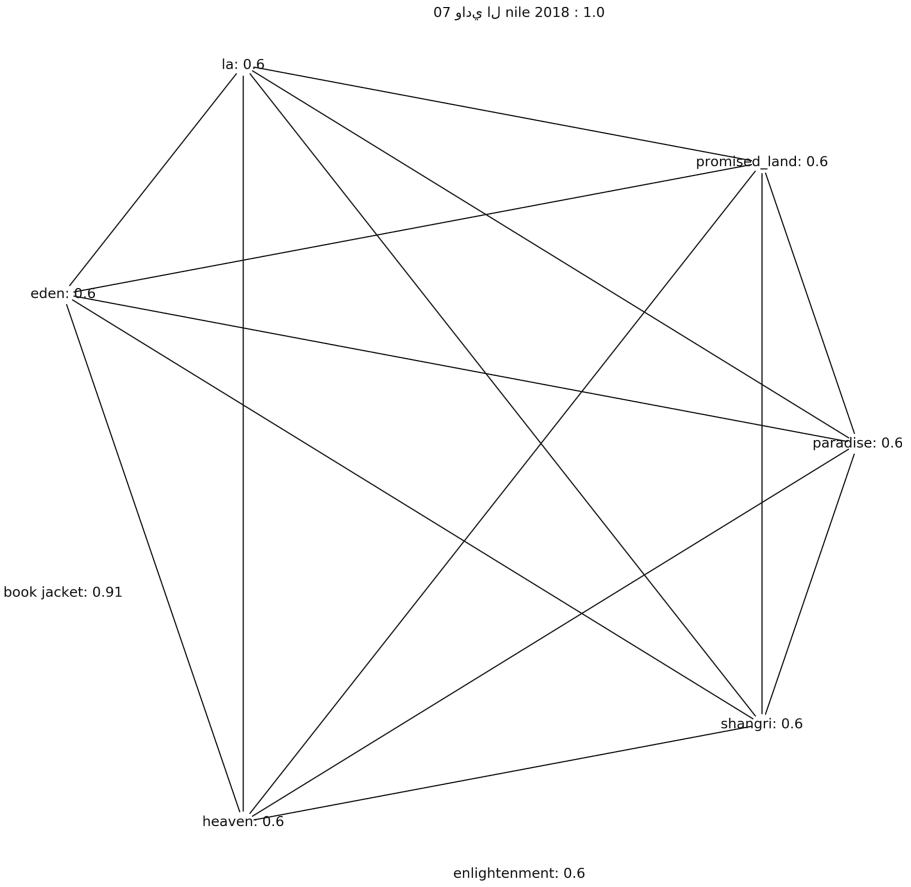


Fig. 4. Genome of a minimal system for $\langle S \rangle = nirvana$ after one iteration

fridaya@uef.fi

interpreted. The maximum number of files demanded for each crawler is also an important parameter that affects the size of the resulting genome as well as the percentage of its contents that will be utilised as queries, since crawlers stop once the desired number of files has been satisfied.

Finally, it should be noted that it is both possible, and in certain cases advisable, to manipulate the genome after each iteration, e.g., filtering out ambiguous terms, 'small' words, numbers, or similar.

For instance, a setup utilising crawlers for YouTube (drawing both user comments and videos), Flickr, WordNet, and MLDb, and demanding just one file per iteration results in the genome illustrated in Fig. 4 for $\langle S \rangle = nirvana$, after one iteration. Yet, a few more iterations later, the genome only features a couple of long phrases in the Tamil language and eventually dies out soon afterwards, since nothing is retrieved for those particular phrases.

A setup featuring crawlers for YouTube, Wikipedia and ConceptNet, with $\langle S \rangle = transcendence$, and again demanding just one file per crawler, results in the following vertices after one iteration (ordered by the highest to the lowest ranking ones):

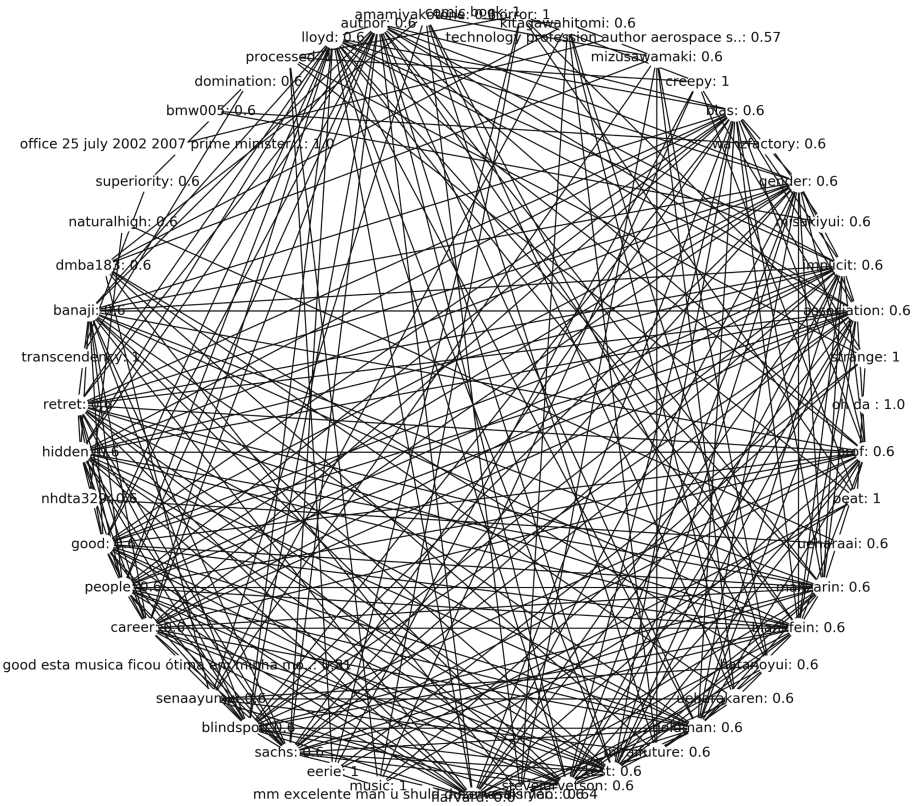


Fig. 5. Genome of a system for $\langle S \rangle = transcendence$ after one iteration

fridaya@uef.fi

‘domination’
‘office 25 july 2002 2007 prime minister atal bihari vajpayee manmohan singh vice president krishan kant bhairon shekhawat preceded’
‘superiority’
‘transcendency’
‘technology profession author aerospace scientist award bharat ratna 1997 padma vibhushan 1990 bhushan 1981 signature website abdulkalam’

After the second iteration the population becomes:

அப்துல் கலாம் (text in Tamil: ‘Abdul Kalam’)
‘musician matt lane assistant engineer gililland howie weinberg audio mastering’

In this particular case the resulting media after the second iteration are the Wikipedia entries for India’s former president Abdul Kalam and for Pantera’s song *Cowboys from Hell*, the official trailer for the *Transcendence* Hollywood film, a video concerning Abdul Kalam’s death, and a religious video featuring a Muslim prayer and text in both Tamil and Urdu. This last video did not have any comments and neither Wikipedia nor ConceptNet returned anything useful for the last genome so the evolution cycle halted immediately and abruptly there.

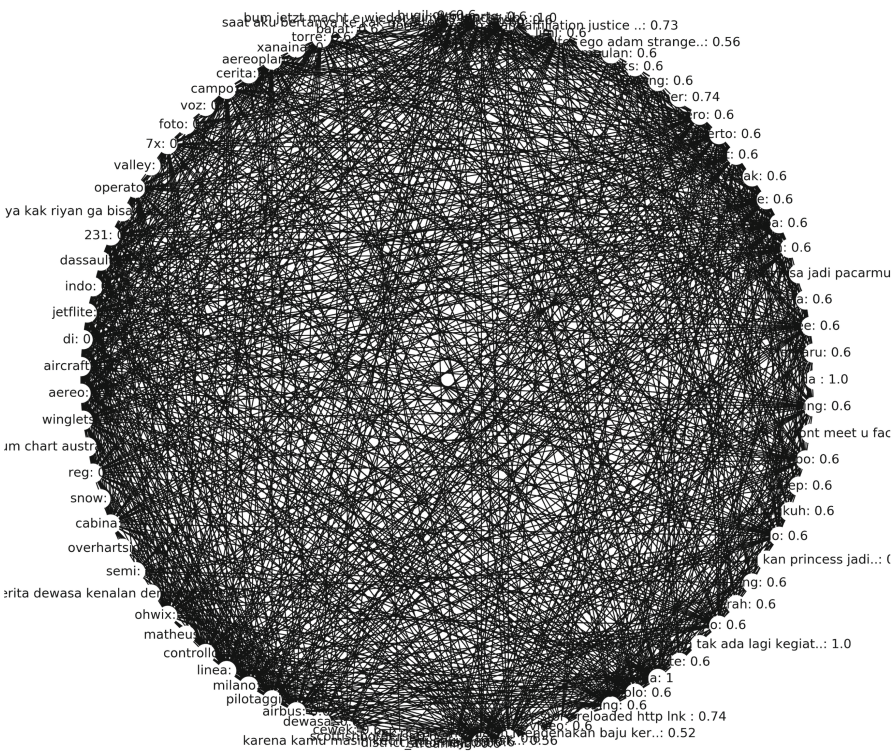


Fig. 6. The previous genome ($\langle S \rangle = \textit{transcendence}$) after the second iteration

fridaya@uef.fi

Systems that only survive for a few iterations are not necessarily useless. They may only return a handful of files; however, these tend to be highly relevant to each other and in certain contexts this is exactly what is needed. More importantly, such systems are much easier and more intuitive to tinker with and fine-tune. Given that counter measures against a population dying out prematurely are trivial to implement[1] such set-ups can be made to even fit situations calling for more than just a handful of files. Still, in some cases inexhaustible populations, dramatic mutations, and higher degrees of ambiguity/randomness are needed. The requirements for our particular implementation fall into this category. Utilising all available crawlers and comprehenders mentioned in the previous section, and demanding a varying number of 1–5 files from each crawler, the genomic structure is massively complex already after one iteration (Fig. 5).

The resulting genome seems to feature tokens in many different languages, esoteric jargon, 'cryptic' non-sensical strings, several 'small' words such as 'oy', '7x', or 'cn', and longer strings of tokens comprising several words and often numbers. This breadth of information guarantees that the population will not die out any time soon. The genome is, indeed, already many orders more complex in the very next cycle, as illustrated in Fig. 6. It should, however, be noted that despite such a complexity, it is still possible to deliver results that are, up to a certain extent, very relevant to each other. As explained in the previous section, in our implementation crawlers both respect associative chains and give preference to the higher ranking vertices. Accordingly, it is both possible and easily enforced to assign very high rankings to just a handful of preferred tokens (e.g., long and descriptive strings, or those returned by some particularly intelligent crawler), so that, as long as they are often successful, it is only them and their offspring that are used. And in case they altogether fail temporarily or more permanently, there will always be a huge pool of subsidiary entries in the population to draw queries from.

The last point explains why, despite the complexity of the graph shown in Fig. 6, the resulting phenotype demonstrates some limited self-similarity. In detail, the resulting files comprise: three 3D models (two having the descriptor "Da Vinci" in their names and the third named "Eye Of Agamotto Doctor Strange"); two spoken fragments of "oh" from Shootka that are meant to indicate "surprise" and "realization", respectively; a spoken Japanese phrase which features a prominent "oh" phoneme; three distinct music compositions (all sharing a toy-piano sound, similar structure, and somewhat similar melodic lines); a free-style rap solo in Brazilian Portuguese, also featuring vocal percussion sounds; the official trailer for the film *Transcendence*; a trailer for the film *Don Q: Oh You The Plug*; a video-clip for the song *Transcendence (Orchestral)* by Lindsey Stirling; three Japanese porn-related images; a few more images showing: an airplane, a man in front of a snowball, an over-weighted teenage girl, three conference attendees one of which is of Asian origin, a forest landscape, two people of Latino origin eating, and a favela in Rio de Janeiro; the lyrics of three songs entitled *bums bums bums bums*, *tenebrae*, and *ain't it strange*; and, finally, Wikipedia entries concerning Abdul Kalam, "Algebraic element", and the soundtrack of the film *Almost Alice*.

There are at least four explicit semantic traits herein, broadly relevant to Japan, Brazil, Hollywood films, and the sound "oh", respectively. A more carefully scrutiny, however, reveals more implicit associations, concerning, for instance, the structure and melodic content of the music compositions retrieved, as well as the content of the

[1] The program could, e.g., be instructed to start a new cycle using some closely related term of the original seed, or to continue from the last 'healthy' genome, but this time employing additional 'auxiliary' and more tolerant crawlers.

fridaya@uef.fi

lyrics—e.g. the German song *bums bums bums bums*, by J.B.O., is from the album *Sex, Sex, Sex* and features 24 occurrences of "oh", while Patti Smith's *ain't it strange* features 41.

6 Conclusion and Future Work

The paper has attempted to address the question of evolvable media repositories, that is, local pools of media files that are retrieved over the World Wide Web and that are ever-renovated with new, related ones in an evolutionary fashion. The proposed architecture is characterised by a number of idiosyncrasies/novelties. Genotypic space is encoded by means of graphs of unicode characters. In this way it is both theoretically possible and pragmatically plausible to represent queries and cross-associations forged in an arbitrary natural language, as well as regarding micro-culture specific idioms, esoteric jargons, machine-oriented encodings, etc. Then, comprehenders and crawlers of any sort can be modularly attached to the system so that it becomes aware of any particular media, repositories and content-based features of interest.

The previous section gave some concrete examples of how the method performs in a given context. It is impossible to generalise regarding the system's overall performance, since it is largely dependent on the particular requirements and implementation specifics, and since there is ample space for *ad hoc* specialisation and fine-tuning. The proposed modular design enables implementors to realise purely text-based and meta-data driven systems, purely content-based retrieval systems pivoting on intelligent feature extraction, and any hybrid combination thereof, with respect to the specific crawlers and comprehenders at play. Then, it is possible to further manipulate the genomic population by means of various filtering or normalisation processes, and in this way guarantee more or less explicit associative chains. More importantly, it seems possible to sustain very big and highly complex populations, but at the same time only utilise a very small fraction of them that is optimised with respect to some particular context. In this way one may carefully engineer systems having the desired degrees of veracity, variety, unpredictability, and semantic coherence.

As a general rule of thumb, both the performance and the particular characteristics of any implementation should be thought of with respect to

(a) what repositories should be used and what kinds of meta-data and para-texts should be taken into consideration;
(b) what particular kinds of language processing is best suitable for processing meta-data/para-texts to generate appropriate genomic material;
(c) the choice of what other comprehenders, if any, should be employed and how they should be best designed to both detect the desired features and somehow blunt the semantic gap; and
(d) whether there should be any normalisation, filtering, sanity checks, or other constraints applied to the genome itself between subsequent iterations.

Given these constraints, the herein described system may deliver ever-renovating local media repositories made to fit a wide array of requirements.

Having formalised a generic, modular and scalable method and already having utilised a few experimental implementations in order to generate audio mashups in real-time, we are currently working towards implementing specialised incarnations that we can modularly attach to more complicated multi-media mash-up systems and synthesizers. *Inter alia*, such a task necessitates that the resulting implementations are

fridaya@uef.fi

robust and adequately stable for public exhibition settings. To boot, we plan to experiment with more sophisticated comprehenders designed to 'understand' arbitrary video, music, audio and 3D data. Although there are relevant research resources zooming in on particular subdomains [1,4,13,33], implementing such comprehenders is both a non-trivial task and an open research question.

References

1. Ankerst, M., Kastenmüller, G., Kriegel, H.P., Seidl, T.: 3D shape histograms for similarity search and classification in spatial databases. In: International Symposium on Spatial Databases, pp. 207–226. Springer, Hong Kong, China, July 1999
2. Biles, J.A.: Autonomous GenJam: eliminating the fitness bottleneck by eliminating fitness. In: The 2001 GECCO Workshop on Non-routine Design with Evolutionary Systems, San Francisco, p. Paper 4, July 2001
3. Bird, J., Husbands, P., Perris, M., Bigge, B., Brown, P.: Implicit fitness functions for evolving a drawing robot. In: Applications of Evolutionary Computation: EvoWorkshops 2008, pp. 473–478. Springer, Heidelberg (2008)
4. Borges, P.V.K., Conci, N., Cavallaro, A.: Video-based human behavior understanding: a survey. IEEE Trans. Circuits Syst. Video Technol. **23**(11), 1993–2008 (2013)
5. Bown, O., McCormack, J.: Taming nature: tapping the creative potential of ecosystem models in the arts. Digit. Creativity **21**(4), 215–231 (2010)
6. Cho, S.B.: Emotional image and musical information retrieval with interactive genetic algorithm. Proc. IEEE **92**(4), 702–711 (2004)
7. Cho, S.B., Lee, J.Y.: A human-oriented image retrieval system using interactive genetic algorithm. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **32**(3), 452–458 (2002)
8. Colton, S.: Automatic invention of fitness functions with application to scene generation. In: Workshops on Applications of Evolutionary Computation, pp. 381–391. Springer (2008)
9. Conrad, M., Pattee, H.: Evolution experiments with an artificial ecosystem. J. Theor. Biol. **28**(3), 393–409 (1970)
10. Cuenca-Acuna, F.M., Nguyen, T.D.: Text-based content search and retrieval in ad-hoc P2P communities. In: International Conference on Research in Networking, pp. 220–234. Springer (2002)
11. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: ideas, influences, and trends of the new age. ACM Comput. Surv. **40**(2), 5:1–5:60 (2008)
12. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: ImageNet: a large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009, pp. 248–255. IEEE (2009)
13. Fu, Z., Lu, G., Ting, K.M., Zhang, D.: A survey of audio-based music classification and annotation. IEEE Trans. Multimedia **13**(2), 303–319 (2011)
14. Geetha, P., Narayanan, V.: A survey of content-based video retrieval. J. Comput. Sci. **4**(6), 474–486 (2008)
15. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Advances in Neural Information Processing Systems, pp. 2672–2680 (2014)
16. Johnson, C.: Fitness in evolutionary art and music: what has been used and what could be used? Evolutionary and Biologically Inspired Music, Sound, Art and Design, pp. 129–140 (2012)

fridaya@uef.fi

17. Koutsomichalis, M., Gambäck, B.: Algorithmic audio mashups and synthetic soundscapes employing evolvable media repositories. In: 6th International Workshop on Musical Metacreation, Salamanca, Spain (2018)
18. Lai, C.C., Chen, Y.C.: A user-oriented image retrieval system based on interactive genetic algorithm. IEEE Trans. Instrum. Meas. **60**(10), 3318–3325 (2011)
19. Laland, K.N., Odling-Smee, J., Feldman, M.W.: Niche construction, biological evolution, and cultural change. Behav. Brain Sci. **23**(1), 131–146 (2000)
20. Lew, M.S., Sebe, N., Djeraba, C., Jain, R.: Content-based multimedia information retrieval: state of the art and challenges. ACM Trans. Multimed. Comput. Commun. Appl. **2**(1), 1–19 (2006)
21. Liu, Y., Zhang, D., Lu, G., Ma, W.Y.: A survey of content-based image retrieval with high-level semantics. Pattern Recogn. **40**(1), 262–282 (2007)
22. McCormack, J.: Open problems in evolutionary music and art. In: Applications of Evolutionary Computing, pp. 428–436 (2005)
23. Mitrović, D., Zeppelzauer, M., Breiteneder, C.: Features for content-based audio retrieval. Adv. Comput. **78**, 71–150 (2010)
24. Nack, F., van Ossenbruggen, J., Hardman, L.: That obscure object of desire: multimedia metadata on the web, Part 2. IEEE MultiMedia **12**(1), 54–63 (2005)
25. Romero, J., Machado, P., Santos, A., Cardoso, A.: On the development of critics in evolutionary computation artists. In: Workshops on Applications of Evolutionary Computation, pp. 559–569. Springer (2003)
26. Rose, S., Engel, D., Cramer, N., Cowley, W.: Automatic keyword extraction from individual documents. In: Text Mining: Applications and Theory, pp. 1–20 (2010)
27. da Silva Torres, R., Falcão, A.X., Gonçalves, M.A., Papa, J.P., Zhang, B., Fan, W., Fox, E.A.: A genetic programming framework for content-based image retrieval. Pattern Recognit. **42**(2), 283–292 (2009). Special issue on Learning Semantics from Multimedia Content
28. Smeulders, A.W., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. **22**(12), 1349–1380 (2000)
29. Smith, J.M., Szathmary, E.: The Major Transitions in Evolution. Oxford University Press, Oxford (1997)
30. Stanley, K.O., Miikkulainen, R.: Evolving neural networks through augmenting topologies. Evol. Comput. **10**(2), 99–127 (2002)
31. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the Inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016)
32. Tangelder, J., Veltkamp, R.: A survey of content based 3D shape retrieval methods. In: Proceedings Shape Modeling Applications, pp. 145–156, June 2004
33. Vishwakarma, S., Agrawal, A.: A survey on activity recognition and behavior understanding in video surveillance. Vis. Comput. **29**(10), 983–1009 (2013)
34. Wan, J., Wang, D., Hoi, S.C.H., Wu, P., Zhu, J., Zhang, Y., Li, J.: Deep learning for content-based image retrieval: a comprehensive study. In: Proceedings of the 22nd ACM International Conference on Multimedia, pp. 157–166. ACM (2014)
35. Zacharis, N.Z., Panayiotopoulos, T.: Web search using a genetic algorithm. IEEE Internet Comput. **5**(2), 18–26 (2001)

fridaya@uef.fi

Ensemble of Multiple Classification Algorithms to Predict Stroke Dataset

Omesaad Rado(✉), Muna Al Fanah, and Ebtesam Taktek

Department of Computer Science,
University of Bradford, Bradford BD7 1DP, UK
{o.a.m.rado, m.m.s.alfanah,
e.a.m.taktek}@bradford.ac.uk

Abstract. Machine learning algorithms became popular in many domains including applications in healthcare. However, in some cases, the datasets classifiers perform poorly due to several reasons. Studies have shown that combining classifiers may help improve performance and obtain better outcomes. Ensemble approach is a technique of combining two or more algorithms to make a robust system for predictions from all the base learners. Ensemble approach works based on supervised learning algorithms in which predictions of various learning algorithms are combined to make an estimation. The simplest type of ensemble learning is to train the base algorithms on random subsets of original data set and after that make a vote for by counting the most common classifications or by computing a form of averaging the predictions of the base algorithms. In this paper, various classifiers have been applied and compared for effective diagnosis of the Stroke data set. The Stroke dataset is used to demonstrate the effectiveness of the Ensemble approach for obtaining good predictions. Experimental results show that classifier Ensemble produces better prediction accuracy.

Keywords: Machine learning · Classification · Ensemble approach

1 Introduction

Classification is one of the main techniques in the field of machine learning. The machine learning model describes the output of the algorithm on certained with data. Thus, this model is used for making predictions. This algorithm can be any machine learning algorithm such as logistic regression, decision tree or neural networks, that accepts labelled data for training. These algorithms, when used as inputs of ensemble methods, are called base models [10].

According to WHO [2], the stroke is one of the common leading causes of death. In this study, Healthcare dataset on the stroke is analysed [3] to predict whether patients have stroke or not.

In this paper an ensemble of Support Vector Machine (SVM) [5], Decision trees (DT) [9, 11] and k-Nearest Neighbours (KNN) [4, 6] as base learners are studied. Some classifiers reported in the literature have consistently lower generalization error than others [12].

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 93–98, 2019.
https://doi.org/10.1007/978-3-030-22868-2_7

fridaya@uef.fi

Das [13] presented a comparative study was performed on four classification methods for diagnosing of the Parkinson's disease. namely, Neural Networks, DMneural, Regression and Decision Tree. Moreover, the performance of the used classifiers was assessed by using different evaluation methods. Yijing et al. [14] proposed an adaptive multiple classifier system named of AMCS to cope with multi-class imbalanced learning. Three components of feature selection, resampling and ensemble learning are considered. five base classifiers and five ensemble rules are applied to construct a selection pool. Another study was also conducted in [15], The system utilized is constructed of an ensemble of Support vector machines (SVM). Methods of the weighted ensemble of SVM, row & column based SVM ensemble and channel selection with optimized SVM's are used. Bock and Poel [16] investigated two rotation-based ensemble classifiers for modeling techniques for customer churn prediction. RotBoost combines Rotation Forest with AdaBoost and then applied on four real-life customer churn prediction. Rotation Forest and RotBoost were compared to a set of well-known benchmark classifiers. Li et al. [17] presented a study of spectral-spatial hyperspectral image classification method by designing hierarchical subspace switch ensemble learning algorithm.

The remaining section of this paper is organised as follows. In Sect. 2, we describe the methodology used for data analysis. Section 3 looks at analysing the results of this study. Finally, areas for the conclusion and future work are identified in Sect. 4.

2 Methodology

In this section, we present a brief background of machine learning techniques and ensemble methods.

2.1 Machine Learning Techniques

Machine learning techniques have been referred to the as prediction based on the models built from existing data. Data mining and machine learning fields are used to discover patterns in the dataset. Classification is a part of the supervised learning technique which is the process to categorize a given data set based on classes using classification algorithms. In this work, we applied the following classification algorithms:

- Support Vector Machine is a supervised machine-learning algorithm for classification via the use of a kernel function [4, 9].
- KNN algorithm is used to predict the class based on the nearest training examples in the space [5].
- DT works based on split point measures to choose split data points that give the best partition of the data. Common measures of a split point are Entropy and Gini index [8, 11].

fridaya@uef.fi

2.2 Ensemble Techniques

The typed of ensemble methods that are used to combine several machine learning techniques are bagging, boosting, and stacking.

2.2.1 Bagging

It is a sampling technique to build multiple models of the same type from different observation samples of the original dataset, e.g. Bootstrap aggregation. In the bagging strategy, the first stage includes creating multiple models which generated using the same algorithm with random sub-samples (bootstrap) of the original dataset.

The second stage is for aggregating the generated models by using well-known methods such as voting and averaging. These methods are used to combine outputs of the classifiers [12].

Averaging means computing the average of predictions from a group of models for regression problem or predicting probabilities for the classification problem [6, 10].

Majority vote refers to choosing maximum vote from multiple model predictions of a classification problem. Each model creates a prediction for each sample of the test set and the final prediction will be the one that receives the biggest number of the votes.

Weighted average means different weights to predictions from multiple models then computing the average of models output. In this method, the prediction of each model is multiplied by the weight and then their average is calculated [1, 7].

2.2.2 Boosting

It is a sequential technique for building the same type of multiple models each of which learns to fix the prediction errors of the previous model, e.g. XGBoost, AdaBoost. Boosting makes an ensemble by training each model with the same dataset and the weights of instances are adjusted according to the error of the last prediction [1].

2.2.3 Stacking

It is an ensemble method where the models are combined using different types of machine learning in layers, each model passes predictions to the layer above it [7].

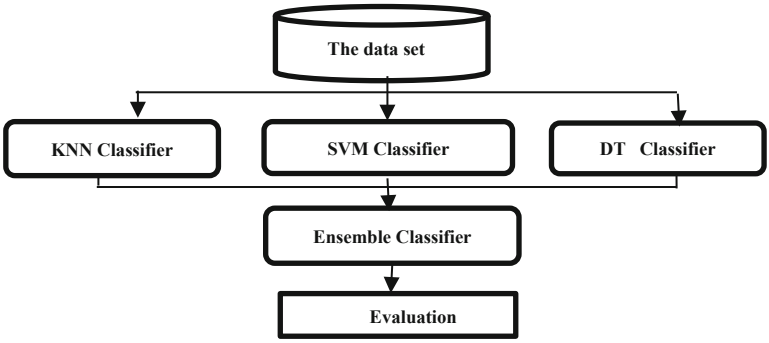


Fig. 1. The model structure

fridaya@uef.fi

3 Results and Discussion

3.1 Dataset

The dataset used in this study presented in Table 1 and clearly described all the features in the data sets. The training data has 43400 instances with 12 attributes, 18601 instances for testing. Table 1 shows the description of the columns of the data. Before building the classification models, handling missing data method and balanced data Rose method were applied on this dataset in pre-processing. The data for this study were retrospectively collected from [3].

Table 1. Data description

N	Feature name	Data type	Description
ID	Numeric	id-Patient	ID
Gender	Categorical	Male, female and other	Gender
Age	Numeric	Age of patient	Age
Hypertension	Categorical	0 - no hypertension, 1 - suffering from hypertension	Hypertension
Heart_disease	Categorical	0 - no heart disease, 1 - suffering from heart disease	Heart_disease
Ever_married	Categorical	Yes/no	Ever_married
Work_type	Categorical	Type of occupation	Work_type
Residence_type	Categorical	Area type of residence (urban/rural)	Residence_type
Avg_glucose_level	Numeric	Average glucose level (measured after meal)	Avg_glucose_level
Bmi	Numeric	Body mass index	Bmi
Smoking_status	Categorical	Patient's smoking status	Smoking_status
Stroke	Categorical	0 - no stroke, 1 - suffered stroke	Stroke

3.2 Results

We used a different classifier and compared outputs with them. The classifiers are SVM, K-NN, Decision trees, Random Forests. Fig. 1 shows the model structure of employing ensemble methods to build multiple base classifiers.

The results obtained from applying the used classifiers to the Stroke dataset are summarised in Table 2. Three performance indicators are used to compare the outcomes of the used methods with respect to the accuracy, MSE, Precision, and F-Measure of different methods. KNN, DT, and SVM as single classifiers. RF, AdBoost and Stacking as ensemble multiple models [12]. In summary, these results show that Stacking technique has the highest average accuracy of 87.58%. Closely the following stacking is Bagging RF with an average accuracy of 86.63%.

fridaya@uef.fi

Table 2. Different ensemble methods for the stroke dataset

Method	Accuracy %	MSE	Precision	F-measure
KNN	84.58	0.1542	0.846	0.846
SVM	77.29	0.2271	0.777	0.772
Decision trees/C4.5	86.10	0.1732	0.861	0.861
Random forest/bagging	86.63	0.178	0.867	0.866
AdaBoost	82.43	0.825	0.825	0.824
Stacking	87.58	0.1912	0.876	0.876

The accuracy of the SVM algorithm was 77.29% as a single classifier but it increased with ensemble methods. RF gave a good accuracy of 86.63%, while AdaBoost gave 82.43% accuracy.

An experimental study and comparison of ensemble learning strategies and measuring their accuracy and performance. This study is to evaluate the performance of some ensemble learning methods with well-known classification algorithms (KNN, SVM, C4.5 and Random Forests) on the Stroke dataset.

4 Conclusion and Future Work

Although of the fact that ensemble methods can help for delivering results with high accuracy, it is frequently not preferred in the domains where interpretability is progressively essential. In any case, the effectiveness of these strategies is obvious, and advantages in fitting applications can be big in fields such as healthcare where even the littlest measure of progress in the accuracy of machine learning algorithms can be important. The Stroke dataset, which is a combination of numerical and categorical features are used in this study as it is common and widely exist in healthcare data.

Ensemble learning approach has been applied for improving the performance of ML algorithms for the Stroke data. Ensemble learning approach gives a better prediction accuracy in detecting whether the patient has a stroke or not.

References

1. Yang, P., Hwa Yang, Y., Zhou, B.B., Zomaya, A.Y.: A review of ensemble methods in bioinformatics. Curr. Bioinform. **5**(4), 296–308 (2010)
2. WHO—Stroke, Cerebrovascular accident. WHO (2015)
3. Healthcare Dataset Stroke Data—Kaggle. https://www.kaggle.com/asaumya/healthcare-dataset-stroke-data. Accessed 10 Aug 2018
4. Wu, X., et al.: Top 10 algorithms in data mining. Knowl. Inf. Syst. **14**, 1–37 (2008)
5. Kantardzic, M.: Data mining: concepts, models, methods, and algorithms
6. Zaki, M.J., Meira Jr., W.: Data Mining and Analysis: Fundamental Concepts and Algorithms
7. Witten, I.H., Frank, E., Hall, M.A.: Data mining. Data Min. **36**(5), 51–52 (2011)
8. Quinlan, J.R., Ross, J.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, Burlington (1993)

fridaya@uef.fi

9. Burger, S.: Introduction to machine learning with R: rigorous mathematical analysis
10. Mahdavi-Shahri, A., Houshmand, M., Yaghoobi, M., Jalali, M.: Applying an ensemble learning method for improving multi-label classification performance. In: Proceedings - 2016 2nd International Conference of Signal Processing and Intelligent Systems, ICSPIS 2016, pp. 1–6 (2017)
11. Singh, A., Thakur, N., Sharma, A.: A review of supervised machine learning algorithms. In: 2016 3rd International Conference on Computing for Sustainable Global Development, pp. 1310–1315 (2016)
12. Breiman, L.: Random Forests (2001)
13. Das, R.: A comparison of multiple classification methods for diagnosis of Parkinson disease. Expert Syst. Appl. **37**(2), 1568–1572 (2010)
14. Yijing, L., Haixiang, G., Xiao, L., Yanan, L., Jinling, L.: Adapted ensemble classification algorithm based on multiple classifier system and feature selection for classifying multi-class imbalanced data. Knowl.-Based Syst. **94**, 88–104 (2016)
15. El Dabbagh, H., Fakhr, W.: Multiple classification algorithms for the BCI P300 speller diagram using ensemble of SVMs. In: 2011 IEEE GCC Conference and Exhibition (GCC), pp. 393–396 (2011)
16. De Bock, K.W., Van den Poel, D.: An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction. Expert Syst. Appl. **38**(10), 12293–12301 (2011)
17. Li, Y., et al.: Joint spectral-spatial hyperspectral image classification based on hierarchical subspace switch ensemble learning algorithm. Appl. Intell. **48**(11), 4128–4148 (2018)

fridaya@uef.fi

A Parallel Distributed Galois Lattice Approach for Data Mining Based on a CORBA Infrastructure

Abdelfettah Idri[1(✉)] and Azedine Boulmakoul[2]

[1] National School of Business and Management, Casablanca, Morocco
abdelfattah.id@gmail.com
[2] LIM Lab. IOS, Computer Sciences Department,
Faculty of Sciences and Technology Mohammedia, Mohammedia, Morocco
azedine.boulmakoul@yahoo.fr

Abstract. Galois lattices are tightly connected to Formal Concept Analysis as they generate a concept hierarchy that helps structuring and clustering closed frequent item sets. When used in Data mining, such a structure can help dealing with large data sets which are very common in such a context. Constructing Galois Lattices is one of the most complex problems in FCA. Our work focuses on the generation process of the Galois Lattice which is based on a parallel distributed approach relying on a CORBA infrastructure. This contribution is part of our global framework aimed at spatial association rules discovery for geomarketing purposes destined to a telecom operator. We will adopt a generation method of association rules that exploits the same concept lattice formerly used to process the closed frequent item sets and which is inspired from the minimal generator concept. The architecture is Manager/Agent based where the agent may encapsulate different concept generation methods given that these ones respect the same concept interfaces.

Keywords: Galois lattice · Association rules · CORBA infrastructure

1 Introduction

Business intelligence and data mining are often related, especially when we are dealing with huge datasets. Nowadays, several contexts necessitate advanced methods and techniques to solve problems in the spatial data mining as the developing of this later is growing so fast [1, 2]. In case of geographic and spatial databases, human perception of decision making may be subject of losing insights. This work is intended to overcome such constraints and reduce the computational complexity by presenting a parallel distributed approach to build up the concept Lattice and by consequent, to extract the spatial knowledge relying on the spatial association rules.

This paper is organized as follows. This first section introduces briefly our geomarketing framework (analyzer) as well as the basics of the formal concept analysis needed for this context. Section 2 presents the sequential algorithm. The parallel distributed architecture and algorithm for constructing the Galois Lattice are proposed in Sect. 3. In Sect. 4, we present some experimental results and examples to illustrate our

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 99–117, 2019.
https://doi.org/10.1007/978-3-030-22868-2_8

fridaya@uef.fi

approach. The association rules generation process is shown in Sect. 5. And finally, this paper ends with conclusions in Sect. 6.

1.1 Global Architecture of the Geomarketing Framework

The architecture of our geomarketing analyzer is component based and integrates the necessary functionalities to support spatial analysis [3–5]. Figure 1 presents a logical view of the system.

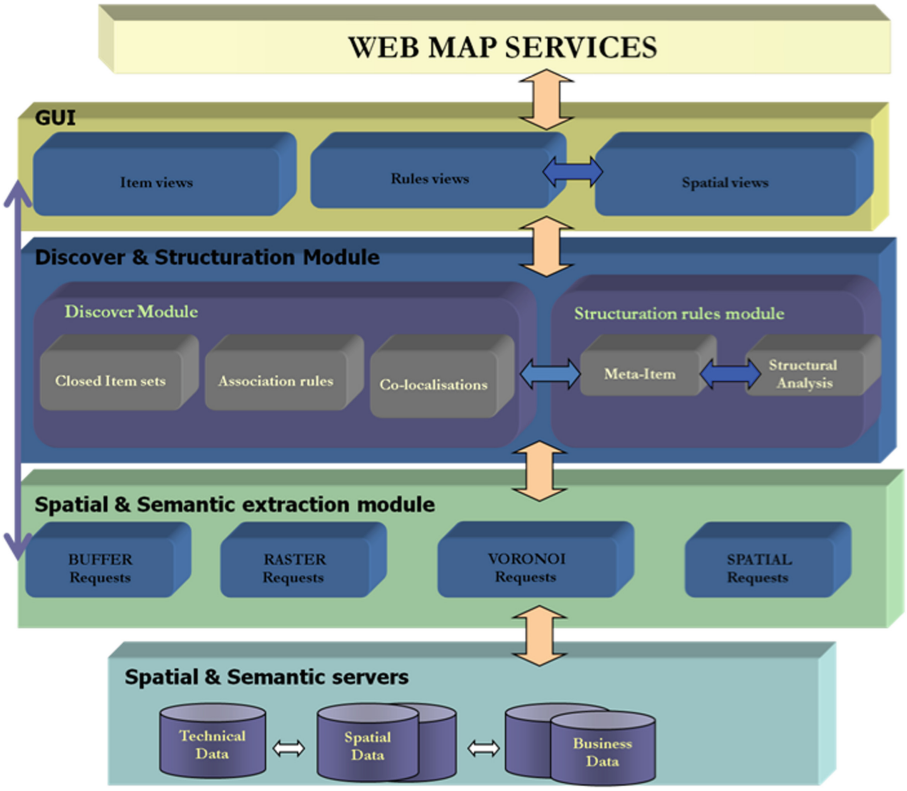


Fig. 1. Global architecture of the geomarketing framework

The adopted architecture is composed of five main layers: spatial and semantic servers, spatial and semantic extraction module, discover and structuration module, graphical user interface and the web mapping services. In this paper, our attention will be given to the discovery module. The detailed description of the other components is out of this scope and will be handled in future work. Rather, a general overview of the fundamental modules will follow.

fridaya@uef.fi

Spatial and Semantic Extraction Module

Requests and Spatial Analysis

Spatial requests include criteria related to the position of objects and spatial relations between them. These relations concern either the same (intra-theme) or different (inter-theme) theme layers. In general, most spatial analysis functions generates new shape files or entity classes (e.g. spatial joins, distance point/node).

Spatial Transaction Extraction

This is an important consuming step in the extraction and the visualization process of spatial knowledge. We propose a model to present spatial transactions such that knowledge discovery methods like association rules can be applied. From within a GIS system, the user can configure his context for spatial extractions (set of layers involved in the knowledge discovery process) and express the appropriate spatial relations by means of a structuring neighboring element. Once a reference layer is defined, we can define then the neighboring objects belonging to this layer by considering the attribute of the given object and the other related objects emerging from the chosen structuring element (buffer, raster or voronoi polygons). In this case, the resulting spatial transactions can be mined by an association rules generator. A description of the different spatial requests follows.

Buffer Request

This request type is aimed at delimiting an area in space around a given service where each of its points should share a certain characteristic we wish to investigate. The area is specified by its radius (see Fig. 2).

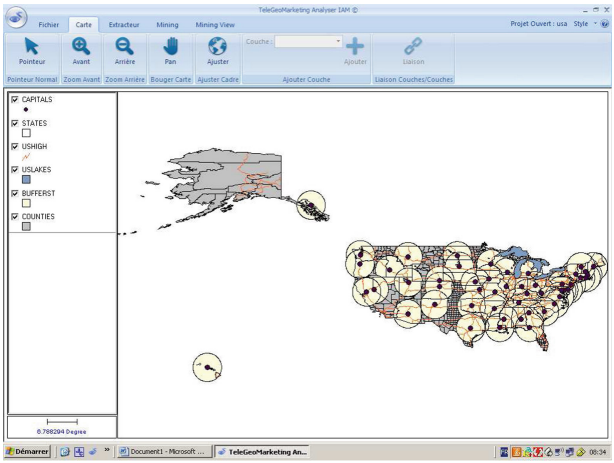


Fig. 2. Buffer extraction

Raster Request

It consists of slicing the space in cells with uniform or variable size (see Fig. 3).

fridaya@uef.fi

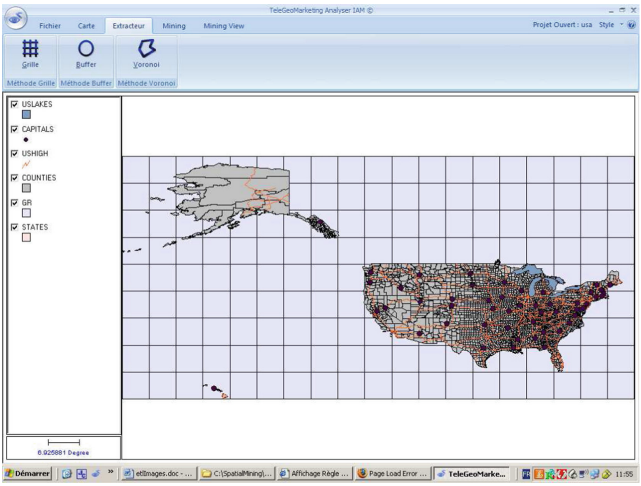


Fig. 3. Raster extraction

Voronoi Request

In mathematics, a Voronoi diagram is a special kind of decomposition of a space determined by distances to a discrete set of points. A Voronoi request allows us to select different theme layers in the space by using only one point in this space.

We recall that a Voronoi diagram is a partitioning into zones defined by considering the distance to specific points called generators identified beforehand so that each point of a given zone based on its generator P is closer to P than any other region's generator (see Fig. 4).

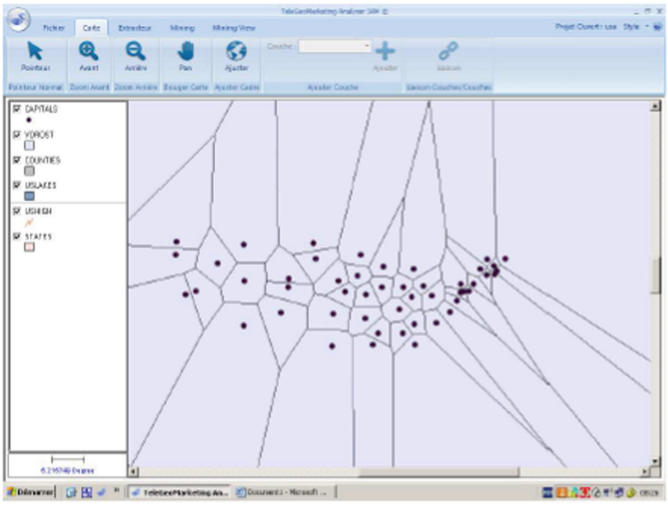


Fig. 4. Voronoi extraction

fridaya@uef.fi

Discover Module

This module is mainly based on the Lattice parallel/distributed approach that will be detailed in this paper. The main concepts of knowledge discovery from within a transactional spatial data base will be explored in detail [6]. The input for our parallel algorithm is composed of the spatial transactions generated by the spatial extraction module and the resulting spatial association rules are passed to the visualization module to be explored.

GUI Module

The graphical user interface allows the user to select both spatial layers and semantic data involved in spatial mining. Besides, the GUI is interfaced with the web through Web map services. From the varieties of interfaces provided by the GUI, we choose the following.

Association Rules Visualization

In the following, an example is given of some spatial association rules generated by the geomarketing framework. Figure 5 exposes the graphical user interface and the spatial association rules representation.

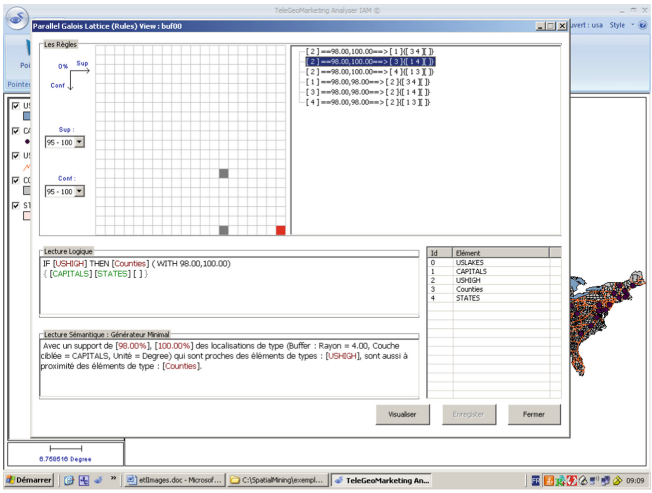


Fig. 5. Spatial association rules visualization

Corresponding Rules

IF [USHIGH] THEN [CAPITALS] (WITH 98.00,100.00) {[Counties] [STATES] [] }
IF [USHIGH] THEN [Counties] (WITH 98.00,100.00) {[CAPITALS] [STATES] [] }
IF [USHIGH] THEN [STATES] (WITH 98.00,100.00) {[CAPITALS] [Counties] [] }
IF [CAPITALS] THEN [USHIGH] (WITH 98.00,98.00) {[Counties] [STATES] [] }
IF [Counties] THEN [USHIGH] (WITH 98.00,98.00) {[CAPITALS] [STATES] [] }
IF [STATES] THEN [USHIGH] (WITH 98.00,98.00) {[CAPITALS] [Counties] [] }

fridaya@uef.fi

The above example represents the association rules generated by the discovery module for a spatial association rule.

Semantic description: with a support of [98.00%], [100.00%] of localizations of type buffer within the range of [USHIGH] elements, are also in the range of [Counties] elements.

Logic description: IF [USHIGH] THEN [Counties] WITH ([98.00%], [100.00%]) {[CAPITALS] [STATES] }{}

Hereafter, the metadata of the analysis method is given (see Table 1 and Fig. 6).

Table 1. Metadata of the used analysis method.

Analysis method details	
Type	Buffer
Radius	4.00
Targeted layer	CAPITALS
Unit	Degree
Processed layers	USLAKES CAPITALS USHIGH Counties STATES
Algorithm details	
Type	Parallel distributed Galois Lattice (association rules generation)
Support	0.00
Confidence	0.00

1.2 Motivation of the Adopted Research Method

In the first place, our interest is the generation of the Closed Frequent Item Sets (CFIS) as this is the most consuming step in terms of CPU processing time. On the other side, our proposed framework is intended to explore these closed frequent item sets repeatedly, every time the final user needs to change his support and confidence configuration to investigate individually some association rules subsets he is interested in. From a functional point of view, the suitable method should generate and hold permanently these items until the user exploration session is closed. If we want additionally to target a larger search space and consequently take in account the totality of these CFIS, then the required method may not directly embed the selective search criteria (support and confidence) to avoid excluding some CFIS from the ending result prematurely and shouldn't rely on a volatile technique in order to guarantee the persistency of these CFIS for future reuse (association rules).

Methods like Apriori [6] based on the item set candidate, Close [11] that adopts the Frequent Pattern Tree (FP-Tree), CHARM [17] and DiffCHARM [18] that uses CHARM-extend and CHARM-Property don't completely meet the requirements stated above. Moreover, the generated result doesn't cover the extension sets (transactions) and isn't hosted in a consistent abstract data type like the Galois Lattice. All these

fridaya@uef.fi

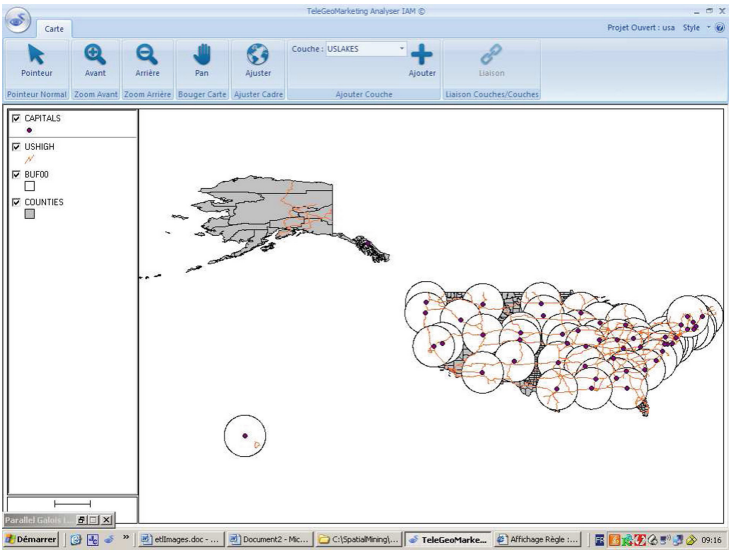


Fig. 6. Corresponding map

methods need to compute the support of CFIS which is often costly. Although CHARM-L [19, 20] generates a Galois Lattice, this later doesn't take in account the transactions as it is required for our purposes.

The aims of this work are:

- The generation of the full set of concepts respecting the constraints support and confidence.
- A scalable solution which can dynamically adapt itself to the generation process needs.
- Possibility of reusing existing methods respecting the context requirements.

Therefore, our approach focuses on a parallel distributed algorithm to allow the scalability of the proposed solution and is based on a Galois Lattice as a fundamental component to ensure the exhaustive generation of concepts (both intentions and extensions). According to this approach, all the CFIS are captured within the Galois Lattice including their corresponding extension (transaction set). The same resulting Galois Lattice will be reused to generate the association rules based on the minimal generator principal as we will see further.

It is obvious that a method that satisfies all the research needs is extremely difficult to find and our approach doesn't make an exception: once the Galois Lattice generated, it must fit in memory to achieve the pretended performance. As its size may be huge, this constraint can impact the algorithm performance. In order to improve the generation process when this situation occurs, a particular memory management strategy has to be applied: persistency management, distributed memory management or other methods. In future work, we will present a specific distributed management approach as well as a similar solution based on the Map/Reduce paradigm.

fridaya@uef.fi

1.3 FCA Terminology and Background

The construction process of Galois Lattice has received a great deal of interest in both Formal Concept Analysis [7–9] and Data Mining areas [10, 11]. The scientific results of these fields are applied in several businesses related to information industry and knowledge discovery. The main reason is that they deal with very large datasets and there is an imminent need of converting such datasets into useful knowledge. In such context, data mining has become an efficient tool for extracting knowledge from these datasets to support decision making. FCA represents in fact a framework for data mining as these two fields are strongly related: there is a one-to-one relation between the intention of a concept and the closed item set [10]. Seen that the mining of large datasets still needs efficient algorithms, our interest focuses on Galois Lattice which allows the generation of closed frequent item sets and consequently association rules. In this paper we propose a parallel distributed Galois Lattice algorithm based on a CORBA infrastructure for generating closed frequent item sets and association rules. We worked on other alternatives like the Map Reduce paradigm, but this will be exposed in future contributions. Our approach is process oriented and ignores the data clustering, partitioning and slicing aspects. This contribution targets the reduction of the computational complexity. Other parallel or distributed methods are more related to data sharing, partitioning or slicing as in [12] which is based on a divide and conquer approach or in [13] which is algorithm specific (Ganter) and exploits some specific partitions.

It is important to consider the relationship between Galois Lattice and data prospection. In fact, there is a certain bijection mapping the Galois Lattice to closed frequent item sets since the intent of a concept coincides with the closed item set notion [14]. We will present hereafter, the basic definitions and concepts on which relies our algorithms [8, 15].

Basic Definitions

Definition 1 Context: In FCA, we call a context a triple (O, M, I) where $O = \{g_1, g_2, \ldots, g_n\}$ is a set of n elements, called objects; $M = \{1, 2, \ldots, m\}$ is a set of m elements called attributes and $I \subseteq O \times M$ is a binary relation between the objects and the attributes.

A context is often represented by a table with objects in colons and attributes in rows as shown in Table 2. We call an object set a set $X \subseteq O$. A set $J \subseteq M$ is called an attribute set. Following the convention, an object set $\{b, d, e\}$ is written as bde and an attribute set $\{3, 4, 6\}$ as 346.

Definition 2 Adjacency List: The set of the common objects of an element $i \in M$ is defined by $nbr(i) = \{g \in O : (g, i) \in I\}$ and is called the adjacency list of i. Similarly, $nbr(g) = \{i \in M : (g, i) \in I\}$ denotes the set of the common attributes of the element $g \in O$ and is called the adjacency list of g. In the table above, we can read $nbr(a) = \{1, 6\}$ and $nbr(1) = \{a, b, c\}$.

fridaya@uef.fi

Table 2. A context example (O, M, I) with $O = \{a, b, c, d, e\}$ and $M = \{1, 2, 3, 4, 5, 6, 7\}$. The table represents the binary relation.

	1	2	3	4	5	6	7
a	x					x	
b	x		x	x	x	x	
c	x			x		x	
d		x	x		x		
e		x					x

Definition 3 attr and obj Functions: The function $attr : 2^O \rightarrow 2^M$ maps a given object set to its common attributes, that's $attr(X) = \bigcap_{g \in X} nbr(g)$ with $X \subseteq O$. The function $obj : 2^M \rightarrow 2^O$ maps a given attribute set to its common objects: $obj(J) = \bigcap_{j \in J} nbr(j)$ with $J \subseteq M$.

Definition 4 Set Closure: An object set $X \subseteq O$ is closed if $X = obj(attr(X))$. An attribute set $J \subseteq M$ is closed if $J = attr(obj(J))$. In Table 2, the set $X = abc$ is closed because $obj(attr(X)) = abc$; $attr(abc) = 16$ and $obj(16) = abc$.

Definition 5 Concepts: A concept is a couple $C = (X, J)$ with $X \subseteq O$ and $J \subseteq M$ where the following statements are verified: $X = obj(J)$ and $J = attr(X)$.

The set X is called the extent of C and is denoted by $X = ext(C)$. The set J is called the intent of C and is denoted by $J = \text{int}(C)$. By definition, both X and J are closed. The set of all concepts related to the context (O, M, I) is denoted by $B(O, M, I)$ or B. The relation defined on B by:

$$(A_1, B_1) \prec (A_2, B_2) \Leftrightarrow A_1 \subseteq A_2 (B_2 \subseteq B_1)$$

where (A_1, B_1) and (A_2, B_2) are concepts belonging to B, is a partial order relation on B.

Definition 6 Galois Lattice: $L = < B, \prec >$ is a complete Galois Lattice.

2 Sequential Algorithm

The Galois Lattice is built up by generating concepts by means of concept successor's identification. The main idea in this process is to start with the parent concept $(O, attr(O))$ and process its children recursively using a Breath First Algorithm (BFS) or a Depth First Algorithm (DFA).

The generation process of the Galois Lattice includes in general three main steps:

- **Data preparation**: In this step, the context is constructed and loaded into memory depending on the chosen model (SLF, transactional, …)
- **Galois Lattice generation**: Using a global and a local Trie (lexicographic classification) organized as follows:
  - **Generation** of the children candidates of a given concept using a local Trie.

fridaya@uef.fi

- **Closure** test of a concept candidate. This test allows us to identify the final children concept list based on the previous list.
- **Existence** test based on the global Trie codified using either the concept intent or the concept extent. This test allows us to avoid processing a concept several times.

– **Galois Lattice visualization:** Provides a graphical visualization of the Lattice and consequently the individual exploration of each Lattice node representing a concept.

These steps are shown in Fig. 7.

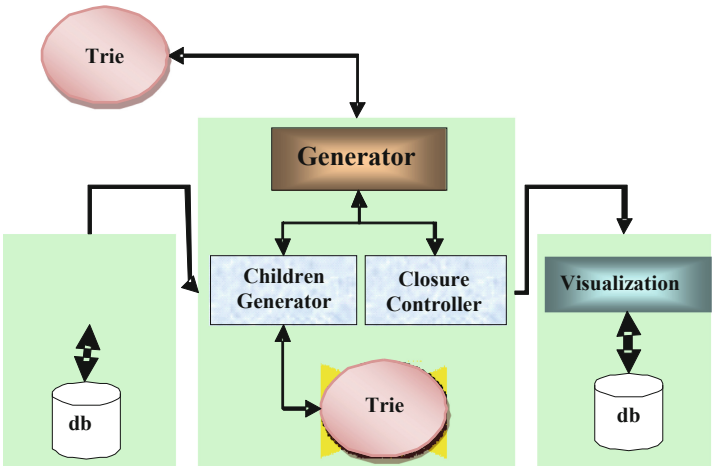


Fig. 7. Architecture of the sequential algorithm model

The corresponding algorithm is given below (Fig. 8) as a model of the sequential approach [15]. It is clear from the closure test in line **7** performed on the children (concept candidates) computed before in line **5** that a concept successor is obtained if this test is verified. To avoid processing a concept several times in order not to impact the algorithm performance, it is necessary to execute the existence test as stated in line **8**.

3 Parallel Distributed Algorithm

It is very common in data mining that the building processes of concepts especially those based on Galois Lattice approaches, often result in huge number of concepts compared to the size of the original database (exponential complexity). Thus, this task may become very hard to process regarding the time and space complexities. Therefore, this process still needs more efficient algorithms and techniques.

Our approach is intended to improve the performance of the sequential algorithm of building the Concept Lattice from a computation perspective (time and space) by focusing on the distribution and parallelism aspects of the mentioned process.

fridaya@uef.fi

```
Sequential Algorithm
1: Compute the top concept C = (O, attr(O));
2: Initialize a queue Q = {C};
3: Compute Child(C);
4: while Q is not empty do
5:   C = dequeue(Q);
     Let X = int(C) and suppose AttrChild(X) =< S1, S2, . . . , Sk >;
6:   for i = 1 to k do
7:      if XSi is closed then
          Denote the concept (obj(XSi),XSi) by K;
8:      if K does not exist then
9:          Compute Child(K);
10:         Enqueue K to Q;
11:       end if
12:     Identify K as a successor of C;
13:    end if
14:  end for
15: end while
```

Fig. 8. A model of the sequential algorithm

3.1 Architecture

The steps described below have leaded us to the conception of the architecture described hereafter.

The first phase was dedicated to identify the independent actions of the algorithm that can contribute to the optimization of the execution time and space. In the second phase we should state whether these actions are well suited for distribution without outperforming the communication between the resulting components. Finally, we have had to analyze the implementation issues of the architecture. The investigation of some existing algorithms like [7, 8] and [15] inspired us to converge to the following actions:

- Generation of the concept children
- Closure test of a set (intent or extent)
- Existence test of a concept

We opted for a **Manager/Agent** model as this later should guarantee the scalability and the distribution of the services on an intuitive way.

The children generation of a concept is a hard task which involves a special algorithm and a local tree (Trie). This task can be delegated to the agents as it may be executed separately. Multiplexing the number of agents results straightforward in runtime reduction.

The same way, the closure test can be easily delegated to the agents.

On the other hand, the existence check of a given concept is achieved by means of a global codification tree (Trie) that conserves the semantic of the concept content. That's why the key is built up from the intent elements or the extent as well since the closure is

fridaya@uef.fi

verified. This task can't be a candidate for distribution because the global tree has to remain shared by all Agents as the generated concepts by each agent are progressively inserted into this tree so that the information is centralized and can be globally accessed for existence control. Thus, this task belongs to the Manager.

The manager communicates with the agents by means of a dispatcher that distributes the task over the agents and a collector that gathers the results sent by these agents. The communication between all these components is achieved using CORBA bus. CORBA allows us to hide the complexity of the data structure used in the algorithm and simultaneously provides some kind of high level programming mechanisms like managing distributed events, supporting asynchronous communication and adopting an Object Oriented programming. Besides, if the right programming language is chosen, then we can gain a lot in terms of performance.

The services provided by the Manager and the Agents are listed below:

Manager:

- Global Tree management (concept insertion, existence check of a concept)
- Communication management (distribution, collection, synchronisation).

Agent:

- Generation of concept children
- Closure test of the concept intent or extent of a concept.

The resulting system architecture is presented in Fig. 9.

Within the CORBA bus, Asynchronous Method Invocation Routers are used to manage the communication between the Manager and the Agents.

3.2 Parallel Distributed Algorithm for Building the Galois Lattice (Fig. 10)

The construction process of the Galois Lattice based on our approach is performed in two main steps assigned to both manager and agents.

First Step:

The agent takes in account the generation of the children related to a given concept as first task. Next, it applies the closure control to each child belonging to the resulting list of the first step so that only the processed concepts having passed the test are sent then to the manager.

Second Step:

The manager sends available concepts from the agenda to the agents selected by the dispatcher in order to be processed. Once the collector receives the results from the agents as a concept list representing the children of the treated concepts, the manager starts then to update the Trie (tree) with the new concepts and establishes the necessary connections between the concept parents and the concept children. This process is repeated until each concept in the agenda is processed.

fridaya@uef.fi

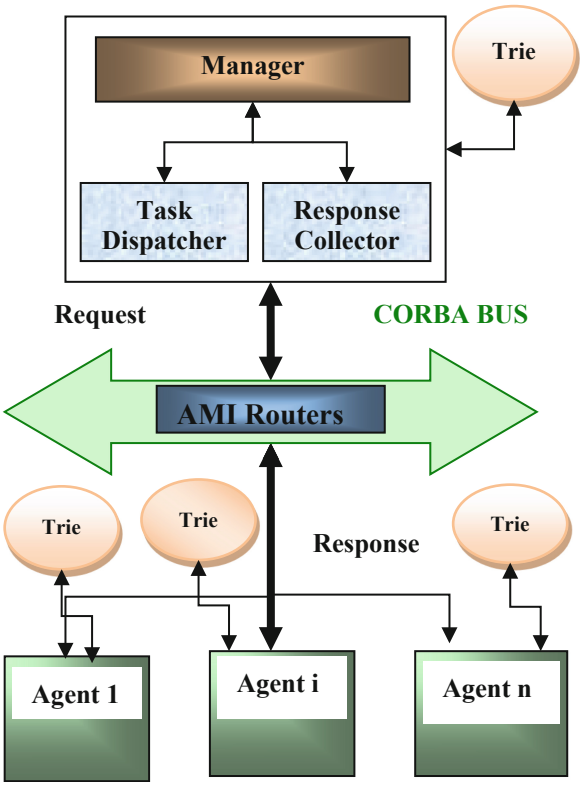


Fig. 9. The distributed architecture

3.3 Design Aspects

In this section, we will present some design aspects to help describing formally our solution.

The main task of building the lattice structure consists of spawning each concept generated by the agents starting with the root concept. As stated above, we adopted a parallel distributed approach to improve the performance of the sequential algorithm, but we need first to verify its consistency.

Distribution:

Generating the children of a given concept is an independent task as it is based on the initial context (or a part of it) and a local tree.

Parallelism:

The order of processing concepts is not relevant for the outcome because of the independency of the children generation. A global tree is used in this case to guarantee the uniqueness of concepts and to maintain the parent-children relation the lattice needs to be built.

fridaya@uef.fi

```
Algorithme parallèle Manager
1: Initialise concept queue Q= {C};  C = (O, attr(O))
2: Initialise Context, Agenda, Trie
5: Initialise and Synchronise Agents
6: while Q is not empty or Agenda is not empty
                or waiting for reponses do
7:    while Q is not empty do
8:      C = dequeue(Q)
9:      Enqueue C to Agenda
10:   end while
11:   for each available Agenti do
12:     if Agenda is not empty then
13:    C = dequeue(Agenda)
14:    Send_request(generateChildren(C)) to Agenti
15:    Mark Agenti as busy
16:     end if
17:   end for
18:   for each buzy agenti do
19:     if response of agenti is available then
20:       get response
21:         for each child in response do
22:              if child not exists in Trie then
23:                  insert child into Trie
24:              Identify child as successor of C
25:              Enqueue child to Q
26:              else
27:              Identify child as successor of C
28:             end if
29:           end for
30:       mark agenti as free
31:     end if
32:  end for
33: end while
```

Fig. 10. The parallel distributed algorithm

Trie Abstract Data Type:

As this data structure is mainly used in our design within both the agent and the manager components, we describe it here because of its importance and for a better insight.

For the tree (Trie), a lexicographic codification is adopted to store both the concepts and the candidate concepts. We need after all to manipulate sets as faster as possible. Such operations include inserting a set in the tree and checking its existence in the tree based only on its elements (semantic key). The key identifying a concept within the tree

fridaya@uef.fi

is then composed from his intent or extent elements seen that these ones are unique for a given concept. An example of the Trie is given below. In this case, two concepts are inserted in the tree: **(abc, 16)** and **(de, 2)**. The extents are given as sets of alphabetical letters for explanation purposes. In our implementation both intent and extent are integers. This tree is codified using an extent key. A precondition for using this tree is that the set elements have to be sorted to avoid duplicated entrees in the tree before inserting them. So every set containing element 'a' uses the node (a) in the tree as its first key element. See the set 'ad' after its insertion which is in fact not a concept as it is not tagged (yellow color) and it doesn't refer any intent. The main benefit of this data structure is the time complexity which is reduced to $O(n)$ where n is the set cardinality in both find and insert actions (Fig. 11).

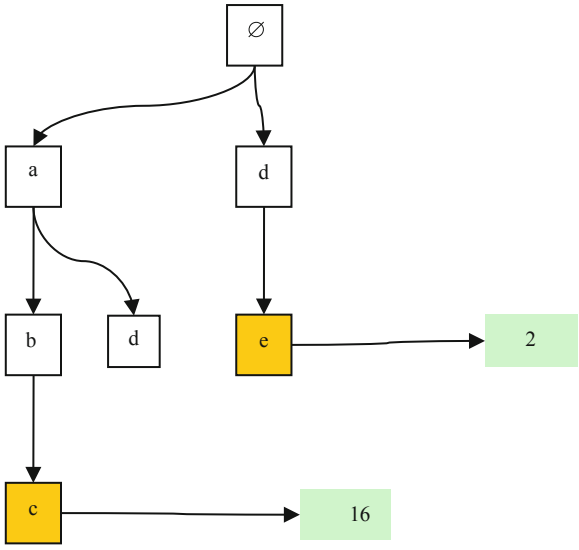


Fig. 11. Trie after inserting the concepts (abc, 16) and (de, 2) from Table 2.

4 Experimental Results of the Galois Lattice Generation

In this paragraph, we will expose some of our results obtained from the experimentations performed. First, an example is given to illustrate the data formats adopted and the visualization tool used at this stage, because when this module is integrated within the global architecture described in the first paragraph, all the visual interfaces are delegated to geomarketing GUI.

Example:

For the visualization, we adopted the Galicia tool [16] for our experiments. As an interface format we have chosen the GSH-XML format. The user can either directly launch the algorithm from the Galicia framework or he can launch it separately outside Galicia and use the generated results in the GSH-XML format in Galicia to visualize it. See the following example to illustrate our architecture. We present in Fig. 12, the

fridaya@uef.fi

supported formats: SLF and transactional. The graphical result in Galicia of the transactional example (Fig. 12b) is shown in (Fig. 12c).

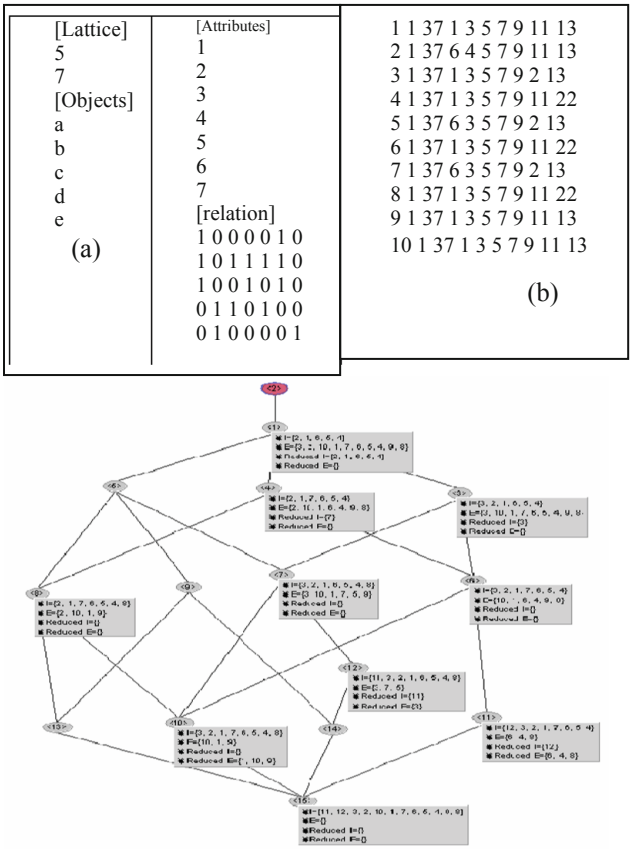


Fig. 12. Example: (a) SLF; (b) Transactional; (c) Galicia visualization

5 Results

In this experiment, we used only one physical computer and by consequent, the agents are virtual components running on the same machine.

- Dataset: "mushroom" file (8124 objects and 119 attributes)
- Number of agents: 3 agents on the same computer
- Number of logic router: 1 AMI router

We ran experiments with the mushroom dataset. The results are given below. The whole file fit easily in memory that's because of the optimised implementation based on bitsets (will be explained in detail in future papers). All the components were running

fridaya@uef.fi

on the same physical computer. This means that these results can be clearly improved by running each agent separately on a physical machine. On the other side, we can increase the performance of the architecture by multiplexing the number of agents in such a way to not outperform the communication between the manager and these agents even if the size of data transported is minimized. Figure 13 shows the execution time in single and distributed mode for the mushroom dataset.

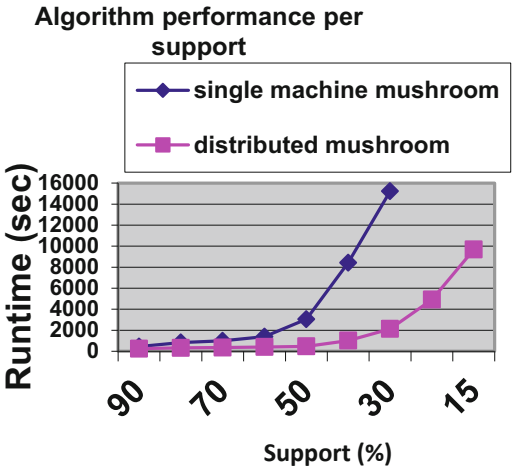


Fig. 13. Algorithm performance in single and distributed mode.

6 Association Rules

Our algorithm generates all frequent item sets since it generates the Galois Lattice (set of concepts). The support is nothing else than the extent cardinality. The generation of the closed frequent item sets becomes evident. But association rules remain the most important knowledge in the database to explore. To achieve this goal we should exploit again the same generated Galois Lattice as mentioned above.

Analysis was done to extract associations rules using a Galois Lattice and one result was the Mirage framework [17]. We took this framework as model and we have implemented it using our logic as this framework is based on Galois Lattice. Figure 14 illustrates the architecture adopted for generating the association rules.

The most time and space consuming task is in fact the Galois Lattice generation process. Our approach is intended to improve its performance based on a parallel/distributed concept. We reused the same generated Lattice to generate the association rules. The key element in this action is the minimal generator of a closed item set which is one of its subsets except that it may not be included in neither of all its direct children regarding the Galois Lattice that holds all the parent/child connections. Consequently, the generation of exact and inexact rules is based on these minimal generators.

fridaya@uef.fi

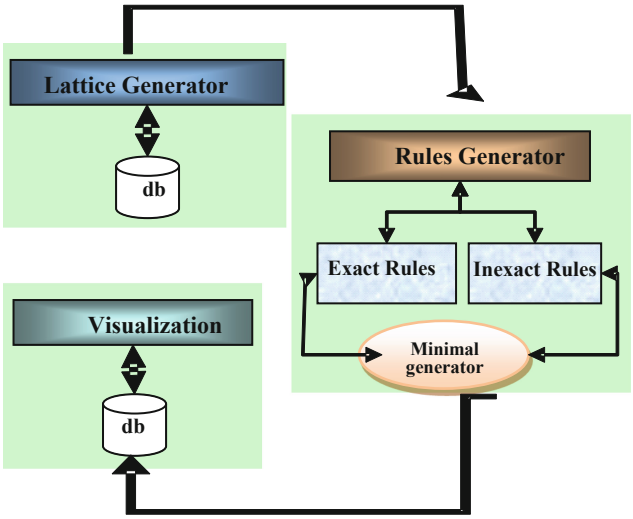


Fig. 14. Association rules generation

7 Conclusion

We presented in this paper a brief overview of our geomarketing framework dedicated for the extraction and the visualization of spatial association rules. Our focus lay on the building process of the Galois Lattice as this later holds all the concepts generated so far and by consequent all the closed frequent item sets. Seen that the spatial association rules are the precious knowledge to discover at the end, we based their extraction on the same generated Galois Lattice.

The distribution of the Galois Lattice construction algorithm allowed us to efficiently handle each phase of this generation process by dissociating its main tasks: children generation (Agents) and the management of the global concepts tree (Manager). The former is time consuming and the second uses a huge amount of memory. We could by means of this architecture test and optimise each of these processes separately. The results are much promising regarding those of the sequential algorithm. On the other hand, the parallelism helped us to improve the performance of the algorithm and give it a scalability dimension by multiplexing the agents whenever it is needed, overcoming this way the natural limits of the sequential algorithm.

References

1. Shekhar, S., Chawla S.: Introduction to Spatial Data Mining, in Spatial Databases: A Tour. Prentice Hall, Englewood Cliffs. ISBN 013-017480-7 (2003)
2. Malerba, D.: A relational perspective on spatial data mining. IJDMMM **1**(1), 103–118 (2008)

fridaya@uef.fi

3. Idri, A., Boulmakoul, A.: Une approche parallèle distribuée pour la génération des motifs fermés frequents basée sur une infrastructure corba. In: Boulmakoul, A. et al. (eds.) Les systèmes décisionnels: applications et perspectives, pp. 197–210. ISBN 978-9981-1-3000-1, ASD, 10–11 octobre 2008
4. Boulmakoul, A., Idri, A.: Une structure logicielle distribuée pour la découverte des règles d'association spatiales. In: Workshop on Decision Systems, vol. 1. ASD (2009)
5. Boulmakoul, A., Idri, A., Marghoubi, R.: Closed frequent item sets mining and structuring association rules based on Q-analysis. In: The 7th IEEE International Symposium on Signal Processing and Information Technology, pp. 519–524, Cairo, Egypt, 15–18 December 2007. ISBN: 978-1-4244-1834-3. https://doi.org/10.1109/isspit.2007.4458017
6. Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. In: Proceedings of ACM-SIGMOD International Conference Management of Data, pp. 207–216 (1993)
7. Ganter, B., Wille, R.: Mathematical Foundations: Formal Concept Analysis. Springer, Berlin (1999)
8. Bordat, J.P.: Calcul pratique du treillis de Galois d'une correspondance. Math. Sci. Hum. **96**, 31–47 (1986)
9. Chein, M.: Algorithme de recherche de sous-matrice première d'une matrice. Bull. Math. R. S. Roumanie **13**, 21–25 (1969)
10. Zaki, M.J., Ogihara, M.: Theoretical foundations of association rules. In: Proceedings of the 3rd ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, pp. 1–7 (1998)
11. Pasquier, N., Bastide, Y., Taouil, R., Lakhal, L.: Closed set based discovery of small covers for association rules. In: Actes des 15èmes journées Bases de Données Avancées (BDA'99), pp. 361–381 (1999)
12. Djoufak, J.F.K.,Valtchev, P., Djamegni, C.T.: A parallel algorithm for lattice construction. In: Ganter, B., Godin, R. (eds.) ICFCA 2005. LNCS (LNAI), vol. 3403, pp. 249–264. Springer, Heidelberg (2005)
13. Baklouti, F., Lévy, G.: Parallel algorithms for general Galois lattices building. In: Proceedings of the WDAS 2003
14. Zaki, M.J., Ogihara M.: Theoretical foundations of association rules. In: Proceedings of the 3rd ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, pp. 1–7 (1998)
15. Choi, V.: Faster Algorithms for Constructing a Concept (Galois) Lattice, Department of Computer Science. Virginia Tech, USA (2006)
16. Galicia Home page. http://www.iro.umontreal.ca/∼galicia/publication.html. Last accessed 12 Sep 2004
17. Zaki, M., Phoophakdee, B.: MIRAGE: a framework for mining, exploring and visualizing minimal association rules. Rensselaer Polytechnic Institute, New York (2003)
18. Zaki, M.J., Hsiao, C.J.: CHARM: an efficient algorithm for closed itemset mining. In: The 2nd SIAM International Conference on Data Mining (2002)
19. Zaki, M.J., Gouda, K.: Fast vertical mining using diffsets. In: 9th International Conference On Knowledge Discovery and Data Mining, KDD '03 (2003)
20. Zaki, M.J., Hsiao, C.J.: Efficient algorithms for mining closed itemsets and their lattice structure. IEEE Trans. Knowl. Data Eng. **17**(4), 462–478 (2005)

fridaya@uef.fi

A Community Discovery Method Based on Spectral Clustering and Its Possible Application in Audit

Hu Jibing[1(✉)], Ge Hongmei[1], and Yang Pengbo[2]

[1] Nanjing Audit University, Nanjing, China
gibenh@nau.edu.cn
[2] Xidian University, Xi'an, China

Abstract. The essence of human is social relation. In computer, relation can be represented by graph. It may be possible to divide a group of individuals into a few of communities according to the intensity of relation among individuals. The relation between individuals in the same community should be strong, while that between individuals in different communities should be weak. That the intensity of relation between two persons is high means that they trust each other; therefore, strong relation can be used to commit crime, for instance, corruption. Consequently, finding out strong relation can help auditor to reduce difficulty of assuring them of reliability of auditing. Seeing that those algorithms that involve iteration process have fatal defect, the authors of this paper introduce a method used to divide a group into a few of communities based on spectral clustering. This method has the advantage of high speed which means that it has favorable performance when it is used to cope with tremendous amount of data.

Keywords: Spectral clustering · Community discovery · Audit

1 Introduction

As government increases the intensity of anti-corruption, corrupt officials don't dare to graft blatantly any longer. They are certain to weaken their tracks in the process of corruption. A simple method of judging whether an official participate in corruption or not is checking his or her revenue and expenditure so as to confirm whether condition C is tenable, where condition C refers to that asset is equal to the result of subtracting overall expenditure form overall revenue (Similarly hereinafter). However, any official does not isolate from the society, which means that the official could take use of his or her relationship network to hide revenue items for successful corruption. Hence, the conclusion about whether an official participates in corruption or not drawn based on checking whether condition C is tenable with regard to the single official is no longer convincing. It is necessary to regard the whole consisting of the official and the individuals who have tight tie with the official as a community. As a result, audit raises the requirement of discovering communities existing in group. In other words, community discovery can be used to support audit.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 118–131, 2019.
https://doi.org/10.1007/978-3-030-22868-2_9

fridaya@uef.fi

In computer, a vertex can be used to represent a person in a group, and graph can be used to represent relation between two persons in the group. For example, that there is connection between two vertices means that there is relatively tight tie between the two persons who are represented by the two vertices. Community discovery is the process of dividing a group of individuals into a few of communities, each of which contains a few of persons. Community discovery should be implemented to ensure that connection in the same community is as tight as possible and that connection between different communities is as loose as possible. After a group of individuals is divided into a few of communities, every person belongs to a community and it is ensured that there is relatively tight tie between two persons in the same community. In auditing, it is necessary to check whether condition C is tenable or not with regard to the whole, which is a community that is a part of the result of community discovery method. If some person is doubted by auditing department, then the community containing this person should be regarded as a whole and it must be executed to check whether condition C is tenable or not with regard to the whole. If this is not the case, the economic exchanges between the doubted person and those who has tight tie with the doubted person will be ignored, which has negative effect on accuracy of audit.

From the perspective of social relation, community discovery can be seen as auxiliary means of auditing. The use of community discovery algorithms, especially those which have high efficiency, may deprive officials involved in corruption of place where they can hide. In view of harm done to economic society by corruption, we believe that it is pressing task to develop community discovery algorithm with high quality.

In statistics, there is a kind of methods called clustering. If individuals to be distributed into communities are regarded as data points, community discovery is quite similar to clustering. Regarding individuals to be distributed into communities as data points is a rational idea, so seeking for appropriate and efficient algorithms can support auditing.

2 Related Research

Community detection or community discovery is a hot research area in recent years. Luca Donetti et al. proposed an efficient and relatively fast algorithm for the detection of communities in complex networks. Their method exploits spectral properties of the graph Laplace matrix combined with hierarchical-clustering techniques, and introduces a procedure to maximize “modularity” of the output [1]. Newman proposed a new algorithm which can be used to discover community structure and give excellent result when tested on computer-generated and real-world networks as well as is much faster, typically thousands of time faster than those algorithms that are computationally demanding and therefore cannot be used to deal with large networks [2]. Aaron Clauset defines both a measure of local community structure and an algorithm that infers the hierarchy of communities that enclose a given vertex by exploring the graph one vertex at a time. This algorithm run in time $O(k^2 d)$ for general graph when d is the mean degree and k is the number of vertices to be explored. For graphs where exploring a new vertex is time consuming, the running time is linear, O(k) [3]. Capocci et al.

fridaya@uef.fi

developed an algorithm to detect community structure in complex networks. Their algorithm is based on spectral methods and takes into account weights and link orientation. Since the method detects efficiently clustered nodes in large networks even when these are not sharply partitioned, it turns to be especially suitable for the analysis of social and information networks [4]. Jordi Duch et al. proposed a novel method to find the community structure in complex networks based on an extremal optimization of the value of modularity [5]. Relatively particular details about this method will be revealed in Sect. 3. Mingming Chen et al. discussed the definition of modularity and reviewed the modularity maximization approaches which were used for community detection in the last decade, and discussed two opposite yet coexisting problems of modularity optimization, and overviewed several community quality metrics proposed to solve the resolution limit problem and discussed modularity density which simultaneously avoids the two problems of modularity, and introduced two novel fine-tuned community detection algorithms that iteratively attempt to improve the community quality measurements by splitting and merging the given network structure. They found that fine-tuned Q_{ds} is the most effective [6].

Yang Gui et al. proposed overlapping community detection on weighted networks (OCDW). In their method, edge weight is defined by combining topological structure and real information. Then vertex weight is induced by edge weight. To obtain cluster, OCDW selects seed nodes according to vertex weight. After finding a cluster, edges in this cluster reduce their weights to avoid being selected as a seed node with high probability [7]. Hu Xinzhuan et al. put forward a new signed network community detection algorithm CD-SNNSP (Community Detection in Signed Networks based on the Node Similarity Degree and Node Participation). Firstly, the concept of Node Influence and Node Clustering Coefficient for signed networks is proposed and then the initial node is selected based on them. Secondly, according to the formula of signed networks node similarity degree, the initial community is formed by the initial node and one of its neighbor nodes which is most similar to the initial node. Thirdly, to determine the order in which the neighbor nodes join the community by the node participation and whether the nodes are assigned to the community by the relative contribution increment [8]. Zhan Wenwei et al. proposed a hierarchical agglomerative community detection algorithm based on similarity modularity through introducing optimized similarity to improve the accuracy. It adopts NMI as the accuracy measurement so that it can overcome shortcoming with regard to accuracy brought by the fact that two nodes that have common neighbors and weak link information may not be merged, which is due to that traditional modularity is adopted for merger communities, meaning that only node link information is considered while neighbor nodes are ignored [9]. Du Hangyuan et al. proposed a new community detection algorithm for overlapping network after designing a centrality measurement model for network nodes. In their algorithm, the cohesion and separation of network nodes are defined at first. Depend on that, centrality measurement is calculated for each node to express its influence on network community structure. Then the nodes with tremendous centralities are selected as community centers. The overlapping features between communities are represented by memberships, and iterative calculation methods for the memberships of non-central nodes are put forward. After that, according to their memberships, all the nodes in the network can be allocated to their possible communities to accomplish the

fridaya@uef.fi

overlapping community detection [10]. Chen Jing et al. proposed the algorithm PMCD (Pairwise Merging Community detection) which resolve questions most algorithms may encounter, such as node error division, large number of sub-communities, and instability of community structure. In PMCD algorithm, network is divided into some small communities by edge weight which is consideration of similarity of community. Similarity between unassigned node and existing communities is computed, and then small communities are extended. The change of the module value is calculated on the basis of community power structure, to determine whether the sub-community for pairwise merge, until the formation of the final community [11]. Zheng Xiangping et al. proposed a community discovery algorithm based on location network by studying the characteristics of urban location network and its difference with traditional social network. The algorithm takes into account the proximity of location, the connection between the locations and the similarity of user's behavior. Firstly, the initial community is divided. Then, the extent of each site belonging to this community is iteratively calculated so as to find significant urban communities [12]. Aiming at the shortages of existing method for calculating the node similarity, Zhang Hu et al. proposed a novel method based on the multi-layer node similarity, which can not only calculate the similarity between nodes more efficiently, but also solve the problem of merging nodes when the node similarity is same. Furthermore, their constructed the community detection model based on the improved calculation method of the node similarity and the measure criteria of connection tightness between groups [13].

3 A Typical Method Used to Detect Communities and Its Defect

Many algorithms can perform detecting communities according to given data. Among them, there is an algorithm to be introduced, namely Extremal Optimization (EO), which adopts modularity as object function to guide the process of dividing large community into small communities. Like some other algorithms, EO is a typical algorithm that involves iteration process.

EO optimizes a global variable by improving extremal local variables. The global variant of this algorithm is modularity (Q), which is also adopted by many other algorithms.

In this algorithm, the contribution of an individual node i to Q of the whole network with a certain community structure is given by

$$q_i = k_{i,c} - k_i \frac{|E_c|}{2|E|}$$

Where $k_{i,c}$ is the number of edges that connect node i to the nodes in its own community c. Notice that $Q = \frac{1}{2|E|}\sum_i q_i$ and q_i can be normalized into the interval [−1,1] by dividing it by k_i

fridaya@uef.fi

$$\lambda_i = \frac{q_i}{k_i} = \frac{k_{i,c}}{k_i} - \frac{|E_c|}{2|E|}$$

Where λ_i, called fitness, is the relative contribution of node i to Q. Then, the fitness of each node is adopted as the local variable.

The detailed steps of this algorithm are omitted. The computational complexity of this algorithm is $O(|V|^2 \log^2 |V|)$, where |V| is the number of vertices in network. We have implemented this algorithm in recursive manner by using C language and then run it on given data. It can be seen that once |V| is larger than a certain number, computer system in which the algorithm is running will halt. It is fatal defect of all methods that involve iteration process. Since auditor must face a large amount of data, which means that |V| is usually quite large, auditor should give up those methods involving iteration process and alternative method must be exploited. Spectral clustering discussed in next chapter, which takes weight of connection into account, is such a kind of method.

4 Use Spectral Clustering to Divide Large Group into Communities

4.1 The Construction of Object Function

It is assumed that some individuals in a group are the objects doubted by auditing department. Compared with the number of all individuals in this group, the number of these individuals may be ignorable. According to above train of thought, it is necessary to regard a subgroup whose center is a certain person as a whole and then check issues such as revenue and expenditure regarding the whole, so as to confirm whether those who are doubted in the subgroup act illegally. The center of this subgroup is, of course, doubted. Since social relation is complicated, a group of individuals that does not include isolated points may contain numerous of persons. Our goal is finding all centers in this group by adopting highly efficient algorithm, leaving others that are not center distributed to different subgroups, each of which is characterized by some center. All the centers are those who are doubted by auditing department. In advance, auditing department is required to give the number of persons who are doubted in this group.

It is assumed that the number of persons in this group is equal to n. There may exist relation within each pair to different degree. Of course, it is allowed that there is no relation between two given persons. A matrix can be used to represent the degree to which there is relation within every couple consisting of two persons, both of which are in this group. The element of the matrix, for example, w_{ij} represents the degree to which there is relation between person i and person j. Here the matrix denoted as W, to which w_{ij} belongs, is called similarity matrix. Obviously, W is a symmetric matrix. If there is no relation between two persons, one of which is marked by p and the other of which is marked by q, then we have $w_{pq} = w_{qp} = 0$. If i and j have been given, then the bigger w_{ij} is, the higher the degree to which there is relation between the two persons, represented respectively by i and j, is. The higher the degree to which there is relation between two persons is, the more frequently they contact with each other. It is one way

fridaya@uef.fi

to assign a value to w_{ij} based on communication data which is mainly characterized by the frequency of communication. As to how a value is assigned to w_{ij} according to communication data, it is not discussed in this paper. In this paper, it is assumed that the value of every element of W is given in advanced. W is a matrix with n × n elements.

$$\mathrm{W} : (w_{ij})_{n\times n}$$

Our goal is to divide a group with n points into k subgroups, where k is a number whose value is also given in advance. The following criterion should be complied with in order to obtain favorable effect of division: The relation between individuals in different subgroups is as loose as possible, while the relation between individuals in the same subgroup is as tight as possible.

Laplace matrix is introduced before we understand how classification algorithm which meets the requirement of above division criterion is deduced. After similarity matrix, denoted as W, is given, a transformation is executed based on W, which results in degree matrix, denoted as D.

$$\mathrm{D} : (d_{ij})_{n\times n}$$

where

$$d_{ii} = \sum_{j=1}^{n} w_{ij}$$

$$d_{ij} = 0, \text{ if i} \neq \text{j}$$

After similarity matrix W and degree matrix D are constructed, Laplace matrix L is easily obtained since L = D − W. Specifically, L is unnormalized Laplace matrix, which has the following nature:

For every vector $\mathbf{f} \in R^n$, the equation $f^{'}Lf = \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2$ is always tenable for the following reason:

$$\begin{aligned} f^{'}Lf &= f^{'}Df - f^{'}Wf \\ &= \sum_{i=1}^{n} d_i f_i^2 - \sum_{i,j=1}^{n} f_i f_j w_{ij} \\ &= \frac{1}{2}\left(\sum_{i=1}^{n} d_i f_i^2 - 2\sum_{i,j=1}^{n} f_i f_j w_{ij} + \sum_{j=1}^{n} d_j f_j^2\right) \\ &= \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2 \end{aligned}$$

fridaya@uef.fi

Clusters are obtained by dividing the group according to the difference between levels of similarity among points, there are two aspects needed to be ensured, one is that the weight between edges in different subgroups is minimized, meaning that level of similarity between points in different subgroups should be fairly low, the other is that the weight between edges in the same subgroup is maximized, meaning that level of similarity between points in the same subgroup should be comparatively high.

After similarity matrix, which is denoted as W, has been given, we need to define a problem of minimizing cutting to divide the graph represented by W. The expression of object function is

$$cut(A_1, A_2, \ldots, A_k) := \frac{1}{2}\sum_{i=1}^{k} W(A_i, \bar{A_i})$$

Where $\bar{A_i}$ is the complementary set of A_i, or the union of all subsets of V except A_i. $W(A_i, \bar{A_i})$ is the sum of edges between A_i and the other subsets of V. We have

$$W(A_i, \bar{A_i}) = \sum_{m \in A_i, n \in \bar{A_i}} w_{m,n}$$

Where $w_{m,n}$ is an element of W.

Our goal of dividing the graph is ensuring that every two structures in any subgraph are similar to each other, where similarity means that the weight of edge is comparatively big on average and that there is connection between every two points in the subgraph as likely as possible, while the number of edges between points in different subgraphs is as small as possible, or the weight of edges between points in different subgraphs is fairly low. Then, our aim can be expressed as

$$\min \operatorname{cut}(A_1, A_2, \ldots, A_k)$$

But if you take a little bit of attention, you will find that this kind of minimizing cutting must definitely cause problem, as indicated in Fig. 1.

As shown in Fig. 1, if min cut $(A_1, A_2, \ldots, A_k)$ is used to cut graph, then cutting the collection of all sample points will result in many discrete graphs with many isolated points. Obviously, this kind of cutting can be finished in shortest period and meet the requirement of minimization operation to the greatest extent. But clearly, this result is not what we want. So other object functions emerge, one of which is Ratiocut. More specifically,

$$Ratiocut(A_1, A_2, \ldots, A_k) = \frac{1}{2}\sum_{i=1}^{k} \frac{W(A_i, \bar{A_i})}{|A_i|} = \sum_{i=1}^{k} \frac{cut(A_i, \bar{A_i})}{|A_i|}$$

where $|A_i|$ is the number of points in A_i.

fridaya@uef.fi

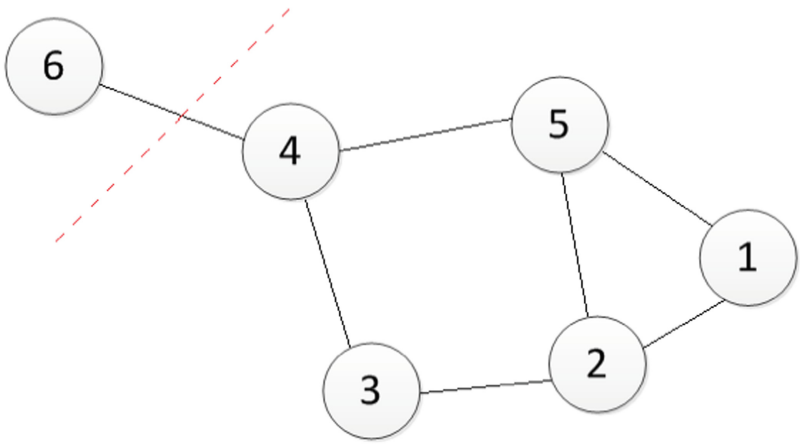


Fig. 1. An iffy graph cutting.

4.2 Unnormalized Spectral Clustering Algorithm

The deduction based on Ratiocut is unnormalized spectral clustering algorithm. Firstly, we check the situation where k is equal to 2 and object function is $\min_{A\subset V} Ratiocut(A,\bar{A})$.

At this time, a vector **f** is defined in this way: $\mathbf{f} = (f_1, f_2, \ldots, f_n)^T \in R^n$.

$$f_i = \begin{cases} \sqrt{|\bar{A}|/|A|}, v_i \in A \\ \sqrt{|A|/|\bar{A}|}, v_i \in \bar{A} \end{cases}$$

Now, the relation between Ratiocut and Laplace matrix is:

$$\begin{aligned} f' Lf &= \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2 \\ &= \frac{1}{2}\sum_{i\in A, j\in \bar{A}} w_{ij}\left(\sqrt{\frac{|\bar{A}|}{|A|}} + \sqrt{\frac{|A|}{|\bar{A}|}}\right)^2 + \frac{1}{2}\sum_{i\in \bar{A}, j\in A} w_{ij}\left(-\sqrt{\frac{|\bar{A}|}{|A|}} - \sqrt{\frac{|A|}{|\bar{A}|}}\right)^2 \\ &= \mathrm{cut}(A,\bar{A})\left(\frac{|\bar{A}|}{|A|} + \frac{|A|}{|\bar{A}|} + 2\right) \\ &= \mathrm{cut}(A,\bar{A})\left(\frac{|\bar{A}| + |A|}{|A|} + \frac{|A| + |\bar{A}|}{|\bar{A}|}\right) \\ &= |V| \times \mathrm{RatioCut}(A,\bar{A}) \end{aligned}$$

fridaya@uef.fi

At the same time, the constraint condition of **f** is:

$$\sum_{i=1}^{n} f_i = \sum_{i\in A}\sqrt{\frac{\left|\bar{A}\right|}{|A|}} - \sum_{i\in\bar{A}}\sqrt{\frac{|A|}{\left|\bar{A}\right|}} = |\mathrm{A}|\sqrt{\frac{\left|\bar{A}\right|}{|A|}} - \left|\bar{A}\right|\sqrt{\frac{|A|}{\left|\bar{A}\right|}} = 0$$

$$\|f\|^2 = \sum_{i=1}^{n} f_i^2 = |\mathrm{A}|\frac{\left|\bar{A}\right|}{|A|} + \left|\bar{A}\right|\frac{|A|}{\left|\bar{A}\right|} = \left|\bar{A}\right| + |\mathrm{A}| = \mathrm{n}$$

Currently, the problem is converted into

$$\min_{A\in V} f' Lf$$

$$\text{s.t. } \mathbf{f}\perp\mathbf{1}, \text{and} ||\mathbf{f}|| = \sqrt{n}$$

Since **f** is discrete and only two values can be assigned to every component of **f**, this division involves selections whose number is exponential function of n. To solve this trouble, we can extend the range of **f** from discrete interval to continuous one. In that way, this problem is actually a linear algebra problem at this time: to find out a specific vector **f**, which can minimize $f' Lf$ and is orthogonal to vector **1** as well as has modulus of $\sqrt{n}$.

According to Rayleigh-Ritz theorem, the solution of this objection function is second least eigenvector of matrix L.

When k is bigger than 2, object function is

$$\min \mathrm{Ratiocut}(A_1, A_2, \ldots, A_k)$$

A transformation is executed with regard to above minimization problem by introducing the indicating vector of $\{A_1, A_2, \ldots, A_k\}$, which is $h_j = \{h_1, h_2, \ldots, h_i, \ldots, h_k\}$ (j = 1,2,…,k), where i is the index of sample and j is the index of subset. $h_{j,i}$ is the indicator of sample i with regard to subset j. Specifically,

$$h_{j,i} = \begin{cases} \frac{1}{\sqrt{|A_j|}}, v_i \in A_j \\ 0, v_i \notin A_j \end{cases}$$

Common understanding about indicating vector is that every subset A_j corresponds to a indicating vector denoted as h_j, every h_j has n elements, each of which represents indicating result of one of the n sample points. If the sample, whose order number is i, is distributed into subset A_j, then the element in h_j whose order number is i is equal to $\frac{1}{\sqrt{|A_j|}}$, otherwise, 0.

fridaya@uef.fi

By further computation, we have

$$\begin{aligned} h_i^T L h_i &= h_i^T (D - W) h_i \\ &= h_i^T D h_i - h_i^T W h_i \\ &= \sum_{m=1} \sum_{n=1} h_{i,m} h_{i,n} D_{m,n} - \sum_{m=1} \sum_{n=1} h_{i,m} h_{i,n} w_{m,n} \\ &= \sum_{m=1} h_{i,m}^2 D_{m,m} - \sum_{m=1} \sum_{n=1} h_{i,m} h_{i,n} w_{m,n} \\ &= \frac{1}{2} (\sum_{m=1} h_{i,m}^2 D_{m,m} - 2 \sum_{m=1} \sum_{n=1} h_{i,m} h_{i,n} w_{m,n} + \sum_{n=1} h_{i,n}^2 D_{n,n}) \end{aligned}$$

Go a step further, we have

$$\begin{aligned} h_i^T L h_i &= \frac{1}{2} (\sum_{m=1} h_{i,m}^2 D_{m,m} - 2 \sum_{m=1} \sum_{n=1} h_{i,m} h_{i,n} w_{m,n} + \sum_{n=1} h_{i,n}^2 D_{n,n}) \\ &= \frac{1}{2} (\sum_{m=1} \sum_{n=1} h_{i,m}^2 w_{m,n} - 2 \sum_{m=1} \sum_{n=1} h_{i,m} h_{i,n} w_{m,n} + \sum_{n=1} \sum_{m=1} h_{i,n}^2 w_{n,m}) \\ &= \frac{1}{2} \sum_{m=1} \sum_{n=1} w_{n,m} (h_{i,m} - h_{i,n})^2 \end{aligned}$$

By making further efforts, which is substituting $h_{j,i}$ mentioned above for those in above equation, the following deductions are consequently obtained:

$$\begin{aligned} h_i^T L h_i &= \frac{1}{2} \sum_{m=1} \sum_{n=1} w_{n,m} (h_{i,m} - h_{i,n})^2 \\ &= \frac{1}{2} (\sum_{m \in A_i, n \in \bar{A}_i} w_{m,n} (\frac{1}{\sqrt{|A_i|}} - 0)^2 + \sum_{m \in \bar{A}_i, n \in A_i} w_{m,n} (0 - \frac{1}{\sqrt{|A_i|}})^2) \\ &= \frac{1}{2} (\in \sum_{m \in A_i, n \in \bar{A}_i} w_{m,n} \frac{1}{|A_i|} + \sum_{m \in \bar{A}_i, n \in A_i} w_{m,n} \frac{1}{|A_i|}) \\ &= \frac{1}{2} (\mathrm{cut}(A_i, \bar{A}_i) \frac{1}{|A_i|} + \mathrm{cut}(\bar{A}_i, A_i) \frac{1}{|A_i|}) \\ &= \frac{cut(A_i, \bar{A}_i)}{|A_i|} \\ &= \mathrm{Ratiocut}(A_i, \bar{A}_i) \end{aligned}$$

In order to take all indicating vectors into consideration, let **H** be equal to $\{h_1, h_2 \ldots, h_k\}$, where h_i is arrayed by column. According to the definition of h_i and the fact that h_i and h_j is orthogonal to each other when i is different from j (in other words, $h_i * h_j = 0$ if i $\neq$ j) as well as that $h_i * h_i = 1$, we have

fridaya@uef.fi

$$\begin{aligned}\text{Ratiocut}(A_1, A_2, \ldots, A_k) &= \sum_{i=1}^{k} h_i^T L h_i \\ &= \sum_{i=1}^{k} (H^T L H)_{ii} \\ &= \text{Tr}(H^T L H)\end{aligned}$$

where "Tr" represents the sum of all elements on diagonal line.

In this way, the minimization problem, which is min Ratiocut $(A_1, A_2, \ldots, A_k)$, is converted into $\arg\min_{H} \text{Tr}(H^T L H)$, s.t. $H^T H = \text{I}$.

Above discussion expresses the train of thought about cutting graph by using rigorous mathematics. Since it is easy to obtain L, the goal with regard to above minimization problem is to seek for a certain H which meets given requirements and therefore minimize $\text{Tr}(H^T L H)$.

It should be noted that obtaining all h_i that make up of H is the prerequisite of seeking for H that meets certain requirements. Every h_i is a n $\times$ 1 vector, and the value of every component in the vector has two choices (0 or $\frac{1}{\sqrt{|A_i|}}$), so the fact that every element has two choices means that there is 2^n situations with regard to h_i whose index ranges from 1 to n. Furthermore, there is $(C_k^1)^n$ situations with regard to the whole H matrix. If n is quite large, this minimization problem is undoubtedly a disastrous NP-hard problem. Obviously, we need to avoid seeking for the optimal solution by traversing all situations.

Another method in which h_i is approximately replaced can be used. When k is equal to 2, the second least eigenvector of L can be used to replace h_i (The least eigenvalue of L is 0, its corresponding eigenvector is **1**. Vector is used in subsequent operation which is classification, but vector **1** has no indicating meaning with regard to classification). When the value of k is arbitrary, what we need to do is obtaining k eigenvectors of L which correspond to k eigenvalues of L, among which the biggest one is smaller than the smallest one in the other eigenvalues of L. These k eigenvectors can be used to construct H. Finally, H is standardized as follows:

$$H_{i,j}^{*} = \frac{H_{i,j}}{\left(\sum_{j=1}^{k} H_{i,j}^2\right)^{1/2}}$$

Now H matrix is finished and the remaining task is clustering samples. It should be remembered that our goal is not seeking for the minimum value of $\text{Tr}(H^T L H)$ but obtaining H which can minimize $\text{Tr}(H^T L H)$. Hence, clustering should be executed on columns. Usually, K-means method is chosen to do this.

fridaya@uef.fi

4.3 A Case Study

It is assumed that the similarity matrix of a group is represented as follows:

$$w = \begin{bmatrix} & 1 & & 16.5 & 17 & 3.4 & \\ 1 & & 9.8 & & 2.5 & & 8 \\ & 9.8 & & 6.8 & 11 & 3.2 & 4.7 \\ 16.5 & & 6.8 & & & 10 & 13 \\ 17 & 2.5 & 11 & & & & 11.2 \\ 3.4 & & 3.2 & 10 & & & \\ & 8 & 4.7 & 13 & 11.2 & & \end{bmatrix}$$

Then its corresponding Laplace matrix is obtained as follows:

$$L = \begin{bmatrix} 37.9 & -1 & & -16.5 & -17 & -3.4 & \\ -1 & 21.3 & -9.8 & & -2.5 & & -8 \\ & -9.8 & 35.5 & -6.8 & -11 & -3.2 & -4.7 \\ -16.5 & & -6.8 & 46.3 & & -10 & -13 \\ -17 & -2.5 & -11 & & 41.7 & & -11.2 \\ -3.4 & & -3.2 & -10 & & 16.6 & \\ & -8 & -4.7 & -13 & -11.2 & & 36.9 \end{bmatrix}$$

It is supposed that auditors are informed that this group should be divided into 3 subgroups in advance. So 3 is assigned to k, which means the three eigenvectors corresponding to three least eigenvalues would be obtained, they are:

$$h_1 = \begin{bmatrix} -0.378 \\ -0.378 \\ -0.378 \\ -0.378 \\ -0.378 \\ -0.378 \\ -0.378 \end{bmatrix}, h_2 = \begin{bmatrix} 0.1167 \\ -0.5361 \\ -0.1902 \\ 0.179 \\ -0.1334 \\ 0.7584 \\ -0.1944 \end{bmatrix}, \text{and } h_3 = \begin{bmatrix} -0.4966 \\ 0.5824 \\ 0.1097 \\ -0.1801 \\ -0.3777 \\ 0.4653 \\ -0.103 \end{bmatrix}, \text{respectively.}$$

After $h_1, h_2,$ and h_3 are normalized, their values are changed to the following ones:

$$h_1^* = \begin{bmatrix} -0.5954 \\ -0.4309 \\ -0.8647 \\ -0.8301 \\ -0.6863 \\ -0.391 \\ -0.8643 \end{bmatrix}, h_2^* = \begin{bmatrix} 0.1838 \\ -0.6122 \\ -0.4351 \\ 0.3931 \\ -0.2422 \\ 0.7845 \\ -0.4445 \end{bmatrix}, \text{and } h_3^* = \begin{bmatrix} -0.7822 \\ 0.6639 \\ 0.2509 \\ -0.3955 \\ -0.6858 \\ 0.4813 \\ -0.2355 \end{bmatrix}$$

fridaya@uef.fi

By taking h_1^*, h_2^*, and h_3^* as inputs of k-means algorithm, auditors could obtain the following classification result:

$$\{1,2,3,4,5,6,7\} \rightarrow \{\{1,4,5\},\{2\},\{3,6,7\}\}$$

which means that the first community contains 3 persons, and the second community contains only one person, while the third community contains 3 persons.

By this time, auditing department can take use of above results to carry on their audit work with low deviation. Auditing department could audit every community, as if it is a single person.

5 Conclusion

Many groups of individuals constitute the whole society and there are complex relations between different groups as well as within the same group. It might be possible to discover these relations because structures exist in complicated network that is made up of many vertices, which can be used to represent individuals in real world. It is urgent to develop efficient method to discover communities within a certain group since relation among individuals can be used to conceal economic crimes and community discovery focuses on finding out relation among individuals. Since those methods involving iteration process have fatal defect that computer system will halt when that kind of algorithm dealing with a large amount of data is running in it, we introduce a community discovery method based on spectral clustering, whose potential application prospect in audit may be attractive. This method shows high speed, from which auditing department can benefit since auditors often face massive data. It should be pointed out that spectral clustering can bring analysis deviation, but this deficiency seems not to prevent it from being widely used, especially by such organizations that cope with tons of data as auditing department.

Acknowledgement. This study was funded by Government Audit Research Foundation of Nanjing Audit University.

References

1. Donetti, L., Munoz, M.A.: Detecting network communities: a new systematic and efficient algorithm. J. Stat. Mech.: Theor. Exp. **10**, P10012 (2004)
2. Newman, M.E.: Fast algorithm for detecting community structure in networks. Phys. Rev. E Stat. Nonlinear Soft Matter Phys. **69**(6), 0661331–0661335 (2004)
3. Clauset, A.: Finding local community structure in networks. Phys. Rev. E **72**(2), 026132 (2005)
4. Capocci, A., Servedio, V.D.P., Galdarelli, G., Colaiori, F.: Detecting communities in large networks. Phys. A **352**(2–4), 669–676 (2005)
5. Duch, J., Arenas, A.: Community detection in complex networks using extremal optimization. Phys. Rev. E **72**, 027104 (2005)

fridaya@uef.fi

6. Chen, M., Kuzmin, K., Szymanski, B.K.: Community detection via maximization of modularity and its variants. Comput. Soc. Syst. **1**(1), 46–65 (2014)
7. Gui, Y., Wenping, Z., Wenjian, W., Haojie, Z.: Community detection algorithm based on weighted dense subgraphs. J. Softw. **28**(11), 3103–3114 (2017)
8. Xinzhuan, H., Jingfeng, G., Shiqian, H., Xiao, C.: Community detection algorithm of node similarity and node participation degree in signed networks. J. Chin. Comput. Syst. **38**(10), 2275–2280 (2017)
9. Wenwei, Z., Jingke, X., Zhixiao, W.: Hierarchical agglomerative community detection algorithm based on similarity modularity. J. Syst. Simul. **29**(5), 1028–1032 (2017)
10. Hangyuan, D., Wenjian, W., Liang, B.: An overlapping community detection algorithm based on centrality measurement of network node. J. Comput. Res. Dev. **55**(8), 1619–1630 (2018)
11. Jing, C., Yun, W.: Research on pairwise merging community detection algorithm based on similarity. J. Chin. Comput. Syst. **39**(7), 1487–1491 (2018)
12. Xiangping, Z., Zhiyong, Y., Guangbin, W.: Community discovery in location network. Comput. Sci. **45**(6), 46–50 (2018)
13. Zhang, H., Yongke, W., Zhizhuo, Y., Quanming, L.: Community detection method based on multi-layer node similarity. Comput. Sci. **45**(1), 216–222 (2018)

fridaya@uef.fi

Refinement and Trust Modeling of Spatio-Temporal Big Data

Lei Zhang[1,2(✉)]

[1] MOE International Joint Lab of Trustworthy Software, East China Normal University, Shanghai 200062, China
lzhang@ce.ecnu.edu.cn
[2] Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China

Abstract. The conventional studies of spatio-temporal data models and their big data applications cannot reliably reflect the large volume, heterogeneity and dynamics of spatio-temporal big data. In this paper, the structure and function expression of spatio-temporal metadata is analyzed. With fused and normalized spatio-temporal reference and data structure, the constraint rules of spatio-temporal big data refinement are proposed. Using the domain specific modeling (DSM) and the data granulation algorithms, an object-oriented modeling language, the thrust modeling of spatio-temporal big data, and the aggregated status correlation of unified model data are established. This work utilizes the trust modeling theory and the spatio-temporal data processing methods and defines a case study that converts spatio-temporal data into dynamic complex big data. This research paves the way for the trust modeling and validation of spatio-temporal big data.

Keywords: Spatio-temporal big data · Refinement · Trust modeling · Domain specific modeling · Granulation

1 Introduction

The Space-time Cube Model, Sequent Snapshots Model, Update Model, and Space-Time Composite Model proposed in the Ph.D. dissertation of Dr. Gail Langran in 1992 provide fundamental modeling theories for the studies of spatio-temporal data [1–4]. Since 2008, new technologies and applications, such as the Internet of Things, cloud computing, smart cities, big data, artificial intelligence, etc., emerge constantly. In the field of earth observation, the global networking enabled by the BeiDou Navigation Satellite System has triggered unprecedented activities in the key technologies and industrial applications of global navigation satellite systems [5]. Remote sensing satellites in the fields of meteorology, oceanography, resources, environment, etc. are developed rapidly by many countries and commerce. Geographic information technologies have also enriched the geographical environment, time and space identifications, and multidimensional attributes, due to the rapid development of navigation satellite system and remote sensing technologies. The "spatial information technology" consists of the conventionally named "3S", i.e., Global Navigation Satellite System

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 132–144, 2019.
https://doi.org/10.1007/978-3-030-22868-2_10

fridaya@uef.fi

(GNSS), remote sensing (RS) and geographic information system (GIS), as well as computer and communication technologies. The dynamic big data with spatial and temporal attributes can be tremendously expanded by all kinds of natural and social observation data which are fused, shared and processed from the spatio-temporal data generated by earth observation and public data (such as city cameras, social media, and personal activities) [6]. Academia has achieved much progress in earth observation data acquisition, processing and application, which promotes the effective development of the spatial information industry. The synthesis and mining of spatio-temporal big data is a critical problem to be solved for the current science and technology development.

The spatio-temporal data model provides a formalism consisting of a notation for describing data of interest, set of operations for manipulating these data and a set of predicates. Spatio-temporal data model is a data model representing the temporal evolution of spatial objects over time [3–5]. If these evolutions are continuous, one speaks about spatio-temporal objects and represents them by spatio-temporal data types like "3S" point and mobile line, etc. Spatio-temporal data types enable the user to describe the dynamic behavior of spatial and transportation objects over time. Mobile object is a spatio-temporal object that continuously changes its location and/or shape. Depending on the particular mobile objects and applications, the movements of mobile objects may be subject to constraints [6].

With the efforts by worldwide institutes and researchers, there has been tremendous progress in spatio-temporal data modeling for the past twenty years. However, due to the diversity and complexity of the reality reflected by the spatio-temporal data and the large volume, heterogeneity and dynamics of the spatio-temporal data, the studies of the models are still defective. To process spatio-temporal data better, the studies of more reliable, efficient and practical technologies in big data processing, management and application are demanded. This paper presents the analysis, design, and study of spatio-temporal metadata, data fusion and refinement, and trusted modeling of spatio-temporal data. This work also expands the research directions of highly trusted modeling and its validation based on data resources with spatio-temporal attributes, and then establishes and completes the fundamental theoretical system for the highly trusted spatio-temporal infrastructure construction, management modes, techniques and operating mechanisms, application service, and standard specification.

2 Structure Design of Spatio-Temporal Data

2.1 Spatio-Temporal Metadata

Measurable and unmeasurable data with spatial and temporal relationships are mainly from GNSS, RS, and sensors, etc. [7]. The formats, processing methods, and expressions of these spatio-temporal data are not the same, showing multi-dimensional, coupled and nonlinear properties [8]. On the other hand, the text, audio, and video data types formed on the Internet of Things and social networks have a linear relationship in the expression of the time stamp. In a geographic information system, spatio-temporal data are defined as natural, humanities and social data based on normalized spatio-temporal reference, directly or indirectly related to geographic elements or phenomena,

fridaya@uef.fi

as shown in Fig. 1. Described by "cubes", spatio-temporal data are geographic space data with time attributes, mainly the geometric and physical characteristics of the earth's surface that can be sensed, recorded, stored, analyzed, and utilized.

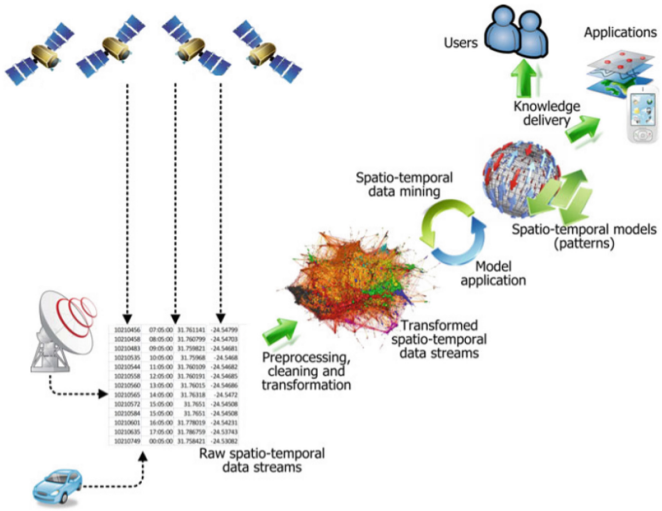


Fig. 1. Spatio-temporal data source

The metadata of spatio-temporal data is the basis to build up spatio-temporal big data in Fig. 2. The spatial location and observation time information acquired from a GNSS receiver can be expressed as:

$$\text{GNSS info} = f(\varphi, \lambda, H, t_p, \pi), \tag{1}$$

where (φ, λ, H) represent the spatial location (latitude and longitude coordinates and elevation value) of the GNSS receiver, respectively, and (t_p, π) represents the observation time and other information of the receiver, respectively.

The object radiation information obtained from the earth observation by air, space, and ground remote sensing devices can be expressed as:

$$\text{RS info} = f(x, y, z, \lambda, t_R), \tag{2}$$

where (x, y) are the spatial location values, (z) is the sensed value with respect to the location (x, y), which is related to the resolution of the spatial location, (λ) is the wavelength of the electromagnetic wave employed, which is related to the resolution of the radiation, and (t_R) is the periodicity of repeated observations on the same object (related to time resolution).

fridaya@uef.fi

GIS can be expressed as:

$$\text{GIS info} = \{i, j, T(A), t_G\}, \tag{3}$$

where (i, j) are the spatial position coordinates used by the system, T(A) is the spatial attributes relevant to the system coordinates (i, j), and (t_G) is the temporal attributes of the system information. The humanities and social data from the Internet of Things and social networks can be written as:

$$\text{More info} = \{C(\alpha, \beta, \gamma), S(\gamma, \psi, \omega), t, \pi\}, \tag{4}$$

where C(α, β, γ) and S(γ, ψ, ω) are the assemblies of the humanities and social attributes, respectively, t is the time stamp of the information acquisition and π represents other multi-source data.

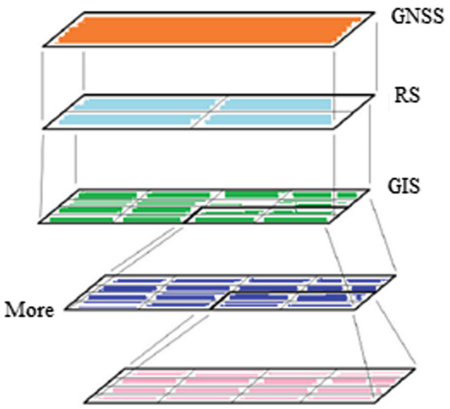


Fig. 2. Structure design framework

Hence, the spatial relationship K(φ, λ, H), temporal relationship T(t_p, t_G), radiation characteristic (λ), humanities attributes C(α, β, γ), social attributes S(γ, ψ, ω) and other data π of the information carriers can be unified and described. The unified function expression of spatio-temporal metadata can be expressed as:

$$\text{Spatio-temporal Metadata} = f\{K(\varphi, \lambda, H), T(t_p, t_G), (\lambda), C(\alpha, \beta, \gamma), S(\gamma, \psi, \omega), \pi\}. \tag{5}$$

The expression can unify spatio-temporal data of multi-phase, multi-scale, multi-type, and multi-source heterogeneity in the same time and space coordinates by dynamic management, comprehend the resolutions of time, space and radiation and the refinement of geography identifications, and realize the effective record, load, share and exchange of spatio-temporal data.

fridaya@uef.fi

2.2 Spatio-Temporal Data Stream

Time is the important and omnipresent variable that is subject to the spatio-temporal data [9]. We use seconds, minutes, days, months, and years to the position and navigation. That is to say that the time-domain and spatial-domain are integrated into the spatio-temporal data. To high-level management, time intervals of quarters of year or a few years could be the meaningful spatio-temporal data on the basis of which one develops any predictive model.

The spatio-temporal data stream items could be viewed as relational tuples with one important distinction: they are time-ordered. Event time of the spatio-temporal data stream tuple is defined by temporal attributes. Shape and location of an object of interest described by a spatio-temporal data stream tuple defined by spatial attributes [10].

Definition 1 (Time Domain): A time domain $\mathbb{T}$ is a pair $(\mathrm{T}; \leq)$ where T is a nonempty set of discrete time instants and $\leq$ is total order on T.

Spatial domain is a set of homogeneous object structures (values) which provide a fundamental abstraction for modelling the geometric structure of real-world phenomena in space. In order to locate an object in space, the embedding space must be defined as well. The formal treatment of spatial domain requires a definition of mathematical space.

Definition 2 (Spatial Domain): A spatial domain $\mathbb{D}_\sigma$ is a set of spatial objects with simple or complex structure.

The spatio-temporal data stream tuple t represents an event, i.e., an instantaneous fact capturing information that occurred in real-world at time instant τ, defined by event timestamp. Event timestamp offers a unique time indication for each tuple, and therefore cannot be either undefined (i.e., event timestamp cannot have a null value) or mutable. In the sequel, we consider explicitly timestamped data streams, ordered by the increasing values of their event timestamps [11]. We distinguish between raw streams produced by the sources and derived streams produced by continuous queries and their operators. In either case, we model the individual spatio-temporal data stream elements as object-relational tuples with a fixed spatio-temporal schema. All spatio-temporal objects share a set of common operations [11]. For this reason, all spatio-temporal data types shall implement common operations, as well as operations that are specific for a particular spatio-temporal data type.

Definition 3 (Trajectory cluster): We denote a trajectory cluster $\mathbb{C}_f$ for the objects in the spatio-temporal trajectory stream $\mathcal{G}_v$ as a 2-tuple $(\mathbb{Q}_m, [\tau_i, \tau_j])$, where $\mathbb{Q}_m$ is the set of the objects in the trajectory cluster, τ_i and τ_j is the start and end time of the trajectory cluster, respectively:

$$\mathbb{C}_f : \exists \mathbb{C}^{i_k}_{[\tau_k,\tau_{k+1}]} | \mathbb{C}^{i_k}_{[\tau_k,\tau_{k+1}]} \in \mathbb{C}_{l_{[\tau_k,\tau_{k+1}]}, \mathbb{Q}_m \subseteq \cap_{k=i}^{j-1} \mathbb{C}^{i_k}_{[\tau_k,\tau_{k+1}]}} \left(\|\mathbb{Q}_m\| \geq min_{\mathbb{Q}} \right), \quad \text{where} \quad \mathrm{k} = \mathrm{i}, \cdots,$$

$\mathrm{j} - 1$, and $min_{\mathbb{Q}}$ is objects number threshold. The line-segment-clusters $\mathbb{C}_{l_{[\tau_k,\tau_{k+1}]}}$ of continuous time intervals constitute a closed trajectory cluster $\mathbb{C}_f$.

fridaya@uef.fi

3 Function Expression of Spatio-Temporal Big Data

3.1 Unifying of Spatio-Temporal Reference

According to the accurate function expression of spatio-temporal metadata, spatio-temporal big data is only worth researching and utilizing when the data are acquired simultaneously, assembled, and unified. In recent years, the compatibility and interoperability of China's BeiDou Navigation Satellite System, the USA's GPS and Russia's GLONASS have been improved continuously, which helps the establishment of a terrestrial reference close to International Terrestrial Reference System (ITRS), and the reduction of the difference between the time systems in the navigation messages of different systems. Especially for near ground applications of spatio-temporal big data, the coordinate systems must be consistent, at least consistent to the ground tracking stations of GNSS, i.e., the time must be synchronous.

Spatio-temporal data are a set of spatio-temporal metadata with a certain periodicity, which is broadcast according to certain encoding rules. Space, time and attribute data are coupled and self-synchronized according to the encoding rules, which realizes the unifying of the spatio-temporal reference of the system itself. Spatio-temporal data are usually broadcast by multiple satellites or signal source systems to single or multiple receiving platforms. This needs at least a GNSS multi-mode receiving system to monitor the deviation of the coordinate system and to send to users for correction or reference as prior information of user navigation parameters. If the time systems are not consistent, multi-system tracking stations are also needed to monitor and broadcast, or more model parameters are needed for real-time estimation and correction. To unify space reference, one needs to build a complete mathematic conversion relationship to eliminate the errors [7].

Assume there are n subsystems with certain conversion relationships in spatio-temporal big data. Their initial states are not all the same. One of the subsystems is randomly picked as the spatio-temporal system platform and copied to the rest (n − 1) subsystems, forming a single-direction coupled series of responding systems. One of the spatio-temporal system platforms (x_0, y_0, z_0) drives all the other (n − 1) responding systems. When the initial states are all the same, the spatio-temporal data of original six-dimensional i-th subsystem $\left(x_i, y_i, z_i, x_i^{'}, y_i^{'}, z_i^{'}\right)$ can be reduced to data of a four-dimensional synchronized spatio-temporal system $\left(x_i, x_i^{'}, y_i, z_i\right)$. This type of synchronized spatio-temporal network algorithm can realize the unifying of the space and time inside spatio-temporal data and fulfill the space-air-ground integration of GNSS and the Internet of Things, providing theoretical support for the spatial and temporal resolution improvement.

3.2 The Characteristics of Spatio-Temporal Big Data

The spatio-temporal description and expression of big data from real-world earth observation require correlation and packaging of attributes, functions and their correlations through spatial-temporal relationships to realize the physical expression of the computer world.

fridaya@uef.fi

Spatio-temporal objects have spatial locations and distribution characteristics. Their data organization needs not only a conventional keyword and sub-keyword index but also require building up a space index for inquiring and search requirements. Spatio-temporal objects have complicated topological relationships, the expression of which is the fundamental of spatial analysis and inference but increases the complexity of maintaining the consistency and completeness of spatio-temporal data. The unstructured characteristic of spatio-temporal data makes it difficult to store and express directly in a general relational database. The geometric coordinate data of spatio-temporal objects are indefinite in length, and there may be complex nested relationships between objects, which cannot meet the paradigm requirements of relational data models. Spatio-temporal data also have characteristics such as massive, multi-source, multi-phase, multi-scale, and non-stationary. This requires a better understanding of the expression of spatio-temporal big data and building of a theoretical framework for spatio-temporal big data mining.

Spatio-temporal big data reflects the temporal and spatial laws of human activities through technologies such as earth observation, that is, the temporal change trend and spatial distribution rule. The support from technologies such as the Internet of Things, virtual reality, and space-air-ground integration observation make spatio-temporal big data more shared and open [12]. The social public is both the users of open big data and the suppliers of big data sharing sources in the spatio-temporal big data sharing scenario. There are hundreds of millions of sensors in the observations of the social public, and the types of natural and social observation data collected vary greatly, which significantly improves the dynamics and richness of earth observation information sources. The social public can effectively participate in the spatio-temporal big data of earth observation, which is also beneficial to the intelligent synthesis of spatio-temporal big data, automatic update of multi-scale spatio-temporal databases, cleaning, analysis, mining, visualization, applicability, and decision supportability of spatio-temporal big data.

4 Refinement of Spatio-Temporal Big Data

4.1 Data Structure Normalization and Fusion

Data always generate effects in certain time and space, which is more prominent when combined with the visual expression of environmental (geographic) elements [13]. Among the spatio-temporal data generated by earth observation technology, the GNSS is instantaneous spatio-temporal data, from which geographic information can be directly acquired. High accuracy is the essential indicator of the quality of such data. Remote sensing (RS) data can be interpreted into the size, shape and spatial distribution characteristics, attribute characteristics, and dynamic characteristics of the objects, which are the most intuitive descriptions and time-sensitive data set. However, the image resolutions of RS are not the same, and the minimum granularity of identified objects is often inconsistent with the GIS data. Therefore, the data structure of GIS geographic information not only needs to support image processing directly but also needs to support the hierarchical fusion of various data of different resolutions.

fridaya@uef.fi

Assume a resolution of 2^j in computer data processing, and Z and R to represent the sets of integer and real numbers, respectively. V_j is a spatio-temporal data subset of $L^2(\mathrm{R})$, i.e. $V_j \subset L^2(R).A_2^j$ is a linear projection operator approximating to $f(\mathrm{t})$ with a resolution of 2^j, i.e. $f(\mathrm{t}) \in L^2(R)$. Since spatio-temporal metadata have strong spatial and temporal correlations, multi-resolution analysis of the data forms the following characteristics:

(1) *Lossless*: One can map the higher resolution spatio-temporal metadata f_j to a lower resolution f_{j+1} with an analysis operator, and thus decomposite the spatio-temporal data f into a series of spaces with gradually reduced resolutions $\{f_0 = f, f_1, f_{2,\cdots}\}$. The information lost during the spatio-temporal big data mapping process $f_j \rightarrow f_{j+1}$ is represented by the detail characteristics $(y_j, \mathrm{j} = 0, 1, 2, \cdots)$. Conversely, the original spatio-temporal big data can be reconstructed by $\{f_0 = f, f_1, f_{2,\cdots}\}$ and $(y_j, \mathrm{j} = 0, 1, 2, \cdots)$ to realize the structure fusion and lossless expression of spatio-temporal big data.

2. *Scalability*: Assuming $\{V_j\} \subset L^2(R)$ is a multi-resolution analysis of $L^2(R)$, then there exists a multi-scale function $\varphi(\mathrm{x}) \in L^2(R)$, the scale function series of which can be defined as $\Phi_j(x) = 2^{-j}\Phi(2^{-j}x)$, $\mathrm{j} \in \mathrm{Z}$. The translation and scaling system $2^{\frac{j}{2}}\varphi_j(x - 2^j n)$ is a canonical orthogonal basis of V_j. These two characteristics help to establish the topologic relationship between point, line, and surface metadata of spatio-temporal big data under the premise of unified spatio-temporal reference and construct a data structure that is compatible with vector and grid.

Hence, the geometric properties or attributes of the same spatio-temporal object are usually heterogeneous in different scales. However, the inclusion relationship between the region in the spatial resolution adjustment and its sub-regions can be expressed based on the same data structure. The fusion and normalization of data structures can establish multi-dimensional heterogeneous basic data organization units and an organizational system correlated by the topological relationship between the parent and child regions. This is also a comprehensive expression of the hierarchical-type and the framing-type organizations of spatial-temporal data.

4.2 The Constraint Rules of Spatio-Temporal Big Data Refinement

There is distortion or loss during the observation of spatio-temporal data. After the spatio-temporal reference is unified, the time series analysis can also introduce additional errors, which affects the accuracy of the fusion model of spatio-temporal big data. In the study of spatio-temporal big data, pre-established data refinement rules and pre-controlled quality of spatio-temporal big data can ensure the reliability and trust of the spatio-temporal data of synchronized time sequence. It is shown in Fig. 3. In the field of mathematics and programming research, a state-based approach often describes the instability of system dynamic equations of spatio-temporal big data applications, i.e., the linear relationship contains pure random trends [14].

fridaya@uef.fi

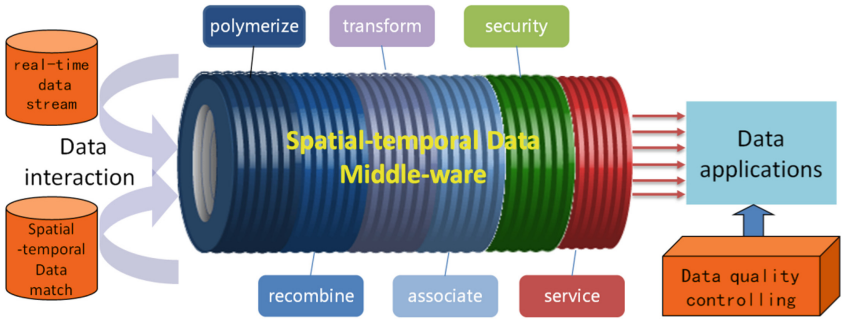


Fig. 3. Refinement processing of spatio-temporal data

Assume a data series x(t) is formed by normalization and unifying of spatio-temporal reference, with a time interval of $t = n\tau$ $(1 \le n \le \infty)$. For a general τ and a specified integer m $(d \le m \le 2d + 1)$, there is a smooth mapping f that meets:

$$f[x(n\tau), \cdots, x(n + m - 1)\tau] = x((n + m)\tau) \quad n = 1, 2, \cdots, \infty, \tag{6}$$

where m is the analytic dimension of spatio-temporal data.

In data refinement, constraint rules are proposed for the convenience of computer multi-resolution analysis. A spatio-temporal big data refinement function $\overline{f_N}$ is constructed for a given finite time sequence (xnτ), that is,

$$\overline{f_N}[x(n\tau), \cdots, x(n + m - 1)\tau] = x((n + m)\tau). \tag{7}$$

In this way, the infinite approximation constraint in the spatio-temporal data refinement is realized, i.e. $f = \overline{f_\infty} = \lim_{N\to\infty} \overline{f_N}$.

Assume m and ν represent the spatio-temporal big data refinement dimension and the time series delay time slice, respectively, and the occurrence probability of each symbol sequence is P_1, P_2, $\cdots$, P_k calculated by the program. The different symbol sequences in the k-th spatio-temporal data in spatio-temporal big data time series x(t), i = 1, 2, ⋯, n can be defined according to the form of Shannon information entropy:

$$H_P(m) = -\sum_{v=1}^{k} P_v \ln P_v. \tag{8}$$

$H_P(m)$ can be normalized using $\ln(m!)$, i.e. $0 \le H_P(m) \le H_P(m)/\ln(m!) \le 1$. The magnitude of H_P value represents the degree of randomization of the time series x(t). The smaller the value of H_P, the more regular the spatio-temporal big data time series are; the larger the value, the more random. The change in H_P reflects and amplifies the changes of data details of x(t) in the spatio-temporal big data refinement process.

fridaya@uef.fi

5 Trusted Modeling of Spatio-Temporal Big Data

There are many good algorithms and study mechanisms for the description of geometric images in spatio-temporal big data, such as the construction of conventional points, lines, surfaces, and radiation in space. However, there are still major limitations to the differentiation management and organization synthesis of the attribute information and geometry in spatio-temporal data cubes. The multi-scale of spatial information reflects the accuracy descriptions of spatial resolution and radiation resolution in remote sensing images. The spatial relationship depends on the structure or mathematical topology described by the spatial entities. The research of pure graphic synthesis is often insufficient for high-accuracy time resolution analysis, and thus the automatic relationship establishment in computer programs and database management is weak. Spatio-temporal big data are rich in the quantity, quality, and attribute indicators of spatial entities. Conventional attribute information analysis is a physical process, which automatically merges and reconstructs selected objects, but cannot guarantee the semantic accuracy of the information description of spatio-temporal data.

The trusted modeling of spatio-temporal big data needs to analyze the attributes importance, attribute uncertainty, attribute table consistency, and attribute reliability in spatio-temporal metadata and their big data [15]. It needs to study the impact of attribute reliability on application decision, to refine data, spatio-temporal attributes and attribute dependencies, and to establish data correlation of spatio-temporal big data. It also needs to evaluate the absolute and relative uncertainties in the data structure, generate decision-making algorithms from data, discover paradigms and logical relationships in spatio-temporal data, and generate a minimum decision and classification algorithms, etc. It then guides the classification and fuzzy boundary division of uncertain spatio-temporal data (especially videos) and realize the expression and processing of data containing multiple spaces, times, radiation, and topological and mapping relationships. It is shown in Fig. 4.

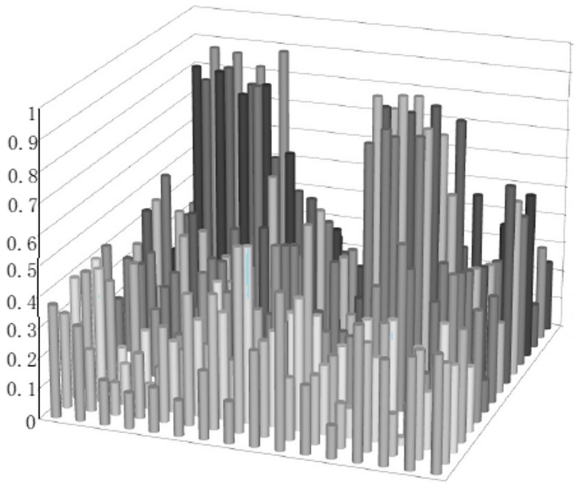


Fig. 4. Data normalization of trusted modeling

fridaya@uef.fi

For integrating spatio-temporal big data and trusted modeling, the trusted modeling method of the domain specific modeling (DSM) is proposed [16–18]. The trusted modeling is used to define spatio-temporal meta-model,to describe contributes,associations and operations of the spatio-temporal objects. The trusted modeling class contributions and the process of the spatio-temporal objects extraction are defined as shown below.

First, add time, space, radiation, humanities, social and other attribute data to UML to form the spatio-temporal metadata set X described in Sect. 1 and the refined data set Y described in Sect. 3.2. If the spatio-temporal big data function D (where X and Y are correlated) meets $\forall x_i \in X$, thus $\mathrm{D}(x_i) = y_i \in Y$, and then D is the tag tuple of X in Y, marked as D(X, Y) or $[x_1 : y_1, \cdots, x_k : y_k]$. A spatio-temporal big data UML model can be defined as:

$$D_{UML} = \{\mathrm{K}(\varphi, \lambda, \mathrm{H}),\ \mathrm{T}(\mathrm{tp}, \mathrm{tG}),\ (\lambda),\ \mathrm{C}(\alpha, \beta, \gamma),\ \mathrm{S}(\gamma, \psi, \omega),\ \pi\}, \tag{9}$$

which contains time identifier, spatial identifier, radiation identifier, humanities identifier, social identifier, and other attribute identifier data sets.

There are both dependencies and certain generalizations between the datasets. Formal semantics can be given by specifying the state of the spatio-temporal big database and the corresponding DSM model. In the spatio-temporal big data trusted modeling, given any data aggregation, the mapping of the instance set can be implemented based on the DSM model, and the spatio-temporal correlation and constraint relationship can be determined.

The spatio-temporal big data clustering with granular computing concepts is created which used gravitational forces as the premise for finding relations between all spatio-temporal metadata. With the premise that similar trusted modeling must belong to the same spatio-temporal big data granules, gravitational forces are used to group together closely related data and find such relevant spatio-temporal granules [19, 20]. When the trusted modeling of the spatio-temporal big data is performed, the refinement results are very promising as obtained performance is very comparable with recent state of the art clustering algorithm, showing that the trusted modeling of granular computing based on the spatio-temporal big data, when applied, does obtain great performance.

6 Summary

This work expands the simple data modeling study with spatial and temporal correlations, to the study of the spatio-temporal big data generated by the earth observation technology and realizes the cross-study and innovation of spatio-temporal big data and highly trusted modeling, especially for the various data sets with space, time, radiation and their topologies and mapping relationships. This research direction can thoroughly facilitate the assembling and strong correlation of global satellite navigation data, geodetic remote sensing data, geographic information identification data, and social public data, etc., and evaluate the uncertainty of spatio-temporal big data with the support of the theoretical system.

fridaya@uef.fi

The goal of this paper is that the classification of spatio-temporal objects refined to the big data categories or classes, and that the trusted modeling and the granular computing approaches based on the spatio-temporal big data are built. In what follows, we consider a specific class of the spatio-temporal big data, i.e. knowledge mining in the context of the trusted data streams produced by the spatio-temporal objects. In other words, we will focus on relevant methodologies and algorithms based on the spatio-temporal big data, fundamental for trusted modeling of trajectories based on their proximity in spatio-temporal domain. So, it will boost the development of highly trusted software and system platforms in the spatial information industry.

Acknowledgments. I would like to express my gratitude to National Key R&D Program of China (No. 2018YFE0101000) for the opportunity given to us by providing us grants to perform our research.

I would also like to thank Prof. Yunxuan Zhou and Prof. Jifeng He in East China Normal University, and all persons who have contributed to the success of our research, either directly or indirectly.

References

1. Pelekis, N., Theodoulidis, B., Kopanakis, I., et al.: Literature review of spatio-temporal database models. Knowl. Eng. Rev. **19**(3), 235–274 (2004)
2. Rocha, L.V.D., Edelweiss, N., Iochpe, C.: Geo Frame-T: a temporal conceptual framework for data modeling. In Proceedings of the 9th ACM International Symposium on Advances in Geographic Information Systems, pp. 124–129. Atlanta (2001)
3. Hunter, G.J., Williamson, I.P.: The development of a historical digital cadastral database. Int. J. Geogr. Inf. Syst. **4**(2), 169–179 (1990)
4. Langran, G., Chrisman, N.R.: A framework for temporal geographic information. Int. J. Geogr. Inf. Geovis. **25**(3), 1–14 (1988)
5. Zhang, W., Li, A.N., Jin, H.A., et al.: An enhanced spatial and temporal data fusion model for fusing Landsat and MODIS surface reflectance to generate high temporal Landsat-like data. Remote Sens. **5**(10), 5346–5368 (2013)
6. Bertone, A., Burghardt, D.: A survey visual analytics for the spatio-temporal exploration of microblogging content. J. Geovis. Spat. Anal. **1**(1), 2 (2017)
7. Zhang, L.: Simulation on C/A codes and analysis of GPS/pseudolite signals acquisition. Sci. China Ser. E Technol. Sci. **52**(5), 1459–1462 (2009)
8. Voisard, A., David, B.: A database perspective on geospatial data modeling. IEEE Trans. Knowl. Data Eng. **14**(2), 226–243 (2002)
9. Yao, J., Vasilakos, A.V., Pedrycz, W.: Granular computing: perspectives and challenges. IEEE Trans. Cybern. **43**(6), 1977–1989 (2013)
10. Oppenheim, A.V., Schafer, R.W.: Discrete-Time Signal Processing, 3rd edn. Pearson Education Limited, Essex (2009)
11. Galić, Z.: Spatio-Temporal Data Streams. Briefs in Computer Science. Springer, New York (2016). ISSN 2191-5768
12. Liu, M., Zhu, J., Zhu, Q., et al.: Optimization of simulation and visualization analysis of dam-failure flood disaster for diverse computing system. Int. J. Geogr. Inf. Sci. **31**(9), 1891–1906 (2017)

fridaya@uef.fi

13. Peuquet, D.J.: Making space for time: issue in space-time data representation. Geo Inf. **5**(1), 11–32 (2001)
14. He, J., Hoare, C.A.R., Sanders, J.W.: Data refinement refined. In: ESOP, pp. 187–196 (1986)
15. Gartner, H., Bergmann, A., Schmidt, A.: Object-oriented modeling of data sources as a tool for the integration of heterogeneous geoscientific information. Comput. Geosci. **27**, 975–985 (2001)
16. Xu, Y., Zhan, H., Yu, J., Sun, L.: Knowledge Modeling Method Based on Domain Specific Modeling Meta-Module, 1005-2895(2012)02-0074-05
17. Kingston, J., Macintosh, A.: Knowledge management through multi-perspective modeling: representing and distributing organizational memory. Knowl.-Based Syst. **13**(2), 121–131 (2000)
18. Wiig, K.M.: Knowledge management: an introduction and perspective. J. Knowl. Manage. **1**, 6–14 (1997)
19. Sanchez, M.A., Castillo, O., Castro, J.R., Rodríguez-Díaz, A.: Fuzzy granular gravitational clustering algorithm. In: North American Fuzzy Information Processing Society, pp. 1–6 (2012)
20. Pedrycz, W., Vukovich, G.: Granular computing with shadowed sets. Int. J. Intell. Syst. **17**(2), 173–197 (2001)

fridaya@uef.fi

Data Analytics: A Demographic and Socioeconomic Analysis of American Cigarette Smoking

Ah-Lian Kor(✉), Reavis Mitch, and Sanela Lazarevski

School of Computing, Creative Technologies, and Engineering,
Leeds Beckett University, Leeds LS6 3QS, UK
A.Kor@leedsbeckett.ac.uk

Abstract. This study attempts to model smoking behavior in the United States using Current Population Survey data from 2010 and 2011. An array of demographic and socioeconomic variables is used in an effort to explain smoking behavior from roughly 139,000 individuals. Two regression techniques are employed to analyze the data. These methods found that individuals with children are more likely to smoke than individuals without children; females are less likely to smoke than males; Hispanics, blacks, and Asians are all less likely to smoke than whites; divorcees and widows are more likely to smoke than single individuals; married individuals are less likely to smoke than singles; retired individuals are less likely to smoke than working ones; unemployed individuals are more likely to smoke than working ones; and as education level increases after high school graduation, smoking rates decrease. Finally, it is recommended that encouraging American children to pursue higher education may be the most effective way to minimize cigarette smoking.

Keywords: Data analytics · Socioeconomic analysis · Business intelligence

1 Introduction

Smoking cigarettes significantly increase an individual's risk of cancer, specifically lung cancer [1]. This report attempts to use American Current Population Survey to statistically model smoking behavior for demographic and socioeconomic identifiers in the hope of providing meaningful recommendations to policy makers and anti-smoking campaigners to moderate cigarette smoking in America. First, this report provides a general review of literature surrounding business intelligence and data analytics. These findings are then applied to public health literature to examine how business intelligence and data analytics have been leveraged in the industry. Then, literature surrounding smoking behavior based on socioeconomic and demographic identifiers is reviewed, and hypotheses are formed. The data collection process will then be examined, followed by a brief review of a social science data lifecycle. Two methods of regression analysis are then conducted, presented, and explained. Finally, recommendations for policy makers and anti-smoking campaigns are presented based on findings.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 145–156, 2019.
https://doi.org/10.1007/978-3-030-22868-2_11

fridaya@uef.fi

2 Review of Literature

2.1 Business Intelligence and Public Health

Business intelligence (BI) is defined as an innovation that leverages data and analytics to assist business planning and decision making [2]. In this sense, "BI is both a process and a product" [3]. The author in [4] adds that according to information technology executives, business intelligence systems are among the most encouraging and valuable technologies in the present and future. The growth in popularity of BI and supporting systems has largely been spurred by the increase in the amount of data (i.e. 'Big Data') available to firms [5]. The rising availability of Big Data has allowed BI to venture into different sectors that have rather short histories of data leverage, such as the global public health industry [6, 7]. The United States' public health sector, for example, began investing information technology systems and electronic health records in the early 2000's to make health care more efficient and affordable [8, 9]. In fact, studies conclude that almost two-thirds of health care companies increased BI spending in 2015; one firm even planned to spend approximately $2 million on BI and analytics in 2015 [10]. By efficiently maintaining electronic patient records of diagnosis and outcome, further creating a large national database, physicians have been able to utilize historical data to mitigate potentially dangerous mistakes and misdiagnoses [8, 11]. The wide-spread use of data to inform decision making has also been utilized by health care industry regulators and policy makers [ibid]. Many studies focus on vice deterrence: discouraging consumption of harmful substances like drugs and alcohol. For example, [12] conducted a study that analyzed data from the National Longitudinal Survey of Youth and concluded that high alcohol tax-rates, and television advertising bans on alcohol content, reduced drinking among youth and premature mortality among adults. Similarly, [13] applied data analytics to aid the UK's Medical Research Council to further develop macro-level alcohol-related health policies to deter binge drinking with high tax.

2.2 Demographics, Socioeconomics and Smoking Behavior

Gender: Research has used analytics to explore the relationships between broad demographics identifiers and general behavioral tendencies like substance abuse. The author in [14] found a statistically significant increase in the likelihood of substance abuse, risky sexual behavior, poor academic marks, and aggression among American youth males compared to youth females. Additionally, [15] found a statistically significant difference in smoking rates among working men compared to their female counterparts. These findings support the hypothesis: men will have higher rates of smoking than women when controlling for other demographic and socioeconomic factors.

Race: Racial differences in smoking habit have also been explored empirically. The author in [16] analyzed data from the 2003 American Current Population Survey (CPS) and found that non-Hispanic white Americans were less likely to attempt to quit smoking than their non-Hispanic black counterparts. Further, [17] used 2011 CPS data and observed that both Hispanics and non-Hispanic blacks were more likely to attempt to quit smoking than their non-Hispanic white counterparts. These studies support the

fridaya@uef.fi

research hypothesis: *white civilians will have the highest predicted probability of being smokers when controlling for other demographic and socioeconomic factors.*

Marital Status: Gender differences in risky behavior make it reasonable to expect differences in smoking habits among individuals with differing marital statuses [14]. The author in [18] found that married individuals were least likely to smoke out of any other marital status (divorced, single, or widowed). The author in [19] similarly found that married individuals were least likely to smoke, followed by single, widowed, and divorced individuals, in that order. These studies support the research hypothesis: married individuals will have the lowest smoking rates when controlling for other demographic and socioeconomic factors.

Children: Research has also briefly explored the effect of children on parent smoking behavior. The author in [20] found that knowledge of the harmful effects of secondhand smoke limits smoking in households with at least one child. Yet, [21] found contradicting results in that the presence of children increases emotional and financial stress among parents, which is linked to higher smoking rates. Acknowledging conflicting results, it can be hypothesized: the presence of children on smoking behavior will be either minimal or statistically insignificant.

Employment: From a theoretical view, studying the effect of employment status on smoking rates can be complicated. There is the economic argument that cigarettes are so-called normal goods – goods for which demand and disposable income are positively correlated [22]. Yet, there is also the psychological argument stating that unemployed, discouraged individuals cope with stress with vices like tobacco [23]. Despite this double argument, the majority of studies indicate that joblessness is highly correlated with substance use, including cigarettes [24]. Therefore, there is support for the hypothesis: unemployed individuals are most likely to smoke cigarettes.

Education: Education has been found to significantly decrease smoking rates. The author in [18] found that less educated people had higher rates of physical nicotine addiction (indicated by biomarkers in blood samples) and general smoking rates. Likewise, [25] found that more educated individuals are more likely to be non-smokers and also more likely to be successful at quitting if they did once smoke. These studies support the hypothesis: years of schooling and smoking prevalence will be negatively correlated.

3 Methodology

3.1 Data Collection and Variables

The data for this study was collected from the 2010 and 2011 United States Current Population Survey (CPS), which has been systematically accumulated in the Integrated Public Use Microdata Series (IPUMS) by [26]. This sub-section will present all variables used in this study, as well as the applied selection and recoding processes used to arrive at the final data set.

fridaya@uef.fi

SMOKER: The CPS collected smoking data with the Tobacco Use Supplement (TUS). The data used for this report is comprised of the three most recent surveys available including this supplement. These surveys were conducted January 2011, August 2010, and May 2010. Upon collection, the raw data included 188,119 responses. The SMOKER variable was derived from the TSMKER variable in the CPS TUS. All observations for which the TSMKER variable had a label of 'Indeterminate' or 'Not in Universe' were dropped from this study, as these values designate blank or ambiguous responses. Dropping these observations reduced the sample to 139,012 responses. The remaining responses were used to derive the SMOKER variable used in this report. Never-smokers and former-smokers were labeled as non-smokers for this report. Former-smokers are considered non-smokers because this report only studies current smoking behavior. The two groups of non-smokers were assigned the value (0) for the derived SMOKER variable. Non-daily smokers and everyday smokers were assigned the value (1) for the SMOKER variable. It is advantageous to have the SMOKER variable as a dichotomous variable (i.e. all responses are either 1 or 0) because it is the dependent variable in this study. These advantages will be discussed in more detail in the Data Analysis section.

CHILD: The CHILD variable in this report was derived from the NCHILD variable used in CPS. The NCHILD variable is an integer indicating the number of children present in the household. Observations indicating 1+ children in the household were assigned a (1) for CHILD. All responses indicating zero children in the household were assigned (0).

FEMALE: The FEMALE variable in this report was derived from the SEX variable used by CPS. CPS assigned males (1) and females (2) in the SEX variable. Although the SEX variable used by CPS is already a binary variable, it had to be re-coded into a 1 or 0 dichotomous variable for proper analysis. All of those who indicated they were female were assigned a value of (1) for this report's FEMALE variable; all males were assigned a value of (0). There was no need for a MALE variable in the report because a FEMALE value of (0) implies that one is a male.

BLACK//ASIAN//HISPANIC//(WHITE): This report selected cases in which individuals indicated only a single race, for the three most populous responses (White, Black, Asian) in order to properly control for race demographics. The race variable was derived from a combination of the RACE and HISPAN variables used in CPS. If respondent indicated Hispanic origin, they were assigned a value of (1) for the HISPANIC variable used in this study, and a (0) for the BLACK and ASIAN variables. If a respondent indicated they were black, they were assigned a value of (1) for the BLACK variable used in this study, and a (0) for the HISPANIC and ASIAN variables. If a respondent indicated they were Asian, they were assigned a value of (1) for the ASIAN variable used in this study, and a (0) for the HISPANIC and BLACK variables. Lastly, a value of (0) for HISPANIC, BLACK, and ASIAN variables implies that the individual is white.

fridaya@uef.fi

MARRIED//DIVORCED//WIDOWED//(SINGLE): The MARRIED, DIVORCED, and WIDOWED variables in this report were derived from the CPS's MARST variable. Individuals indicating being married and living with their spouse were assigned a value of (1) for MARRIED and a (0) for DIVORCED and WIDOWED. The same logic was applied to those indicating they were divorced or widowed. Cases for which the respondents indicated being married but not living with their spouse, and those who indicated being separated, were not selected for the study due to ambiguity. Lastly, observations that have a value of (0) for MARRIED, DIVORCED, and WIDOWED were single and never married, so no SINGLE variable was created.

FOREIGN: The FOREIGN variable in this report was derived from the CPS's NATIVITY variable. Cases of unknown nativity were not selected for this study. Individuals who indicated being born outside of the US were assigned a value of (1) for FOREIGN. Individuals who indicated being born in the US were assigned a value of (0) for FOREIGN, regardless of the nativity of their parents. The dichotomy of this variable diminished the need to create a NATIVE variable.

UNEMPLOYED//RETIRED//(WORKING): The UNEMPLOYED and RETIRED variables in this report were derived from the CPS's EMPSTAT variable. Unemployed, experienced workers were assigned a value of (1) for EMPLOYED and a (0) for RETIRED. Unemployed, new workers (e.g. recent university graduates) were not included due to minimal responses. Those not in the labor force due to retirement were assigned a value of (1) for RETIRED and a (0) for UNEMPLOYED. Those employed and working are represented by having values of (0) for both UMEPLOYED and RETIRED.

Code	Label	Jan 11	Aug 10	May 10
000	NIU or no schooling	·	·	·
001	NIU or blank	X	X	X
002	None, preschool, or kindergarten	X	X	X
010	Grades 1, 2, 3, or 4	X	X	X
011	Grade 1	·	·	·
012	Grade 2	·	·	·
013	Grade 3	·	·	·
014	Grade 4	·	·	·
020	Grades 5 or 6	X	X	X
021	Grade 5	·	·	·
022	Grade 6	·	·	·
030	Grades 7 or 8	X	X	X
031	Grade 7	·	·	·
032	Grade 8	·	·	·
040	Grade 9	X	X	X
050	Grade 10	X	X	X
060	Grade 11	X	X	X
070	Grade 12	·	·	·
071	12th grade, no diploma	X	X	X
072	12th grade, diploma unclear	·	·	·
073	High school diploma or equivalent	X	X	X
080	1 year of college	·	·	·
081	Some college but no degree	X	X	X
090	2 years of college	·	·	·
091	Associate's degree, occupational/vocational program	X	X	X
092	Associate's degree, academic program	X	X	X
100	3 years of college	·	·	·
110	4 years of college	·	·	·
111	Bachelor's degree	X	X	X
120	5+ years of college	·	·	·
121	5 years of college	·	·	·
122	6+ years of college	·	·	·
123	Master's degree	X	X	X
124	Professional school degree	X	X	X
125	Doctorate degree	X	X	X

Fig. 1. Possible responses for the education variable

EDUCATION: The *_EDUC* variables in this study required the most extensive recoding process. They were derived from the EDUC variable used in CPS. The range of possible responses as shown in Fig. 1 [26]. The responses indicate the highest level of education attained. Responses with an (X) indicate at least one respondent identified with that label as the highest level of education achieved. Cases of individuals with (Grade 4 or less) as their highest level of education were not included due to the minimal responses. The remaining observations were coded into nine variables (see Fig. 2).

fridaya@uef.fi

CPS CODE(S)	HIGHEST LEVEL OF EDUCATION ACHIEVED	*_EDUC* VARIABLE ASSIGNED A VALUE OF 1*
020	Grades 5 or 6	*GRADE_EDUC*
030	Grades 7 or 8	*MIDDLE_EDUC*
040, 050, 070, 071	Some High School	*HSDROP_EDUC*
073	High School Degree	*HSGRAD_EDUC*
081	Some University^	*UNIDROP_EDUC*
091, 092	Associate's Degree	*ASSOC_EDUC*
111	Bachelor's Degree	*BACH_EDUC*
123, 124	Graduate Degree	*GRAD_EDUC*
125	Doctorate Degree	*DOC_EDUC*
* All other *_EDUC* variables then are assigned a value of 0		
^ College and University are synonomous in the United States		

Fig. 2. 9 variables for EDUCATION

AGE: The AGE variable in this study was derived from the AGE variable used in CPS. Respondents aged (18–64) were selected for this survey, as one can purchase tobacco in the US at age 18 [27]. In order to examine how smoking behavior fluctuates with age increase for this age range, 18 has been subtracted from each CPS AGE. Thus, the new AGE variable represents how many years over the age of 18 a respondent is.

3.2 Data Lifecycle

[28] provides a model for the data lifecycle specifically pertaining to data analysis in social sciences. Given the prevalence of demographic and socioeconomic variables used in this report, it becomes appropriate to adopt this model to study the data lifecycle of CPS. The lifecycle of data is depicted in Fig. 3 [ibid]. The data lifecycle of CPS data begins at the Study Concept. Here, the actual CPS survey is designed and data aggregation is determined. The next stage is Data Collection. Here, the physical CPS survey is distributed to a sample of the American population. The data is then coded for computer readability during the Data Processing stage. This is seen in the CPS's use of codes for its education variable (as shown in the *_EDUC* table shown previously). After coding, Data Archiving occurs and the data is included with previous data. In the case of CPS, this would be adding the survey results to the results of prior surveys, maintaining the same coding scheme. During the Data Distribution stage, public access to data is given. This was seen by the IPUMS's ability to collect the CPS data. The next two stages are included in this report. The CPS data was found on the IPUMS site (i.e. Data Discovery phase) and the Data Analysis phase is what will follow in the next sections. Repurposing the Data would be done by the CPS data collectors in the future if they wished to combine or restructure the data to add to its usefulness.

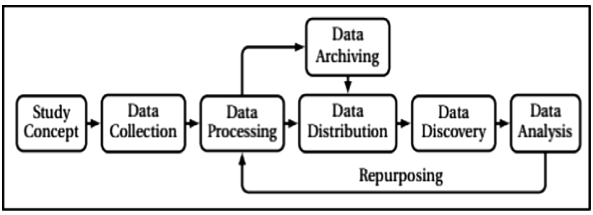


Fig. 3. Data lifecycle

fridaya@uef.fi

3.3 Data Analysis Techniques

This study will report both descriptive and inferential statistics. Descriptive statistics describe and summarize data [31]. This can be done be measuring central tendencies, measuring variance from those central tendencies, and presenting histograms [*ibid*]. Inferential statistics involve employing more complex statistical methods in order to learn about the data, which usually involves a level of confidence [*ibid*]. One of the inferential techniques used in this study was simple linear regression. In simple linear regression, the coefficients of the independent variables included in the regression indicate the increase in the dependent variable for a 1-unit increase in the independent variable. Because this study has a binary variable, *SMOKER*, as the dependent variable, the independent variable coefficients will reflect the increase or decrease in predicted probability of an individual for a 1-unit increase in the independent variable [29]. Because many of the independent variables used in the study are also binary variables, the coefficients show the change in predicted probability of smoking when one is assigned a value of 1 (i.e. a 1-unit increase) for that variable, holding all else constant. The other inferential technique used in this study was binary logistic regression. This method is suitable for regression analysis for which the dependent variable is a binary variable. Instead of changes in predicted probability as given by the simple linear regression, the binary logistic regression will provide a set of relative odds of being a smoker versus the base case [30]. The base case is comprised of all the 'left out' variables for the binary variables, or the 0-case for continuous variables (such as *AGE*). In this report, the base case is an 18-year-old (*AGE* = 0) male that is white, native-born, childless, single, and employed. The *HS_GRAD* variable will also be dropped out of regressions to avoid an error. Thus, the base case male referenced above will have a highest achieved education level as a high school degree. The base case will be referenced many times in the next section.

Label	Mean
SMOKER	0.1624752
CHILD	0.4511625
NO_CHILD	0.5488375
FEMALE	0.4845049
MALE	0.5154951
HISPANIC	0.1129687
ASIAN	0.0450537
BLACK	0.0896325
WHITE	0.7523451
MARRIED	0.5898196
DIVORCED	0.1189394
WIDOWED	0.0191998
SINGLE	0.2720413
FOREIGN	0.1424625
NATIVE	0.8575375
GRADE_EDUC	0.0131859
MIDDLE_EDUC	0.0093733
HSDROP_EDUC	0.0555204
HSGRAD_EDUC	0.2828101
UNIDROP_EDUC	0.1928467
ASSOC_EDUC	0.1071850
BACH_EDUC	0.2211967
GRAD_EDUC	0.1020200
DOC_EDUC	0.0158619
UNEMPLOYED	0.0790220
RETIRED	0.0610739
EMPLOYED	0.8599042
AGE	24.5496360

Fig. 4. Mean of each variable

4 Findings

4.1 Summary Statistics

The mean (average) for each variable is shown in Fig. 4. The base case variables have been added purely for this descriptive section and will be removed for regression analysis. The variation from the central tendencies are irrelevant for binary variables, and thus are not included in the table. For every binary variable, the mean is simply the percentage of observations that have a value of 1 expressed in decimal form. For example, a *SMOKER* mean of roughly 0.162 implies that approximately 16.2% of the sample in this report are classified as current smokers. By the same logic, a mean of 0.451 for *CHILD* implies that 45.1% of the sample have at least one child. Recall that an age of 0 implies one is 18. Therefore, a mean *AGE* of approximately 24.5 implies

fridaya@uef.fi

that the average age of individuals in the sample is roughly 42.5 years old. The table shows a few points of interest: there are more men than women in the sample by a slim margin; whites represent over 75% of the sample; just under 59% of the respondents are married; over 85% of the sample was born in the US; the overwhelming majority are employed; and a high school degree (which is the base case for education) is the most common level of educational attainment. The next section will use regression methods to discover more about the data set.

4.2 Inferential Statistics

As previously mentioned, two types of regression techniques will be used in this report. The following sub-sections will discuss the results of these statistical methods.

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	0.29642	0.00302	98.02	<.0001
CHILD	CHILD	1	0.01676	0.00218	7.69	<.0001
FEMALE	FEMALE	1	-0.03218	0.00193	-16.68	<.0001
HISPANIC	HISPANIC	1	-0.10597	0.00366	-28.92	<.0001
ASIAN	ASIAN	1	-0.02866	0.00524	-5.47	<.0001
BLACK	BLACK	1	-0.06660	0.00344	-19.34	<.0001
MARRIED	MARRIED	1	-0.05459	0.00288	-18.98	<.0001
DIVORCED	DIVORCED	1	0.05911	0.00373	15.85	<.0001
WIDOWED	WIDOWED	1	0.02250	0.00751	3.00	0.0027
FOREIGN	FOREIGN	1	-0.03976	0.00350	-11.36	<.0001
GRADE_EDUC	GRADE_EDUC	1	-0.01951	0.00896	-2.18	0.0293
MIDDLE_EDUC	MIDDLE_EDUC	1	0.01992	0.01012	1.97	0.0490
HSDROP_EDUC	HSDROP_EDUC	1	0.07539	0.00448	16.81	<.0001
UNIDROP_EDUC	UNIDROP_EDUC	1	-0.06108	0.00282	-21.62	<.0001
ASSOC_EDUC	ASSOC_EDUC	1	-0.08208	0.00343	-23.92	<.0001
BACH_EDUC	BACH_EDUC	1	-0.15647	0.00273	-57.25	<.0001
GRAD_EDUC	GRAD_EDUC	1	-0.18235	0.00352	-51.78	<.0001
DOC_EDUC	DOC_EDUC	1	-0.19225	0.00781	-24.61	<.0001
UNEMPLOYED	UNEMPLOYED	1	0.09520	0.00359	26.51	<.0001
RETIRED	RETIRED	1	-0.01088	0.00434	-2.51	0.0122
AGE	AGE	1	-0.00042403	0.00009387	-4.52	<.0001

Fig. 5. Simple linear regression

Simple Linear Regression: The results of simple linear regression are shown in Fig. 5. Recall the base case is an 18-year-old (AGE = 0) male that is white, native-born, childless, single, employed, and has a high school degree as his highest level of education. The intercept is the predicted probability of being a current smoker for this base case. Thus, the base case has a predicted probability of just under 0.30 of being a current smoker. This can be compared to the sample SMOKER mean of just 0.16. The binary independent variable parameter estimates represent the change in predicted probability when assigned a (1) in that category versus the base case, when all else is held constant. Therefore, negative coefficients indicate that having a value of (1) assigned for that category decreases the predicted probability of being a current smoker. For example, a FEMALE coefficient of approximately −0.032 indicates that being a female decreases the predicted probability of being a current smoker by 3.2% points when compared to a male of the same race, age, education, etc. By the same logic, the probability of being a current smoker for high school drop outs is approximately 7.5% points higher than their high school graduate counterparts. Likewise, Hispanics were found to have the lowest predicted probability of being a current smoker out of any race. Perhaps one of the most interesting results is that the predicted probability of smoker for those who have finished high school decreases with each additional level of education. For the AGE

fridaya@uef.fi

variable, the coefficient indicates the change in predicted probability of being a current smoker with a 1-year increase in age. A negative coefficient for this variable indicates that smoking rates decrease with age. Interestingly, every single variable used in the simple linear regression are statistically significant variables at a 95% confidence level, as measured by the p-values in the right-most column in the regression table above.

Binary Logistic Regression: Though a different technique, the binary logistic regression should confirm the findings found in the simple linear regression. A graphical result of this regression is shown in Fig. 6. The odds ratio line at 1.0 indicates identical odds of being a smoker versus the base case for each category. Therefore, variables which have an odds ratio that are placed to the left of the 1.0 odds ratio line have lower odds of being a current smoker versus the base case. The farther to the right of the 1.0 line, the more likely an individual identifying with that category is to be a smoker; the farther to the left, the less likely. The graph depicts some additional interesting findings: those with children are more likely to smoke than those without; both divorced and widowed individuals are more likely to smoke than single ones; high school drop outs have the highest odds of being a current smoker out of any education level; and the effect of education on smoking rates after high school graduation explained in the simple linear regression can be seen on this graph by the southwest-moving odds estimates after HSDROP_EDUC for remaining education levels - doctoral degree holders are over four times less likely to be smokers than high school dropouts.

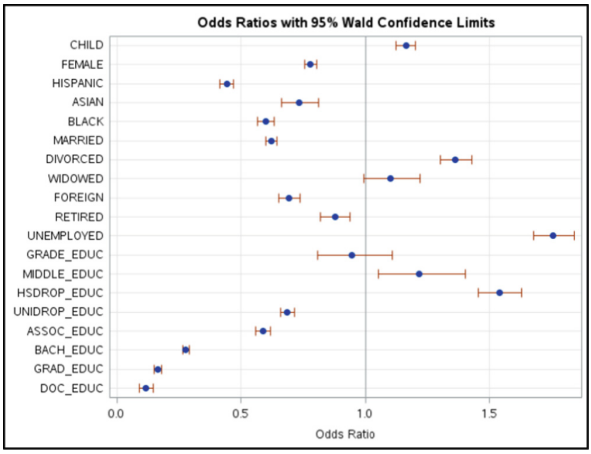


Fig. 6. Binary logistic regression

5 Recommendations

There are a number of conclusions that policy makers can draw from these regression results, if their goal is the reduce smoking rates in the US. Given the increased odds of being a current smoker when at least one child is in the house hold, policy makers may wish to educate parents on the harmful effects of secondhand smoke on children, as well as the potential example the parents are setting. Additionally, with the knowledge that, based on this report's sample, whites are the most likely to smoke than any other race, anti-smoking campaigns may wish to be launched in areas that are primarily

fridaya@uef.fi

white. Friends and family of divorced and widowed individuals may wish to encourage the avoidance of tobacco use in loved ones stressed with such great change. Unemployment offices may also wish to launch anti-smoking campaigns. Perhaps the largest area for intervention is in education. While the large majority of the variables used in this study are out of the control of the respondent (e.g. race, gender, etc.), the individual, to a certain extent, can control the amount of education they attain. Anti-smoking agencies may wish to team-up with anti-high school dropout campaigns if they want to most effectively minimize future smoking behavior.

6 Conclusion

This study attempted to model smoking behavior in the US using Current Population Survey data (CPS) from 2010 and 2011. An array of demographic and socioeconomic variables is used to explain smoking behavior of the surveyed individuals. The study found that those with children are more likely to smoke compared to those without children; females are less likely to smoke than males; Hispanics, blacks, and Asians are all less likely to smoke than whites; divorcees and widows are more likely to smoke than single individuals; married individuals are less likely to smoke than singles; retired individuals are less likely to smoke than working ones; unemployed individuals are more likely to smoke than working ones; and as education level increases after high school graduation, smoking rates decrease. Further studies may wish to use more current data once it becomes available and more advanced statistical techniques. Researchers may also wish to add additional explanatory variables, like residence or income.

References

1. US Department of Health and Human Services: The health consequences of smoking: a report of the Surgeon General (2004). https://www.ncbi.nlm.nih.gov/pubmed/20669512. Accessed 21 Jan 2019
2. Elbashir, M., Collier, P., Sutton, S.: The role of organizational absorptive capacity in strategic use of business intelligence to support integrated management control systems. Acc. Rev. **86**(1), 155–184 (2011)
3. Jourdan, Z., Rainer, R., Marshall, T.: Business intelligence: an analysis of the literature. Inf. Syst. Manag. **25**(2), 121–131 (2008)
4. Fink, L., Yogev, N., Even, A.: Business intelligence and organizational learning: an empirical investigation of value creation processes. Inf. Manag. **54**, 38–56 (2017). ScienceDirect, EBSCOhost. Accessed 2 May 2017
5. Trieu, V.: Getting value from Business Intelligence systems: a review and research agenda. Decis. Support Syst. **93**, 111–124 (2017). ScienceDirect, EBSCOhost. Accessed 2 May 2017
6. Davidson, A.J.: Creating value: unifying silos into public health business intelligence. Front. Public Health Serv. Syst. Res. **4**(2), 1 (2015). Publisher Provided Full Text Searching File, EBSCOhost. Accessed 2 May 2017
7. Thayer, C., Bruno, J., Remorenko, M.B.: Using data analytics to identify revenue at risk. Healthc. Financ. Manag. **67**(9), 72–78, 80 (2013)

fridaya@uef.fi

8. Steward, M.: Electronic medical records. J. Legal Med. **26**(4), 491–506 (2005). Academic Search Complete, EBSCOhost. Accessed 2 May 2017
9. Lyke, R.: Healthcare reform: an introduction. CRS Report for Congress. Congressional Research Service, 7-5700, R40517 (2009). http://fpc.state.gov/documents/organization/126771.pdf. Accessed 2 May 2017
10. Eddy, N.: Analytics Investment among Health Care Organizations Grows (2015). http://www.eweek.com/it-management/analytics-investment-among-healthcare-organizations-grows.html. Accessed 2 May 2017
11. Kohli, R., Tan, S.: Electronic health records: how can is researchers contribute to transforming healthcare? MIS Q. **40**(3), 553–574 (2016)
12. Hollingworth, W., Ebel, B.E., McCarty, C.A., Garrison, M.M., Christakis, D.A., Rivara, F. P.: Prevention of deaths from harmful drinking in the United States: the potential effects of tax increases and advertising bans on young drinkers. J. Stud. Alcohol **67**(2), 300–308 (2006)
13. Brennan, A., Meier, P., Purshouse, R., Rafia, R., Meng, Y., Hill-Macmanus, D.: Developing policy analytics for public health strategy and decisions-the Sheffield alcohol policy model framework. Ann. Oper. Res. **1**, 149 (2016)
14. Chun, H., Mobley, M.: Gender and grade-level comparisons in the structure of problem behaviors among adolescents. J. Adolesc. **33**, 197–207 (2010)
15. Syamlal, G., Mazurek, J., Dube, S.: Brief report: gender differences in smoking among U.S. working adults. Am. J. Prev. Med. **47**, 467–475 (2014)
16. Fagan, P., Augustson, E., Backinger, C.L., O'Connell, M.E., Vollinger Jr., R.E., Kaufman, A., Gibson, J.T.: Quit attempts and intention to quit cigarette smoking among young adults in the United States. Am. J. Public Health **97**(8), 1412–1420 (2007)
17. Soulakova, J., Li, J., Crockett, L.: Race/ethnicity and intention to quit cigarette smoking. Preventive Medicine Reports **5**(C), 160–165 (2017). C, p. 160, Directory of Open Access Journals
18. Pennanen, M., Broms, U., Korhonen, T., Haukkala, A., Partonen, T., Tuulio-Henriksson, A., Laatikainen, T., Patja, K., Kaprio, J.: Smoking, nicotine dependence and nicotine intake by socio-economic status and marital status. Addict. Behav. **39**(7), 1145–1151 (2014)
19. Cho, H., Khang, Y., Jun, H., Kawachi, I.: Marital status and smoking in Korea: the influence of gender and age. Soc. Sci. Med. **3**, 609 (2008)
20. Mills, A.L., White, M.M., Pierce, J.P., Messer, K.: Home smoking bans among US households with children and smokers: opportunities for intervention. Am. J. Prev. Med. **41**(6), 559–565 (2011)
21. Halterman, J.S., Conn, K.M., Hernandez, T., Tanski, S.E.: Parent knowledge, attitudes, and household practices regarding SHS exposure: a case-control study of urban children with and without asthma. Clin. Pediatr. **49**(8), 782–789 (2010)
22. Ruhm, C.J.: Are recessions good for your health? Q. J. Econ. **115**(2), 617–650 (2000)
23. Harris, K.M., Edlund, M.J.: Self-medication of mental health problems: new evidence from a national survey. Health Serv. Res. **40**(1), 117–134 (2005)
24. Henkel, D.: Unemployment and substance use: a review of the literature (1990–2010). Curr. Drug Abuse Rev. **4**(1), 4–27 (2011)
25. De Walque, D.: Does education affect smoking behaviors? J. Health Econ. **26**, 877–895 (2007)
26. Flood, S., King, M., Ruggles, S., Warren, J.R.: Integrated public use microdata series, current population survey: version 4.0 [dataset]. University of Minnesota, Minneapolis (2015). http://doi.org/10.18128/D030.V4.0

fridaya@uef.fi

27. Sullivan, L.W.: SMOKING AND HEALTH: A National Status Report 2nd Edition – A Report to Congress (PDF). nlm.nih.gov. U.S. Department of Health and Human Services (1986). https://profiles.nlm.nih.gov/ps/access/NNBBVP.pdf. Accessed 4 May 2017
28. Ball, A.: Review of data management lifecycle models. University of Bath (2012). https://purehost.bath.ac.uk/ws/portalfiles/portal/206543/redm1rep120110ab10.pdf. Accessed 21 Jan 2019
29. Long, J.S., Freese, J.: Regression Models for Categorical Dependent Variables Using Stata. Stata Press, College Station (2006)
30. Tranmer, M., Elliot, M.: Binary Logistic Regression. Cathie Marsh Centre for Census and Survey Research (2008). http://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2008/2008-20-binary-logistic-regression.pdf. Accessed 4 May 2017
31. Santucci, A.C.: Data Description in Research. Salem Press Encyclopedia of Health, Ipswich (2016)

fridaya@uef.fi

Designing a System for Integration of Macroeconomic and Statistical Data Based on Ontology

Olga N. Korableva[1], Olga V. Kalimullina[2(✉)], and Viktoriya N. Mityakova[3]

[1] Institute of Regional Economic Studies of Russian Academy of Science, St. Petersburg State University, St. Petersburg, Russian Federation
[2] The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, St. Petersburg, Russian Federation
chemireva@mail.ru
[3] St. Petersburg Electrotechnical University "LETI" (ETU), St. Petersburg, Russian Federation

Abstract. The article describes a process of designing software for the aggregation of data (macroeconomic and statistical indicators) from distributed heterogeneous sources and their analysis based on the previously developed ontology of innovation activity and economic potential. The software includes a data aggregation system (supporting the user's markup process of PDF, HTML, and XLS documents and texts for further automated collection), an ontology automatic replenishment system and a system of semantic search for data subsets according to certain criteria.

Keywords: Data collection · Algorithm · Methodology · Ontology

1 Introduction

Within the framework of the previous stage of the study, the methodology and algorithms were developed for collection of economic data and bringing them to the structure of the developed ontology of the innovative activity and economic potential [1].

This article describes the design of the software for macroeconomic and statistical data collection based on the previously described algorithms and methodology. This software makes it possible to create a knowledge base for building the models for prediction of target macroeconomic indicators, in particular, the potential indicators of the economic growth and innovation activity in the Russian Federation.

1.1 Analog Overview

The systems and approaches to the integration of heterogeneous data based on ontologies have been considered in the works of a number of authors.

In [2], the author describes an approach to the harmonization of semantic models used by financial organizations and their respective terms, as well as simplification of

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 157–165, 2019.
https://doi.org/10.1007/978-3-030-22868-2_12

fridaya@uef.fi

the process of integration of data from heterogeneous sources with the help of the information system using semantic Web technologies. The main component of this system is the global ontology of the financial sphere [3]. The global ontology interacts with the ontologies of data sources based on the contents of relational databases, documents in natural language, and spreadsheets.

In [4], a new approach to solving the problem of integration of multiple ontologies is considered to ensure the compatibility and representation of data and knowledge in intelligent information systems. To solve the problem of semantic heterogeneity of data and knowledge and provide access to heterogeneous information from different subject areas, a modified approach is proposed for building a resultant ontology for several initial ones at the level of matching concepts, relationships, and attributes [5].

The author in [6] contributes to a formal and semi-automatic approach to the development of ontologies based on the formal concept analysis (FCA) method with the goal of integration of the data, containing the implicit and ambiguous information. The FCA method improved within the work supports the development of ontologies by extracting the concepts from attribute descriptions of the objects, which allows automating some ontology development processes. A study of several non-trivial industrial datasets and the experimental results obtained show that the proposed method offers an effective mechanism to enable the enterprises to interrogate and verify heterogeneous data and create the knowledge that meets business needs.

The author in [7] presents a framework that facilitates the access of scientists to historical and cultural data on food production and the commercial trading system during the times of the Roman Empire distributed across various data sources. The proposed approach uses on the ontology-based data access paradigm (OBDA), where different data sets are actually integrated by the conceptual level (ontology), which provides the user with a single access point and an unambiguous conceptual view.

The author in [8] describes the process of evolutionary development of an ontology, developed for a data integration system called ONDINE (Ontology based Data Integration), which supports the processes of obtaining, commenting and querying the experimental data extracted from scientific documents. The main element of the ONDINE system is an ontology, which makes it possible to annotate and request the experimental data. The ontology may be transformed to be adapted to the changes in the domain, new ways of use and annotated data. The article presents the process of evolutionary development of the ontology, which takes the ontology as an input information in a consistent state, determines and applies some changes and manages the consequences of these changes, creating an ontology in a consistent state.

The author in [9] provides an overview of systems and approaches to integration of the heterogeneous data based on ontologies prior to 2001. Unlike this work, the databases and ontologies are considered as data sources.

The overwhelming majority of existing projects are focused on ontological modeling of the business processes. There are practically no works considering the ontological modeling of the economic growth potential and innovative activity indicators within the country. This confirms the relevance and scientific novelty of the research conducted by the authors.

fridaya@uef.fi

1.2 System Design

To support the process of collection of unstructured and weakly structured data from distributed heterogeneous sources, a prototype system for macroeconomic and statistical data collection (MSDS) is being developed. In [1], the methodology and algorithms for finding the sources and collection of economic data and their reduction to the structure created within the framework of the ontology of innovative activity and economic potential, as well as semantic technologies that are used in the construction of the MSDS, are described. This paper describes the MSDS design process based on them. The first step in designing a system is the elaboration of its use cases.

2 Results

2.1 The UML Diagram of the Use Cases for the MSDS

The UML diagram of the use cases is shown in Fig. 1. To identify a user and save the individual settings, the user must first register and then log in to the system. Further, the user with the help of the system solves the high-level tasks: data collection from various types of sources, viewing of collected and preprocessed data, or searching for the data subsets according to certain criteria.

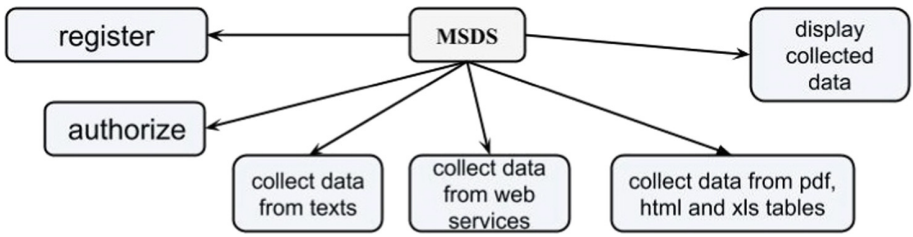


Fig. 1. The UML diagram of the use cases for the MSDS.

2.2 Features of Data Sources for the MSDS

Unfortunately, most of the sources found for this work provide data in the form of PDF tables. The PDF format is intended for uploading the information in a human-readable form and is not a machine-readable format. In addition to the problems with formatting, the tables contain data in an arbitrary structure. All this makes the task of automatic data collection from PDF/XLS documents difficult and time-consuming. In this regard, it was decided to introduce the stage of semantic annotation of the document by the operator for the tables, in which the cells are matched with the ontology entities.

2.3 General Architecture of the MSDS

The MSDS is designed as a client-server web application with which the user interacts through a web interface. The general MSDS architecture is shown in Fig. 2. The web

fridaya@uef.fi

application contains a system for collecting data from various types of sources (PDF, HTML, XLS, web services), part of which is the browser extension for semantic annotation of HTML documents by the user (tables and texts). With the help of this system, a web application receives potential RDF triples that undergo primary processing in an ontology automated replenishment system and are stored in a knowledge base (KB). The basis of the KB is the ontology of the innovation activity and economic potential. The data is also requested from it to be displayed in the web interface by the system of semantic search for data subsets.

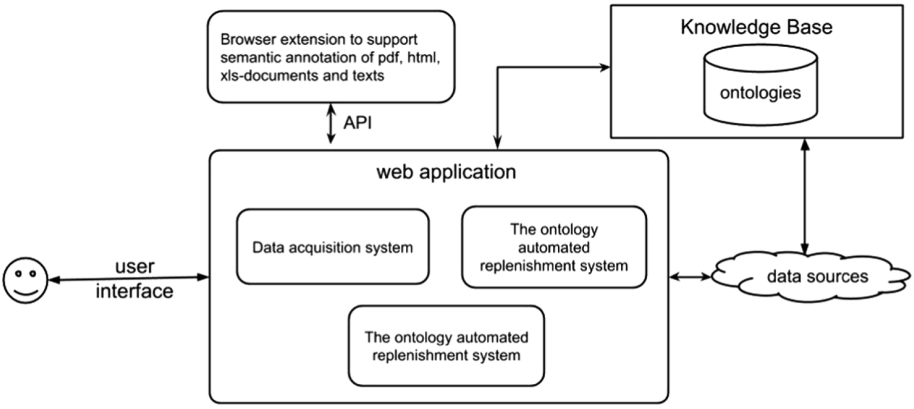


Fig. 2. General architecture of the MSDS.

2.4 The UML Diagram of Data Collection System Use Cases

The data collection system works in conjunction with a browser extension to support the semantic annotation of HTML documents. The user work is as follows. First, the user uploads into the system a PDF, HTML, or XLS document containing the table. The documents in the PDF and XLS formats are converted to HTML as the most convenient format for further processing.

Next, the user selects a cell in the table, which is a potential instance of the ontology class. Using the browser extension, the user selects the ontology class to which it belongs (if necessary, the user creates a class). The user selects the properties one by one (creates them if necessary) and adds their values from the table. If a conflict arises (there is already such an instance with a different value), then the user resolves it.

In order not to perform a similar action manually, for example, for the cells of the next line, the user can record the macros (a sequence of actions when annotating a part of the table) and thereby automate the process for the remaining part.

For example, the table shows the energy intensity of GDP for various regions of the Russian Federation. The user presses the macros record button and then proceeds to annotate the table. The user selects a cell in which the value of the energy intensity of GDP is indicated. Then, in the system interface, the user selects the ontology class “Energy intensity of GDP”, filling in the appropriate field in the form. Next, the user

fridaya@uef.fi

selects the value “Leningrad Region” specified in the table and in the interface selects the property “indicated for the region”, thus filling in the next form field. The similar actions are performed for other properties (for example, quarter, year, which are also properties of the class “Energy intensity of GDP” ontology). After annotating a single row of the table, the user can run the macros for the subsequent rows, so as not to repeat similar actions.

The UML diagram of use cases is shown in Fig. 3.

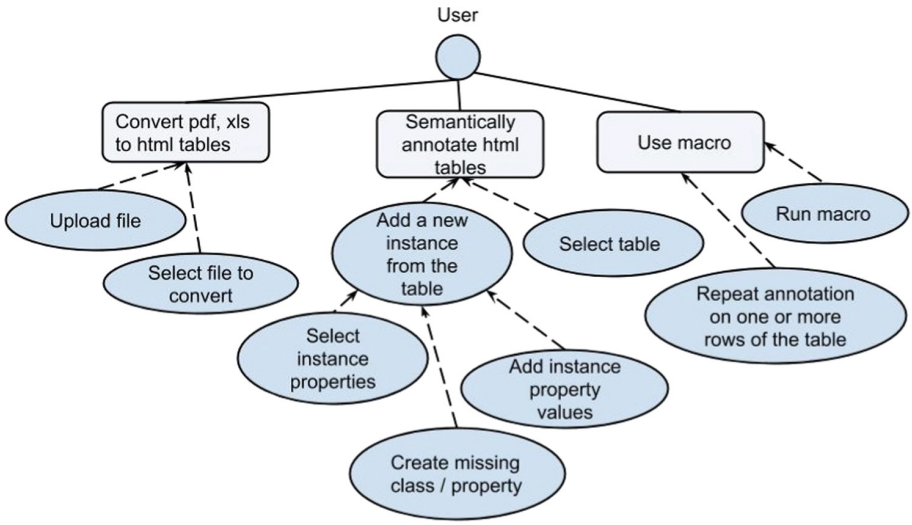


Fig. 3. The UML diagram of the use cases.

2.5 The UML Diagram of the Use Cases for the Ontology Automatic Replenishment System

Potential RDF triples that are preprocessed before being saved to the ontology (resolution of data conflicts, deletion of duplicate information) are fed to the input of the ontology automatic replenishment system. The UML diagram of the use cases is shown in Fig. 4. Data from web services is also collected using the ontology automatic replenishment system.

2.6 The UML Diagram of the Use Cases of the System of Semantic Search for the Data Subsets with Specified Properties

The user works with the system of semantic search for the data subsets with the specified properties as follows. The user can select the following display modes:

(1) Tabular (by analogy with the OLAP cube for relational databases). It is convenient to build around a certain class of ontology, then the dimensions are its properties (including connections with other classes), and the values of the properties, respectively, are in the cells.

fridaya@uef.fi

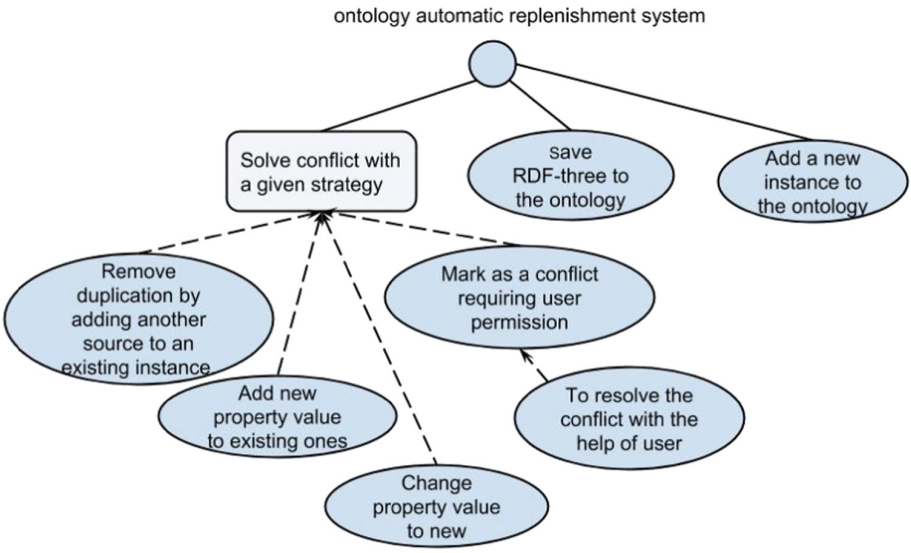


Fig. 4. The UML diagram of the ontology automatic replenishment use cases.

(2) Block-structured. The user can select a class from the list, after which all its entities will be displayed, as well as the links to other classes that can be accessed for viewing.
(3) Graph-structured.

An oriented graph is a convenient way to display classes, instances, and connections of an ontology, while the vertices are the ontology classes, the connections are the relations between the classes. The relations "is the subclass" and "contains an instance" make it possible to build a hierarchy of classes and instances. The user can open and hide the classes, subclasses, and instances for viewing, going down or going up the hierarchy.

The user can also sort the entities by the selected property value, filter the data by one or more selected property values (for example, year, industry), conduct a full-text search and save the specified filters.

The UML diagram of use cases is shown in Fig. 5.

2.7 The UML Diagram of the Components of the MSDS

The UML diagram of the components of the MSDS is shown in Fig. 6. Let us consider it in more detail. To work with ontology, the Jena library is used, which provides the Jena API and allows a user to store the ontologies in the database. The ontology infrastructure is common for all the MSDS subsystems and is implemented using a set of components and interfaces (see the TripleDBConnector component).

fridaya@uef.fi

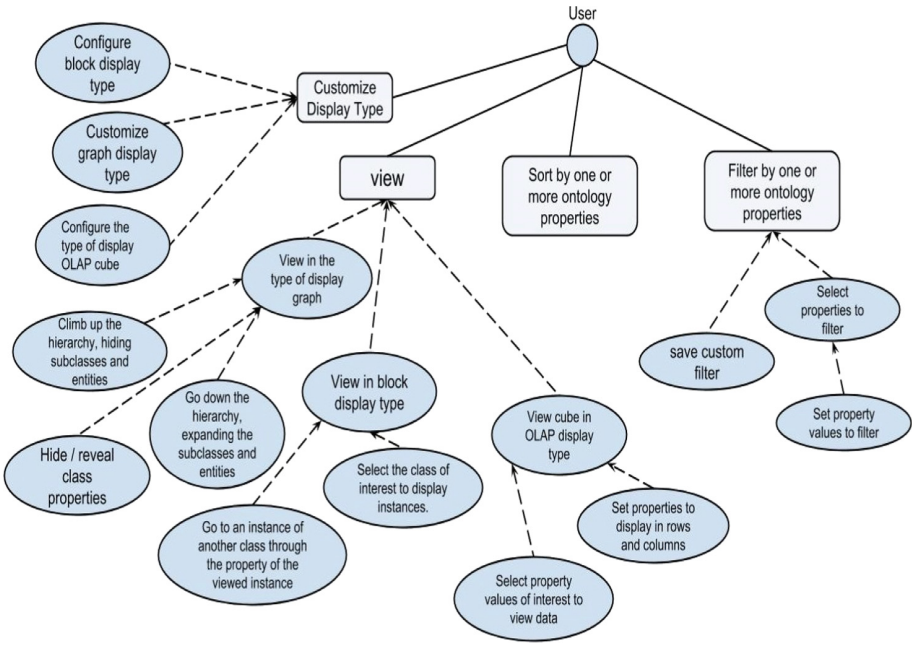


Fig. 5. The UML diagram of the use cases of the system of semantic search for data subsets with specified properties.

The data collection system allows the user to convert PDF documents into HTML (using the "easyConverter" library [10]) and provides the API to the browser extension for semantic annotation of the documents by the user. For this, the system interacts with the knowledge base to obtain a list of classes, properties, and instances. After annotation, potential RDF triples are obtained for addition to the ontology.

The ontology automated replenishment system is intended for primary data processing, i.e. reducing the potential triples (PotentialRDFTriples interface) to a form suitable for adding to the ontology (RDFTriples interface): duplicate triples are removed, data conflicts are resolved, etc.

The semantic data subset search system formulates the filters, sorting and display mode selected by the user in terms of a SPARQL query and sends them to the KB (TripleDBConnector component). The subsystem also provides the API for displaying the data on the client (UI API). The web interface is planned to be implemented using modern popular libraries ReactJs [11] and Bootstrap [12].

fridaya@uef.fi

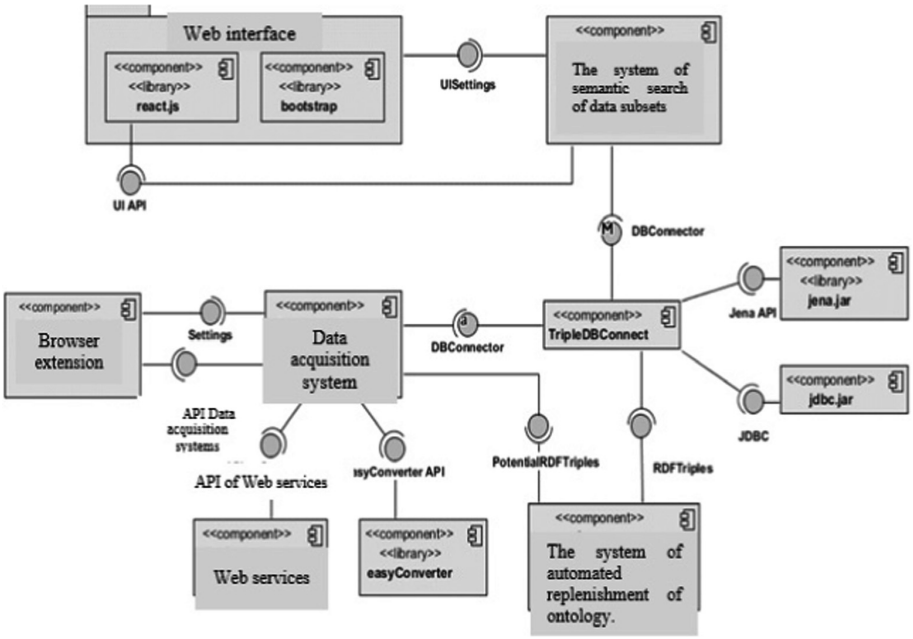


Fig. 6. The UML diagram of the MSDS components.

3 Conclusion

This paper presents the process of designing the software for semantic search, collection and analysis of macroeconomic and statistical data based on the ontology of innovation activity and economic potential, using the previously developed methodology and algorithms. The software includes a data collection system (supporting the user's markup process of PDF, HTML, XLS documents and texts for further automated data collection), the ontology automatic replenishment system and the system of semantic search for data subsets according to certain criteria.

The software for collection of the macroeconomic and statistical data based on ontology makes it possible to create a knowledge base for building the predictive models for target macroeconomic indicators, in particular, the indicators of the economic growth and innovation activity potential of the Russian Federation.

Acknowledgements. This research is supported by RFBR (grant 16-2912965\18).

References

1. Korableva, O.N., Kalimullina, O.V., Mityakova, V.N.: Innovation activity data processing and aggregation based on ontological modelling. Paper presented at the 2018 4th International Conference on Information Management, ICIM, pp. 1–4 (2018)

fridaya@uef.fi

2. Petrova, G.G., Tuzovskiy, A.F.: Financial organization information system based on semantic Web technologies. In: Collection of Works of the XIII All-Russian Scientific and Practical Conference of Students, Postgraduates and Young Scientists, pp. 87–89. Tomsk (2016)
3. Korableva, O.N., Razumova, I.A., Kalimullina, O.V.: Research of innovation cycles and the peculiarities associated with the innovations life cycle stages. Paper presented at the Proceedings of the 29th International Business Information Management Association Conference - Education Excellence and Innovation Management Through Vision 2020: From Regional Development Sustainability to Global Economic Growth, pp. 1853–1862 (2017)
4. Bova, V.V.: Ontological model of data and knowledge integration in intellectual information systems, Izvestiya SFU. Technical science, №. 4(165) (2015)
5. Wolter, U., Korableva, O., Solovyov, N.: The event bush method in the light of typed graphs illustrated by common sense reasoning. In: Dynamic Knowledge Representation in Scientific Domains, pp. 320–353 (2018). https://doi.org/10.4018/978-1-5225-5261-1.ch014
6. Gaihua, Fu.: FCA based ontology development for data integration. Inf. Process. Manage. **52**, 765–782 (2016)
7. Calvanesea, D., Liuzzod, P., Mosca, A.: Ontology-based data integration in EPNet: production and distribution of food during the Roman Empire. Eng. Appl. Artif. Intell. **51**, 212–229 (2016)
8. Ibanescu, L., Buche, P., Dervaux, S.: Ontology evolution for an experimental data integration system. Int. J. Metadata Semant. Ontol. **11**(4), 231–242 (2016)
9. Wache, H., Vögele, T., Visser, U.: Ontology-based integration of information - a survey of existing approaches. In: Workshop: Ontologies and Information, pp. 108–117 (2001)
10. http://www.pdfonline.com/convert-pdf-to-html/. Accessed 9 Oct 2018
11. https://reactjs.org/. Accessed 15 Oct 2018
12. https://react-bootstrap.github.io/. Accessed 3 Oct 2018

fridaya@uef.fi

A Modified Cultural Algorithm for Feature Selection of Biomedical Data

Oluwabunmi Oloruntoba and Georgina Cosma(✉)

School of Science and Technology, Nottingham Trent University, Nottingham, UK
oluwabunmi.oloruntoba2017@my.ntu.ac.uk, georgina.cosma@ntu.ac.uk

Abstract. An important step in developing predictive models is determining the best features to be used for building the models. Feature selection algorithms are frequently adopted to remove non-informative and redundant features from the dataset before building the predictive models. However, it can be a significant challenge to determine the features which would make the best predictors, particularly in large datasets. Importantly, building a model with the best subset of features can make it more interpretable and efficient. This paper proposes a novel feature selection algorithm which is based on the cultural meta-heuristic optimization algorithm. The modified Cultural Algorithm was developed and optimized for achieving high accuracy and Area Under the Curve. The quality of the selected features was assessed using the performance of a Support Vector Machine classifier. The proposed algorithm was tested on five benchmark datasets and achieved an average accuracy of 0.923 and an Area Under the Curve of 0.898 across all datasets. Although the performance of the proposed modified Cultural Algorithm was comparable with that of the standard Genetic Algorithm, the modified Cultural Algorithm required a smaller number of features and shorter execution time than the Genetic Algorithm.

Keywords: Cultural algorithm · Machine learning · Feature selection · Biomedical data

1 Introduction

Biomedical data has experienced exponential growth in recent years partly due to availability of smart and wearable devices and reduction in the cost of gene sequencing [12]. Biomedical scientists are faced with the challenge to manage and analyse massive amounts of data [8]. These datasets are often characterised by a small number of samples and a large number of features, many of which may be irrelevant and useless features [14]. Efficient processing of such high-dimensional biomedical data could potentially enhance health care services and increase research opportunities [5].

In order to achieve reliable inferences, it is important to identify and separate the important features from the irrelevant features, a process called feature

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 166–177, 2019.
https://doi.org/10.1007/978-3-030-22868-2_13

fridaya@uef.fi

selection. Feature selection has been increasingly used to reduce noise, computational time and resources, and increase the accuracy of classification models of biomedical data [2]. A reduced feature set leads to an efficient learning model that can be easily interpreted.

Feature selection methods can be classified as filter, wrapper and ensemble. The selection of features by filter-based models is independent of classification algorithms. The models often use statistical tests to determine the features that are more predictive of the output. Wrapper-based feature selection models work together with the classification model. This model selects a subset of features and determines its suitability based on the predictive performance of a classifier. The performance of the classifier is then used to select the best subset of features which give the optimal predictive performance. Ensemble feature selection methods select an optimal subset of features by combining multiple subsets of features using various classifiers [14].

To determine the best subset of features that can give the optimal classification or predictive performance, an exhaustive search of the feature space is required – which can be a complex process for biomedical datasets because they often have a large number of features [10]. A possible solution to the complex search process is to utilize the exploratory capabilities of meta-heuristics. Meta-heuristics are iterative generation search algorithms that are problem independent and can generate optimal solutions by iteratively improving the quality of candidate feature sets based on some evaluation criteria [17]. Meta-heuristic based hybrid feature selection algorithms including, Genetic Algorithm [17], Gravitational Search Algorithm [1], Cultural Algorithm (CA) can be used to tune parameters of classifiers as well as to select the best feature subset for optimal classification performance [13].

However, most meta-heuristic-based models, except CA, suffer from time complexity. CA can improve the time complexity of the search process by using knowledge of previous results to guide the search [17]. CA has been applied to solve engineering design problems [17], and schedule tasks in cloud scheduling operations [1]. CA algorithms have been shown to significantly improve the convergence rate of Genetic Algorithms (GA) and Particle Swarm Optimization [3].

Basic versions of CA have been known to suffer from low accuracy [16]. To solve these, several variations to Cultural Algorithms have been developed. Yan et al. [18] developed a cultural differential evolution algorithm with multiple populations which involve the use of segregation of the population based on common characteristics prior to the application of genetic operators. This version of CA was applied to image matching and compared with a Genetic Algorithm, and the traditional CA. Results showed that the segregation operator in the modified CA helped increase the precision and stability of the algorithm achieving accuracy rate of 98% with average execution time of 9.53 s. A multi-population Heritage-Dynamic Cultural Algorithm (HDCA) was developed by Hlynka and Kobti [4]. HDCA incorporates ancestral heritage dynamics thus allowing for selection of parent solutions from different populations to produce offspring. When applied to functions from the CEC'14 Test Suite single objective real-parameter numer-

fridaya@uef.fi

ical optimization test function [7], HDCA showed it has greater search potential than the traditional CA due to its support for migration between populations.

To determine an optimal vibration control strategy for a vehicle's active suspension system [15], Fuzzy-PID controller was designed by combining CA with niche algorithm. A niche algorithm attempts to converge to multiple solutions within a single run. The algorithm segments the population into multiple disjointed sets such that there is at least one candidate solution in the area within each local optima [15]. There is limited information on the suitability and application of Cultural Algorithms to select optimal features in biomedical datasets.

This paper proposes a modified Cultural Algorithm for feature selection which is experimentally evaluated using a number of biomedical datasets. The paper is structured as follows. Section 2 presents the architecture of the conventional Cultural Algorithm. Section 3 explains the techniques of the proposed Cultural Algorithm. Section 4 describes the datasets and experiment methodology. The results obtained are presented and discussed in Sect. 5, and Sect. 6 provides a conclusion and future work.

2 Cultural Algorithm

Cultural Algorithm (CA) is a computation model inspired by the process of cultural evolution in human society [1]. It defines acceptable means of solving problems within a society and defines the rules of interaction between the members of the population. Information that is common to the population are stored in a common area and it is used to direct the process of evolution within the society. The knowledge and characteristics shared by the population are collectively referred to culture. This knowledge is centrally stored and used to guide the behaviors and actions of the society. The CA consists of three concepts: (1) a population of individuals, (2) a collection of knowledge called the Belief Space, and (3) a rule of communication which defines the control of knowledge between the two spaces. The population space is often modelled after any population-based evolutionary algorithm, including the Genetic Algorithm and Evolutionary Algorithm [11]. The belief space consists of any number of combination knowledge types including situational knowledge, normative knowledge, topographical knowledge, domain knowledge, and historical knowledge. Situational knowledge contains information about the best solution in the population over the course of generations. The range of possible values a solution can take are stored in the normative knowledge. The topographical knowledge contains information about the topography of the solution space while historical information about the search process is stored in the historical knowledge. Information about the problem domain is stored in the domain knowledge. Knowledge within the cultural society can be expressed as any combination of the five knowledge sources [1].

Individuals are assessed using a performance function. The fitness of each individual represents its problem-solving experience. A proportion of individuals are selected to impact the belief space, and this is controlled by an acceptance

fridaya@uef.fi

function. The update function determines how the selected knowledge will be included in the belief space. The knowledge in the belief space is then used to determine the individuals to be selected for the next generation through an influence function. Thus, at each generation, feedback from the population is used to update the belief space and vice versa. This is a form of dual inheritance system which has been shown to improve the rate of evolution [1].

3 Proposed Algorithm

A new Cultural Algorithm (CA) is proposed for the task of feature selection, as illustrated in Algorithm 1. The proposed CA starts with a population of randomly generated individual solutions represented as binary strings of fixed length equivalent to the number of features in the dataset. The binary value of 1 or 0 indicates that the feature is to be selected or ignored respectively. Two knowledge spaces are used including situational and normative knowledge. CA was optimized for accuracy and the Area Under the Receiver Operating Characteristic curve (AUC). The fitness of an individual is either the accuracy (where optimization is performed for accuracy) or the AUC (where optimization is performed for AUC). The acceptance function considers the solutions to be added to the belief space as the best 30% of the population. The update function updates the situational knowledge with the maximum of the best solution in the population and the current solution in the situational knowledge. The normative knowledge is also updated with the current worst and best solutions. The influence function applies the change operator to individuals with fitness less than the solution in the situational knowledge.

Random selection is performed to select a percentage of the total genes. A change operator, applied to the selected genes, changes the values to 0. The application of the change operator results in a new population [15]. The objective is to determine the least number of features that would give the best predictive performance. The feature selection algorithm attempts to find the set of features with the best classification performance. Thus, the objective function is defined as maximum accuracy or maximum AUC where optimization is done for accuracy or AUC respectively.

A Support Vector Machine (SVM) classifier with a linear kernel function was adopted to assess the performance of the selected subset of features. SVM is a binary classifier that finds the separation boundary with the greatest possible margin between positive and negative class instances. The datasets tested in this experiment are provided in Table 1.

4 Datasets and Experiment Methodology

The section describes the biomedical datasets and the experiment methodology adopted for evaluating the proposed modified CA for feature selection and to compare its performance to the GA for feature selection.

fridaya@uef.fi

Algorithm 1. Pseudocode for proposed CA

```
begin
generation = 0;
initialise Population
initialise BeliefSpace
loop
assess every individual
if fitness(individual) < best(BeliefSpace) then
    apply change operator
select best(Population)
update(BeliefSpace)
generation = generation + 1
until termination condition is reached
end
```

4.1 Datasets

The datasets adopted for the experiments are benchmark datasets which are commonly used in evaluating algorithms and machine learning models. Table 1 shows the characteristics of the datasets used in this experiment. All the datasets were standardized using z-score transformation to minimize variations within the data. Z-score converted all features to a common scale with an average of zero and standard deviation of one. Table 2 shows the change operator value for each dataset.

Table 1. Dataset characteristics and class distribution for all datasets.

	Class	Breast	Heart	Hepatitis	Leukaemia	Ovarian
Number of features		30	13	19	72	100
Sample size		569	270	155	7129	216
% of samples	1	62.74	55.56	20.65	65.28	56.02
	2	37.26	44.44	79.35	34.72	43.98

4.2 Evaluation Measures

Accuracy, AUC, True Positive Rate (TPR) and False Positive Rate (FPR) were the evaluation measures used in the comparisons of the prediction models. Accuracy is defined as the ratio of all correctly identified positive and negative instances/classes. The TPR, also called Recall or Sensitivity, is a measure of the fraction of actual positives that are correctly identified. FPR is the measure of the fraction of negatives incorrectly identified as positives.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$TPR = \frac{TP}{TP + FN} \tag{2}$$

fridaya@uef.fi

$$FPR = \frac{FP}{FP + TN} \tag{3}$$

TP is the total number of True Positive predictions, TN is the total number of True Negative predictions, FP is the total number of False Positive predictions, and FN is the total number of False Negative predictions.

AUC is a performance measure that calculates the area under the Receiver operating characteristic (ROC) curve, and it is a measure of how well predictions are ranked and not their absolute values. The AUC provides an aggregate value across all possible classification thresholds. The AUC calculates the probability that a randomly chosen positive sample will be ranked higher than a randomly chosen negative one by a classifier.

4.3 Experiment Methodology

Every predictive model (i.e. classification model) was evaluated using the k-fold cross-validation method where the value of k was set to 10 (k = 10), to minimize test errors due to bias or variance [16]. With k-fold cross validation, a dataset is randomly partitioned into k equal size subsamples. Out of the k subsamples, k-1 subsamples are used as training data, and a single subsample is used as the validation data for testing the model. The cross-validation process is then repeated k times (where k is equal to the number of folds), with each of the k subsamples used once as the validation data. The results from each k fold are averaged to produce a single performance estimation. Hence, this method ensures the performance of the classifier is not skewed towards the majority class.

Since all the datasets were imbalanced (i.e. the number of samples in each class varied) random under sampling was applied to reduce the imbalance in the training dataset [9]. Using random under sampling involved elimination of random samples from the majority class to ensure that there is an equal distribution of the target classes in the training data.

The modified Cultural Algorithm (CA) and a Genetic Algorithm (GA) were developed and optimized for achieving high accuracy and AUC. The quality of the selected features was assessed using the performance of a Support Vector Machine classifier (SVM). Experiments were carried out using the Cultural Algorithm with SVM classification Accuracy (SVM(Acc)) as an objective function (CA+SVM(Acc)); Cultural Algorithm with SVM(AUC) as an objective function (CA+SVM(AUC)); and the Genetic Algorithm with SVM classification Accuracy (SVM(Acc)) as an objective function (GA+SVM).

The experiments were carried out in MATLAB 2017a running on Windows 10 operating system on a Lenovo desktop with core i5-7400 3 GHz and 8 GB of RAM. For each model evaluated, population size is set to 30. The average performance of each model was evaluated over 10 runs, consisting of 30 generations per run.

fridaya@uef.fi

Table 2. Change operator value for each dataset.

Dataset	Change operator value
Breast	0.30
Heart	0.40
Hepatitis	0.03
Leukaemia	0.03
Ovarian	0.03

5 Results and Discussion

The models were optimised for Accuracy (CA+SVM(Acc)) and AUC (CA+SVM(AUC)) and were compared to a GA that used an SVM classifier (GA+SVM) proposed by Xu et al. [16]. The mean values for the Accuracy, AUC, TPR and FPR are shown in Table 3. This section discusses the results of the CA and the GA taking into consideration the SVM's overall performance across the datasets; the number of features the feature selection algorithms needed for achieving their highest performances; and the execution times of the algorithms.

Comparison Based on Overall Performance Across the Datasets: For a model to be robust, it should have a very high accuracy and AUC value. Also, the ratio of correctly identified positive and negative cases should be high. This is of particular importance while building predictive models for biomedical datasets. One desirable trait is that the model must be able to correctly identify both positive and negative cases so as to give correct predictions as well as avoid wrong diagnosis. Thus, a robust model would have TPR = 1 and FPR = 0 in addition to high accuracy and AUC values. Table 3 shows the model names where GA+SVM denotes the GA with SVM Accuracy as the objective function to be maximised. CA+SVM(Acc) denotes the Cultural Algorithm with SVM Accuracy as the objective function to be maximised; and CA+SVM(AUC) denotes the Cultural Algorithm with SVM Area Under the Curve as the objective function to be maximised.

GA+SVM returned the highest predictive accuracy over all datasets, slightly outperforming CA+SVM(Acc) by 0.028 points, and CA+SVM(AUC) by 0.036 points. The results obtained by the GA+SVM and CA+SVM models were comparable. Comparing CA+SVM(Acc) and CA+SVM(AUC) the results are very close, but CA+SVM(Acc) returned on average a slightly higher Accuracy and TPR, and a lower FPR than CA+SVM(AUC).

Figures 1 and 2 show the boxplots for the accuracy values and AUC values across all datasets for the models GA+SVM, CA+SVM(Acc) and CA+SVM(AUC) across the 10 runs. The boxplots show that GA+SVM achieved the best performances. It is also notable that the CA+SVM(AUC) model did not perform as good as the other models on the Hepatitis and Leukaemia datasets.

fridaya@uef.fi

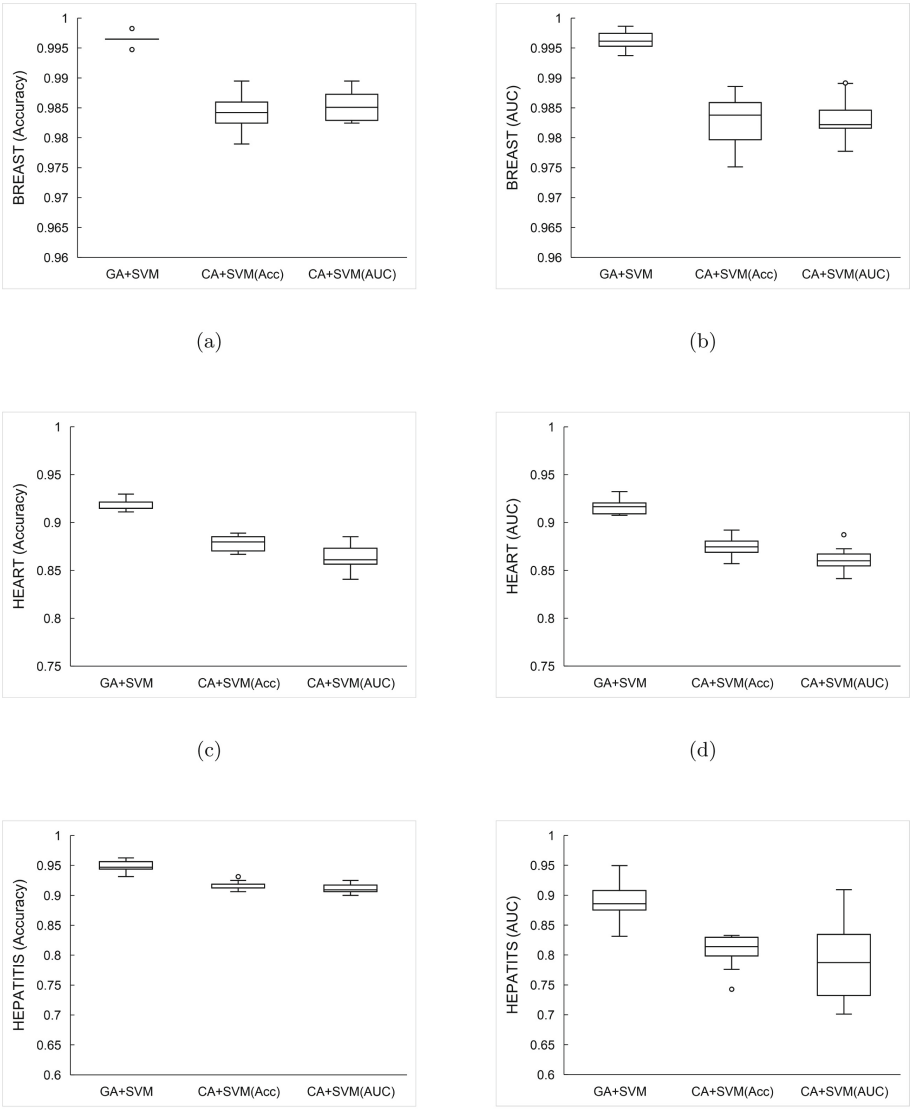


Fig. 1. Accuracy and AUC for (a and b) Breast (c and d) Heart (e and f) Hepatitis datasets.

A previous study has shown that the performance of the evolutionary process can be improved by concentrating the strong solutions through a series of elimination of the weaker solutions [6]. Thus, the slightly lower performance values of the CA+SVM compared to the GA+SVM might have been attributed to higher levels of weaker solutions within its population.

fridaya@uef.fi

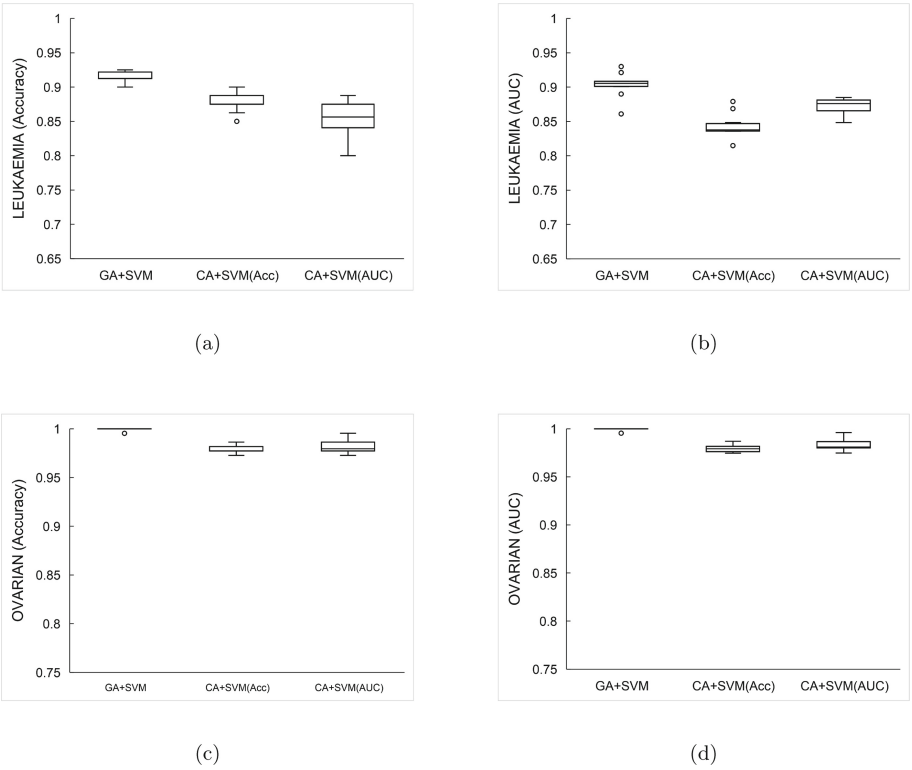


Fig. 2. Accuracy and AUC of the (a and b) Leukaemia (c and d) Ovarian datasets

Comparison Based on the Number of Features Selected: The optimal number of features relative to the total feature size selected by the models are shown in Fig. 3. The best model is the one which used the least number of features to achieve maximum predictive performance. A uniform basis of comparison is required to compare the selected features for the models across all datasets. Therefore, all selected features were converted into corresponding percentage values relative to their total feature size. Across all the datasets, the GA+SVM model selected the highest number of features using at least 50% of the feature size to achieve optimal predictive performance. For the Breast dataset, the percentage of features required to achieve the best predictive performance across all models were above 50%. CA+SVM models optimized for accuracy and AUC identified fewer number of features required to achieve optimal classification performance than GA+SVM model for Leukaemia and Ovarian datasets. In contrast, only CA+SVM(Acc) selected the least number features for both Heart and Hepatitis datasets.

Comparison Based on Execution Time: The average execution time (excluding the initialization time) for each model across 10 runs are shown in

fridaya@uef.fi

Table 3. Mean results of each approach applied to the datasets

	Method	Accuracy	AUC	TPR	FPR
Breast	GA+SVM	0.996	0.996	0.994	0.001
Heart	GA+SVM	0.918	0.916	0.900	0.067
Hepatitis	GA+SVM	0.948	0.891	0.979	0.205
Leukaemia	GA+SVM	0.914	0.903	0.822	0.043
Ovarian	GA+SVM	1.000	1.000	1.000	0.001
Average mean	GA+SVM	0.955	0.941	0.939	0.064
Average std.	GA+SVM	0.041	0.053	0.077	0.084
Breast	CA+SVM(Acc)	0.984	0.983	0.977	0.011
Heart	CA+SVM(Acc)	0.878	0.875	0.851	0.101
Hepatitis	CA+SVM(Acc)	0.917	0.806	0.98	0.377
Leukaemia	CA+SVM(Acc)	0.879	0.842	0.661	0.041
Ovarian	CA+SVM(Acc)	0.979	0.98	0.985	0.026
Average mean	CA+SVM(Acc)	0.927	0.897	0.891	0.111
Average std.	CA+SVM(Acc)	0.052	0.081	0.14	0.153
Breast	CA+SVM(AUC)	0.985	0.983	0.974	0.008
Heart	CA+SVM(AUC)	0.864	0.862	0.837	0.113
Hepatitis	CA+SVM(AUC)	0.913	0.794	0.976	0.387
Leukaemia	CA+SVM(AUC)	0.854	0.873	0.624	0.051
Ovarian	CA+SVM(AUC)	0.982	0.983	0.989	0.023
Average mean	CA+SVM(AUC)	0.919	0.899	0.88	0.116
Average std.	CA+SVM(AUC)	0.063	0.083	0.156	0.157

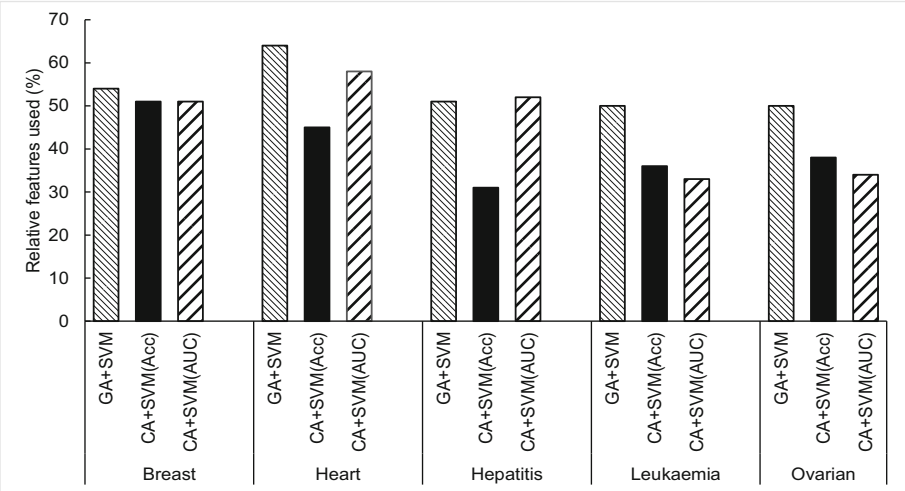


Fig. 3. Percentage (%) of features selected per dataset.

fridaya@uef.fi

Table 4. Mean execution time across datasets in seconds(s)

Datasets	GA+SVM	CA+SVM(Acc)	CA+SVM(AUC)
Breast	650	51	56
Heart	566	93	30
Hepatitis	501	85	63
Leukaemia	928	126	133
Ovarian	741	86	87
Average	677.2	88.2	73.8
Std.	166.675	26.696	38.829

Table 4. GA+SVM model had the highest execution time across all datasets. In general, the average execution time of the CA+SVM models were over 80% faster in achieving optimal performance compared to the GA+SVM model. In addition the CA+SVM models had, on average, a much lower execution time than GA+SVM. The standard deviation of the execution times for the CA+SVM algorithms is much lower than that of GA+SVM, and hence the execution times are less distributed and closer to the average.

6 Conclusion

This paper explores the suitability of a modified Cultural Algorithm for feature selection. The model was compared with a Genetic Algorithm and tested on various biomedical datasets. The quality of features selected was assessed using support vector machine with a linear kernel. In future, the performance of other kernel functions could be investigated.

Predictive performance values show that the modified Cultural Algorithm is comparable to the Genetic Algorithm. The Cultural Algorithm required much less execution time than the Genetic Algorithm. The number of features selected by the proposed Cultural Algorithm was smaller compared to Genetic Algorithm in most datasets.

The modified Cultural Algorithm proposed in this study is promising as a feature selection model for biomedical datasets. A form of biomedical data which was not considered in this study is image-based data which is often generated as outputs of common medical procedures like medical scans and x-rays.

Future work will focus on investigating the performance of this model on such data. Also, the performance of the modified Cultural Algorithm with other datasets and classification algorithms will be explored.

References

1. Azad, P., Nima, J.N.: An energy-aware task scheduling in the cloud computing using a hybrid cultural and ant colony optimization algorithm. Int. J. Cloud Appl. Comput. (IJCAC) **7**(4), 20–40 (2017)

fridaya@uef.fi

2. Dorigo, M., Gambardella, L.M.: Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans. Evol. Comput. **1**(1), 53–66 (1997)
3. Gogna, A., Tayal, A.: Metaheuristics: review and application. J. Exp. Theor. Artif. Intell. **25**(4), 503–526 (2013)
4. Hlynka, A.W., Kobti, Z.: Heritage-dynamic cultural algorithm for multi-population solutions. In: 2016 IEEE Congress on Evolutionary Computation (CEC), pp. 4398–4404, July 2016
5. Kim, J., Groeneveld, P.W.: Big data, health informatics, and the future of cardiovascular medicine. J. Am. Coll. Cardiol. **69**(7), 899–902 (2017)
6. Li, A., Xi, H., Liu, Q., Dong, L.: The operator of genetic algorithms to improve its properties. **4**(3)
7. Liang, J.J., Qu, B.Y., Suganthan, P.N.: Problem definitions and evaluation criteria for the CEC 2014 special session and competition on single objective real-parameter numerical optimization. Technical report (2013)
8. Margolis, R., Derr, L., Dunn, M., Huerta, M., Larkin, J., Sheehan, J., Guyer, M., Green, E.D.: The national institutes of health's big data to knowledge (BD2K) initiative: capitalizing on biomedical big data. J. Am. Med. Inform. Assoc. **21**(6), 957–958 (2014)
9. Mazurowski, M.A., Habas, P.A., Zurada, J.M., Lo, J.Y., Baker, J.A., Tourassi, G.D.: Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance. Neural Netw. **21**(2–3), 427–436 (2008)
10. Min, S., Lee, B., Yoon, S.: Deep learning in bioinformatics. Brief. Bioinform. **18**(5), 851–869 (2017)
11. Reynolds, R.G., Ali, M., Jayyousi, T.: Mining the social fabric of archaic urban centers with cultural algorithms. Computer **41**(1), 64–72 (2008)
12. Scruggs, S.B., Watson, K., Su, A.I., Hermjakob, H., Yates, J.R., Lindsey, M.L., Ping, P.: Harnessing the heart of big data. Circ. Res. **116**(7), 1115–1119 (2015)
13. Sebald, A.V., Fogel, L.J.: Evolutionary programming. In: Evolutionary Programming, pp. 1–386. World Scientific, September 1994
14. Ling, W., Haoqi, N., Ruixin, Y., Vijay, P., Michael, B.F., Panos, M.P.: Feature selection based on meta-heuristics for biomedicine. Optim. Methods Softw. **29**(4), 703–719 (2014)
15. Wang, W.I., Song, Y., Xue, Y., Jin, H., Hou, J., Zhao, M.: An optimal vibration control strategy for a vehicle's active suspension based on improved cultural algorithm. Appl. Soft Comput. **28**, 167–174 (2015)
16. Xu, W., Zhang, L., Gu, X.: A Novel cultural algorithm and its application to the constrained optimization in ammonia synthesis. In: Li, K., Li, X., Ma, S., Irwin, G.W. (eds.) Life System Modeling and Intelligent Computing, pp. 52–58. Springer, Berlin (2010)
17. Yan, X., Li, W., Chen, W., Luo, W., Zhang, C., Liu, H.: Cultural algorithm for engineering design problems (2012)
18. Yan, X., Song, T., Wu, Q.: An improved cultural algorithm and its application in image matching. Multimed. Tools Appl. **76**(13), 14951–14968 (2017)

fridaya@uef.fi

Exploring Scientists' Research Behaviors Based on LDA

Benji Li[1], Weiwei Gu[1,2], Yahan Lu[3], Chensheng Wu[4], and Qinghua Chen[1(✉)]

[1] School of Systems Science, Beijing Normal University, Beijing 100875, People's Republic of China
qinghuachen@bnu.edu.cn
[2] School of Informatics, Computing, and Engineering, Center for Complex Networks and Systems Research, Indiana University, Bloomington 47408, USA
[3] HSBC Business School, Peking University, Shenzhen 518055, People's Republic of China
[4] Beijing Institute of Science and Technology Information, Beijing 100044, People's Republic of China

Abstract. The progress of complex networks provides new perspective and methods for investigating on the subject "science of science", and relative literature has made a lot of remarkable discoveries. But this research framework is still insufficient with only network topology information of scientific literature and lack of considering true relationships between papers. Thanks to information technology, the distance between two contents can be evaluated by mapping them into a high-dimensional space based on the theme of each paper, which is automatically created by analyzing the abstract of paper via Latent Dirichlet Allocation (LDA). This paper analyzed some data from APS dataset with measures of distances between article pairs. The analysis of the pattern of scholars carrying out their research shows that, most scholars have concentrated rather than extensive research areas. Especially scientists' new researches are highly relevant to previous studies. Besides, this paper shows the scholar's citation behavior is not random. The citation behavior is basically concentrated in a limited interval, and the reference preference reaching a maximum at a certain distance. It implies there is a "most proper citation distance". Furthermore, they have a strong preference to cite the academic masters' articles.

Keywords: LDA · Science of science · Paper distance · Scientists' development pattern

1 Introduction

"Science of science" is a subject devoted to the study of characters and development of the entities and their relationships during the development of science. Very early discussions on the characteristics of scientific development primarily involved the fields of philosophy and naturology [1]. With the emergence and development of some objective and quantitative means or skills, great progress has been made. In this history, the main features are some small-scale statistical studies. For example, the evolution of

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 178–189, 2019.
https://doi.org/10.1007/978-3-030-22868-2_14

fridaya@uef.fi

popular research keywords, changes in hot research fields and papers' co-citation problems [2]. At the same time, scientometrics had been proposed and got some attention [3]. However, the research field had not received enough attention in academic circles, because there was no relatively complete theory and research methodology.

In the past twenty years, the rapid development of complex networks has provided new perspectives and methods for the research on science of science. This study based on network analysis mainly focuses on the scholars' cooperation network and science citation network [4–6]. The scholars' cooperation network is based on relationships among scholars. The nodes are the scholars and the edges represent cooperation among them. This network structure reflects the relationships among scholars. The science citation network is based on articles, the nodes are the articles and the edges are citations among them. It reveals the spread and development of knowledge as well as the relationships among fields of knowledge. This study based on complex networks can be divided into two parts. The first part is about the study of network properties while the second part is about the study of science. The first study focuses on the cooperation network and the science citation network's function, structure and evolution. It aims to discover community structure [7], rich club effect [8], fractal structure [9], triangle structure [10], and network development [11–13]. The second part mainly focuses on evaluation and prediction. It aims to evaluate academic articles, scholars, journals, schools and research institutions. For example, Duan proposed a weighted WPageRank algorithm for different reference times [14]. Considering the number of published articles and their cited times, Hirsch proposed H index [15]. Radicchi measured the influence of a scholar based on the cited diffusion process of the network [16]. Zhou established a bipartite network. This bipartite network contains both academic collaboration and citation information between articles. By discussion of this diffusion process, scholars and articles can be evaluated simultaneously [17]. Gualdi evaluated the quality of the paper through random walk processes on science citation networks [18]. Other studies focused on the allocation of credit between co-authors [19] and the selection of representation articles [20].

Although the study of science based on network analysis has gained abundant results, there are still several unresolved issues. These studies rely too much on cooperation and citation relationships. Each cited article contributes differently to the paper and there may exist "false references". Besides this, the citation relation is directed toward previous articles and can't cite following articles. This may lead to an incomplete structure and make the analysis of the problem one-sided. In addition, an advantage of complex network method is to grasp the macro structural information of the scientific literature but not deeply with each article's content. A large amount of literature was simplified into homogeneous nodes without information. The result was that the complex network method to application is still lacking in recommendations [21] and forecast [22]. Some research works have addressed these problems, and they actively promote the method of content analysis as a supplement. This enhanced the results of classification and clustering and opened up a new direction of inquiry.

Thus, here we introduce a new academic approach which is a new perspective to solve those problems. In this approach we discuss the transfer or development of scholars' research fields and the selection of reference materials. This approach is

fridaya@uef.fi

created through LDA which is a widely used natural language processing technology. This paper is organized as follows: In Sect. 2, the brief introduction to LDA and the data used in this paper is introduced. In Sect. 3, we conduct the analysis about the characteristics of the scholars' research fields and the selection of citations based on the distance between articles. In last Section, the conclusions and discussions are exhibited.

2 Methods and Materials

2.1 LDA

The vectorization of texts to facilitate subsequent computation is an important research issue of Natural Language Processing (NLP). Various methods have already been put forward, including TF-IDF, VSM, LSA, PLSA, LDA, and Doc2Vec, etc. Among them, LDA [23, 24] is a typical algorithm and has been successfully used to extract the hidden thematic structure in large documents. The intuition behind this model is that documents represent random mixtures of latent topics, where each topic is characterized by a distribution of words. In this model, each document exhibits multiple topics and each word of a document supports a certain topic.

$$P(word \mid document) = \sum [P(word \mid topic) \cdot P(topic \mid document)]. \quad (1)$$

LDA is a generative model. The intuition behind this model is that documents are represented as random mixtures of latent topics, where each topic is characterized by a distribution of words. Let α and η be the prior parameters for the Dirichlet document-topic distribution and topic-word distribution respectively. Assume there are K topics and β is a matrix with dimension $K*V$ where V is the number of words in the vocabulary (all the words in the corpus D). Each β_k is a distribution of the vocabulary. The topic assignment for the d-th document are z_d, where $z_{d,n}$ is the topic assignment for the n-th word in the d-th document. Finally, the observed words for document d are w_d, where $w_{d,n}$ is the n-th word in document d, which is an element from the fixed vocabulary. Using the above notations as presented in Fig. 1, LDA process can be described as follows:

(1) For each topic k, draw $\beta_k \sim \mathrm{Dir}(\eta)$.
(2) Give the d-th document d in corpus D, draw $\theta_d \sim \mathrm{Dir}(\alpha)$.
(3) For the n-th word in the d-th document $w_{d,n}$.
   (a) Draw $z_{d,n} \sim \mathrm{Multi}(\theta_d)$;
   (b) Draw $w_{d,n} \sim \mathrm{Multi}\left(\beta_{z_{d,n}}\right)$.

Here, Dir (.) is the Dirichlet distribution and Multi (.) is the multinomial distribution.

fridaya@uef.fi

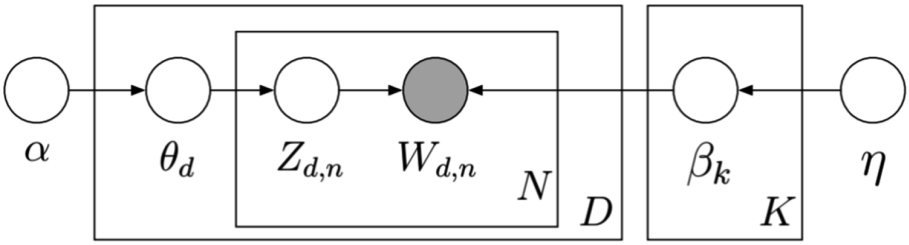


Fig. 1. Graphic structure of LDA process. This exhibits the adjusting process of parameters in LDA method.

2.2 Data Source and Processing

Data Source. The data is provided by the American Physical Society (APS), includes many physics journals, such as PHYSICAL REVIEW LETTERS, PHYSICAL REVIEW X, PHYSICAL REVIEW (A-E). We used 462,999 articles from year 1893 to 2009. The data format is listed in Table 1.

Table 1. An example of the data format.

Field name	Example
Title	Scaling behavior and surface-plasmon resonances in a model…
Doi	https://doi.org/10.1103/physrevb.48.6658
Author	X. Zhang; D. Stroud
Printed_time	1993/9/1
Received_time	1993/5/24
Citing_doi	https://doi.org/10.1103/physrevlett.46.375; https://doi.org/10.1103/physrevb.15.5733;…
Pacscode	82.70.-y;42.40.Jv;77.90.+k
Abstract	We calculate the ac dielectric function of a model Drude…

Processing. The abstract is a summary of the author's research content, so we can approximately represent each article's content with its abstract. Before we start to apply the topic vector representation for each article with LDA, we need to clean the data first. The cleaning process is as follows:

Convert Case. We translated all the words in the abstracts into lowercase. For example, "Dog" and "dog" indicate the same word "dog", if we treat them as different characters, it is not proper and may cause errors.

Remove Stop Words and Punctuations. The stop words such as "the", "an", "of", "a", "and" and punctuations have little meaning on the distribution of topic. So we got rid of those meaningless words.

Normalized Text. In order to improve the efficiency of text processing, we changed the plural form of each word into the singular form. We also turned all words' tense to the

fridaya@uef.fi

present. After data cleaning, the revised abstract usually contains only nouns, verb phrases, and other words which are essential to the articles' topic or content. Since LDA is a "bag-of-words" model, it does not take the order of words into account, so the data cleaning process doesn't affect the LDA model.

Embed Abstracts into High-Dimensional Space. Next we build a document matrix where each row represents an article's abstract. Then we fed the documents matrix to LDA model to acquire topic vector for each article. After training process, we got topic vector for each article that is high-dimensional representation.

3 Analysis and Results

Through LDA model, we got topic vector θ_i with 50 dimensions for i-th article, with each dimension indicates the probability under this topic. The sum of vector θ_i is equal to 1 or $\Sigma\theta_{i,k} = 1$. Here $\theta_{i,k}$ is k-th member in the vector θ_i. We then measured the difference between two articles through Euclidean distance between their vectors. The distance between articles i and j is

$$dis_{i,j} = \left\{ \sum \left(\theta_{i,k} - -\theta_{j,k} \right)^2 \right\}^{0.5}. \tag{2}$$

Obviously, $dis_{i,j} = dis_{j,i}$. Based on those definition, we discussed the authors' research model as well as the articles' citation problems.

3.1 Authors' Research Development Mode

After the LDA calculation, each article gets a vector representation, so that each article can be regarded as a point in the academic space, and each author's academic development is a trajectory in the academic space, and each point in the trajectory corresponds to one of his published articles. Figure 2 shows the career track of the two authors whose number of papers is about thirty. Based on this, we can analyze the research scopes of scholars and the migration of academic ideas. Besides, we tried to discuss the author's research scope and the relevance of their previous and present studies through the distance between the articles. Figure 3 shows the distance distribution of the different articles. In the first subplot (a), we selected two articles from all papers completely randomly and calculated the distance between them. In the case of random papers by the same author (b), we first randomly choose the author, then we calculated the distance between the randomly selected articles of that author. In the same author successive papers case (c), we ordered the author's articles by the publishing date, and then we calculated the distance between articles by chronological order. We found that those subplot's distributions are subject to the normal distribution with $u = 0.4845$ and $\sigma^2 = 0.0081$ for case (a); $u = 0.3643$ and $\sigma^2 = 0.0098$ for case (b); $u = 0.3571$ and $\sigma^2 = 0.0095$ for case (c). In Fig. 2(d), from the cumulative distribution functions (CDF), we also found that the mean value of the random authors' papers is greater than that of the same author's random papers and the mean of same author's random papers is greater than that of the same author's successive papers.

fridaya@uef.fi

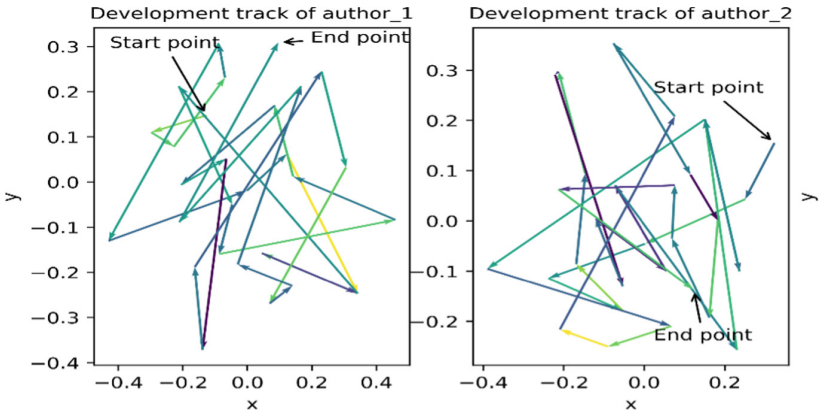


Fig. 2. Development track maps of two different authors, the changes in their academic fields and ideas can be seen from the map.

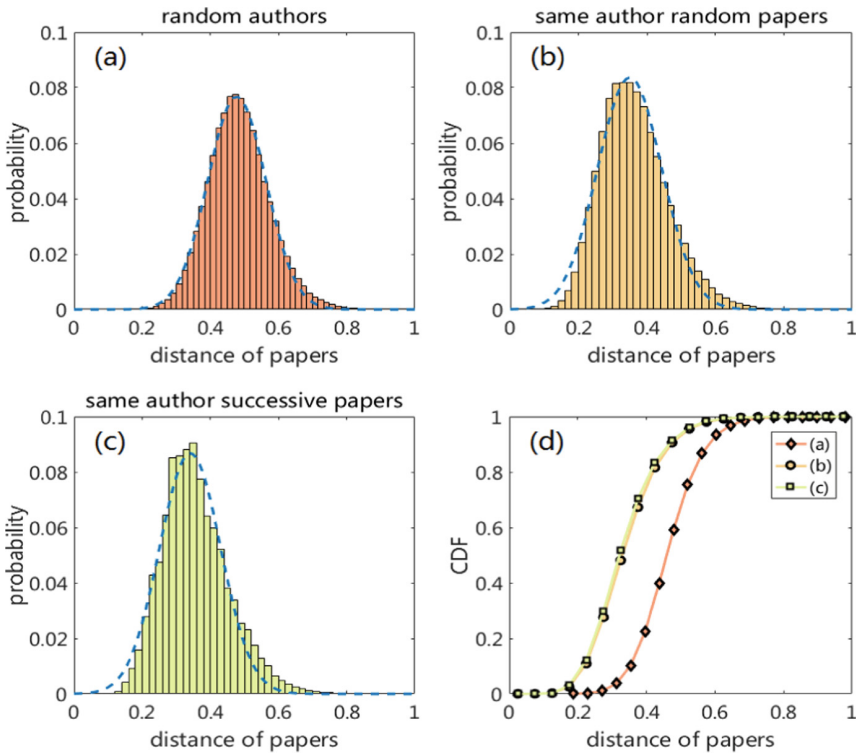


Fig. 3. Distributions of articles' distance in different cases. Subplots (a), (b), and (c) show the distributions of distance of randomly chosen papers, random papers of the same author and adjacent papers of the same author respectively. Subplot (d) shows the cumulative distribution functions.

fridaya@uef.fi

The small distance between the same author's previous and present articles is easy to understand. This phenomenon reflects the inheritance of thought for an author. For example, if a scientist is working in a field, he needs to stay in this field for several years. During those years, scientists may publish several academic papers in this field. Based on the same question, the study area and approaches are similar, thus making continuous publications of higher relevance.

The same authors' articles ordered by time have the minimum distances, while the randomly selected papers have the maximum distance. Two factors account for this. First, in a given period of time, the scientists' research direction is likely to belong to the same project, so there will be some words and ideas repeated in similar articles. Secondly, different authors have different research fields, contents and methods, which will inevitably lead to different authors' article distance greater than the same author's article distance, especially when the authors' research fields vary greatly from each other.

3.2 Property of Authors' Citation

When scientists write papers, they usually cite some published literature to highlight the significance of their research or reveal what technologies they adopted. Citation is the process of knowledge inheritance. Generally speaking, articles that investigate the same type of problem or have a similar knowledge background will have a greater probability of being referenced.

For any article i, represent the number of its citations as m, and randomly select m articles whose publication dates are earlier than the publication date of article i. Therefore, we get the random citation list $V_Ref_Rand_i$ of the article i and repeat the process 5000 rounds. The number of citations at different distances in the actual model and the control model are counted separately, and are recorded as realRef and randRef. The distribution and statistical indicators of these two statistics are as follows (the red inverted triangle indicates realRef and the blue circle indicates randRef) (Fig. 4):

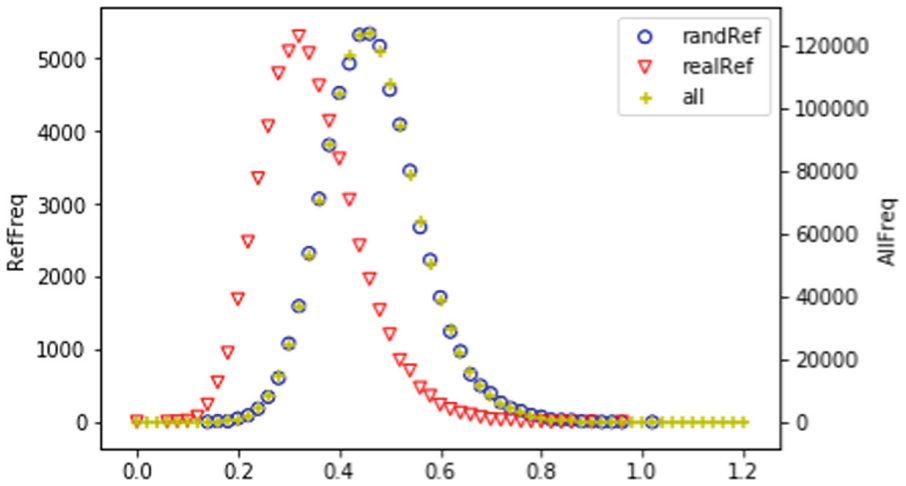


Fig. 4. Comparison of citation distributions under actual situation and random hypothesis. Red inverted triangles present the real citations. Blue circles are for randomly chosen citation. Yellow plus are for randomly chosen paper pair.

fridaya@uef.fi

Intuitively, the shapes of the frequency distribution images of realRef and randRef are similar, but the shape of realRef is more concentrated. Through statistical analysis, we can find that the mean and median of realRef are both 0.34 and the standard deviation is 0.10, which are less than the corresponding indicators of randRef. At the same time, realRef is right-biased (the skewness is 1.26, much larger than 0), and the kurtosis is 3.09. After JB test, it can be considered that realRef does not obey the normal distribution at 99% confidence level; while randRef has zero skewness, unbiased, the kurtosis is 1.8, and can be considered to be a normal distribution. The Kolmogorov-Smirnov test is performed on both, and the K-S statistic is 0.58, and the p value is close to 0. It can be considered that randRef and realRef do not belong to the same distribution, that is, the scholar's citation behavior is not random.

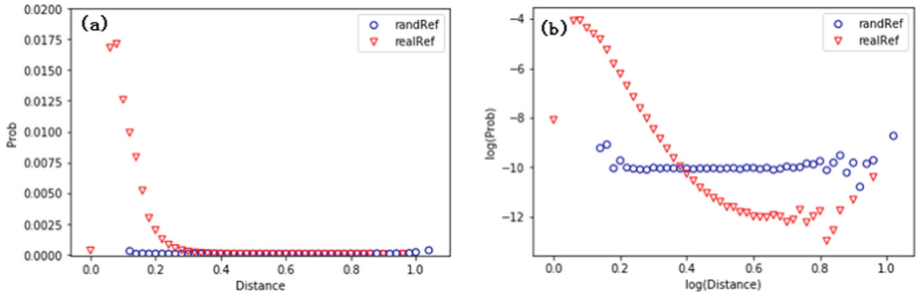


Fig. 5. Comparison of citation distributions under actual situation and random. Red inverted triangles present the real citations. Blue circles are for randomly chosen citations.

Figure 5 shows the distributions of the reference probabilities at different distances in the actual situation and the hypothetical situation. Subplot 5(b) shows the image after taking the logarithm of the probability. It can be seen that the reference probability of the hypothetical case (randRef) is basically stable and insensitive to the change of the distance; while in the actual situation (realRef) scholars' reference behavior is obviously affected by the distance of the article pairs, and the reference preference rises first and then decreases as the distance of the article pair increases. The reference preference reaches a maximum at a distance of about 0.1, and the citation behavior is basically concentrated in the interval [0, 0.4], and the articles those are farther away (the content difference is larger) are rarely cited.

3.3 Citation to Different Types of Articles

From the above research, we have learned that the scholar's citation behavior is not random. So what factors are related to scholars' citation behavior? Here, we assume that different types of articles will affect the scholar's citation behavior and obtain the following results.

fridaya@uef.fi

In this paper, we discussed 3 types of distance. The subplot 6(a) shows the citation of academic masters' articles. The masters are defined as the scholars who rank in the top 1000 according to the number of their published articles in our database. The citation to those masters is 63,245 in total. The second subfigure 6(b) citation for the author's articles shows the relationship of the reference probability and citation distance between the author's own articles. When an author writes papers, there may be a high probability to cite his or her previous articles. There are 19,408 of this type of citations in our database. Figure 6(c) represents citation for other articles means the authors cite neither academic masters' nor their own articles. There are 758,827 of this type of citations in our database.

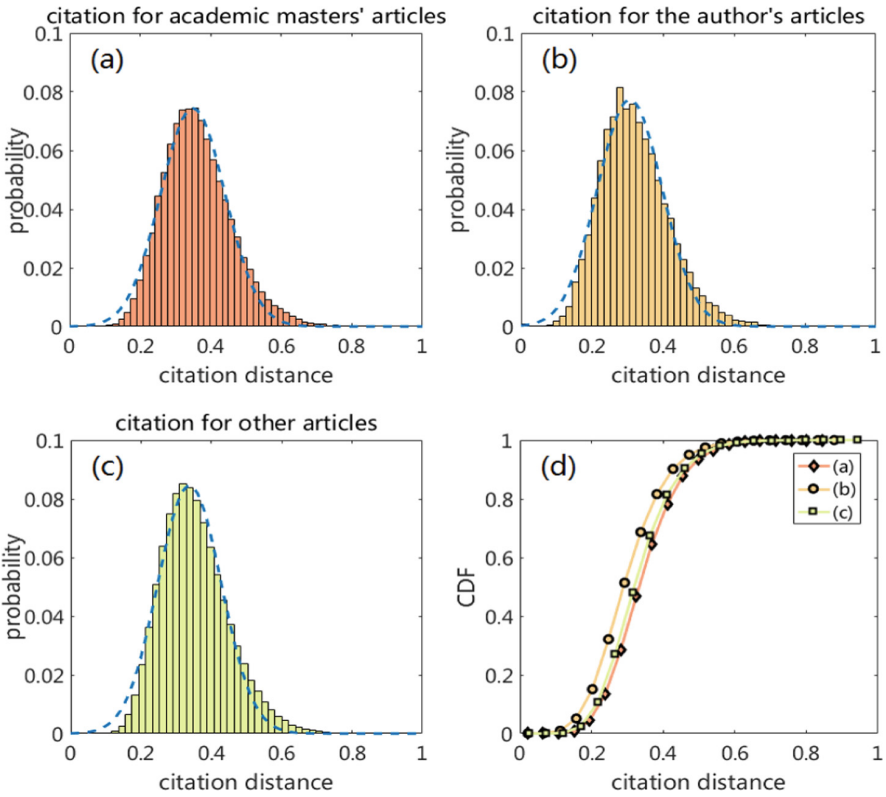


Fig. 6. Citation distance distribution. Subplots (a), (b), and (c) show the citation of academic masters' articles, the author's own articles and the other articles. Subplot (d) is CDFs for them.

The distributions of those subfigures are subject to the normal distributions with $u = 0.3630$ and $\sigma^2 = 0.0094$, for case (a); $u = 0.3211$ and $\sigma^2 = 0.0095$ for case (b); $u = 0.3534$ and $\sigma^2 = 0.0092$ for case (c). The cumulative distribution function (d) shows the citation on the author's articles has the minimum mean value, while the citation on academic masters' articles has the maximum mean value. This indicates scientists' new research is highly relevant to previous studies, and they have a strong preference to cite the academic masters' articles.

fridaya@uef.fi

4 Discussion

The changes or evolution of a scholar's research field have always been a hot topic in science. In this article, we embedded the article's abstract into a high dimensional space through the LDA model. Different from the citation network and cooperation network which mainly focus on network structure analysis, this paper puts forward a new method which is based on the content of the articles. In addition, this paper projects the topic vectors of the scholars to the low-dimensional space and draws the scholars academic career tracks, so that we can analyze the research scopes of scholars and the migration of academic ideas.

We found that the movement distances of scholars are relatively small and the distance of one same author's adjacent papers is different from the randomly selected papers. It showed that most scholars focus on their own study fields and the newly published articles of the scholars are highly correlated with their previous studies.

Besides, by comparing frequency distributions of citation distance in the real situation and the random model, we can conclude that the scholar's citation behavior is not random. Furthermore, when we explore the distance between citing and cited articles, we found that most articles have a bigger distance from the articles of academic masters than the distance from the average scholars' articles, which indicates that scholars prefer to cite classical papers or masters' articles. This is a strong evidence to show the strong preference to cite the academic masters' articles. Furthermore, there is "most proper citation distance" phenomena.

What's more, our work is a creative attempt, because we maps the abstracts of articles from the journals of APS to vector space and discuss some of the scholar's behavior. We are currently working on some related work by using other data and we believe that this law also applies to other disciplines. Furthermore, this working framework combining network and semantic analysis is a good starting point. In the future, we will use other data sources to indicate the success status of scholars and cluster or classify the scholars' development tracks so that the significant differences in behavioral patterns between successful scholars and general scholars can be studied. Therefore key characteristics determining the success of scholars can be found and thus we can give optimization suggestions for the scholar's academic process.

These results are based on the distance of the articles in the generated so called academic space. This article creatively represents the content analysis based on an article's abstract to get the deep information behind the article. Thus the results based on contents are more reliable. Our paper discusses the social problem based on combination of big data technology (LDA) and statistics (physics and mathematics). It is a meaningful attempt and we believe the paper should be of particular interest to the participants of the conference as it proposed a new perspective to research this human behavior problem.

Acknowledgments. We appreciate comments and helpful suggestions from Prof. Zengru Di, Jiang Zhang and Honggang Li. And we appreciate the reviewers' comments so that we can make the article logically clearer and more complete. This work was supported by Chinese National Natural Science Foundation (71701018, 61673070 and 71671017), and Beijing Normal University Cross-Discipline Project.

fridaya@uef.fi

References

1. Bernal, J.D.: Science in History, 3rd edn. Watts, London (1965)
2. Hou, H., Qu, T., Liu, Z.: The rise and development of science studies in China's universities. Stud. Sci. Sci **27**, 334–339 (2009). (in Chinese)
3. Mingers, J., Loet, L.: A review of theory and practice in scientometrics. Eur. J. Oper. Res. **1**, 1–19 (2015)
4. Albert, R., Barabási, A.-L.: Topology of evolving networks: local events and universality. Phys. Rev. Lett. **85**(24), 5234–5237 (2000)
5. Otte, E., Rousseau, R.: Social network analysis: a powerful strategy, also for the information sciences. J. Inf. Sci. **28**, 441–453 (2002)
6. Wu, J.S., Di, Z.R.: Complex networks in statistical physics. Prog. Phys. **24**, 18–46 (2004). (in Chinese)
7. Li, Y., He, K., Bindel, D., Hopcroft, J.: Uncovering the small community structure in large networks: a local spectral approach. Comput. Sci. **36**, 658–668 (2015)
8. Berahmand, K., Samadi, N., Sheikholeslami, S.M.: Effect of rich-club on diffusion in complex networks. Int. J. Mod. Phys. B **32**, 1850142 (2018)
9. Gualdi, S., Yeung, C.H., Zhang, Y.C.: Tracing the evolution of physics on the backbone of citation networks. Phys. Rev. E **84**(4), 046104 (2011)
10. Wu, Z.X., Holme, P.: Modeling scientific-citation patterns and other triangle-rich acyclic networks. Phys. Rev. E **80**(3), 037101 (2009)
11. Karrer, B., Newman, M.E.: Random acyclic networks. Phys. Rev. Lett. **102**(12), 128701 (2009)
12. Medo, M., Cimini, G., Gualdi, S.: Temporal effects in the growth of networks. Phys. Rev. Lett. **107**(23), 238701 (2011)
13. Barabási, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek, T.: Evolution of the social network of scientific collaborations. Physica A Stat. Mech. Appl. **311**(3–4), 590–614 (2002)
14. Duan, Q.F., Zhu, D.H., Wang, X.F.: Citation literature ranking method based on improved PageRank algorithm. Inf. Stud. Theor. Appl. **1**, 26 (2001). (in Chinese)
15. Dunnick, N.R.: The H index in perspective. Acad. Radiol. **24**, 117–118 (2017)
16. Radicchi, F., Fortunato, S., Markines, B., Vespignani, A.: Diffusion of scientific credits and the ranking of scientists. Phys. Rev. E **80**(5), 056103 (2009)
17. Zhou, Y.B., Lü, L., Li, M.: Quantifying the influence of scientists and their publications: distinguishing between prestige and popularity. N. J. Phys. **14**, 033033 (2012)
18. Gualdi, S., Medo, M., Zhang, Y.C.: Influence, originality and similarity in directed acyclic graphs. EPL (Europhys. Lett.) **96**(1), 18004 (2011)
19. Shen, H.W., Barabási, A.-L.: Collective credit allocation in science. Proc. Natl. Acad. Sci. **111**(34), 12325–12330 (2014)
20. Niu, Q., Zhou, J., Zeng, A., Fan, Y., Di, Z.: Which publication is your representative work? J. Informetrics **10**(3), 842–853 (2016)
21. Lü, L., Zhou, T.: Link prediction in complex networks: a survey. Physica A Stat. Mech. Appl. **390**(6), 1150–1170 (2011)
22. Sinatra, R., Wang, D., Deville, P., Song, C., Barabási, A.-L.: Quantifying the evolution of individual scientific impact. Science **354**(6312), aaf5239 (2016)

fridaya@uef.fi

23. Blei, D.M., Ng, A., Jordan, M.: Latent Dirichlet Allocation. J. Mach. Learn. Res. **3**, 993–1022 (2013)
24. Yuan, J., Zheng, Y., Xie, X.: Discovering regions of different functions in a city using human mobility and POIs. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 186–194. ACM: New York, USA, Beijing, China (2010)

fridaya@uef.fi

Check for updates

Hybrid of Filters and Genetic Algorithm - Random Forests Based Wrapper Approach for Feature Selection and Prediction

Pakizah Saqib(✉), Usman Qamar, Andleeb Aslam, and Aleena Ahmad

Department of Computer and Software Engineering, CEME, National University of Sciences and Technology, Islamabad, Pakistan
pakizahfatima@gmail.com, andleebaslam0@gmail.com, usmanq@ceme.nust.edu.pk, aahmadl7@ce.ceme.edu.pk

Abstract. Data is too diverse. The diversity of data does not just exist in terms of dimensionality but also of varied datatypes. To extract most useful information from datasets and to improve the prediction accuracy, feature selection is of great importance in data mining. This paper is proposing a hybrid feature selection methodology with the motivation of producing most relevant feature subset and better predicting accuracy. A wrapper composed of Genetic Algorithm (GA), a heuristic search tool and Random Forest (RF) as a predicting model, keeping in view the optimality of Genetic Algorithm and predictive accuracy of Random Forests, is suggested. And for the purpose of creating a reduced search space for GA-RF wrapper, set of filters methods are used, which generate a reduced subset of features by weight assignment and filtration through threshold criteria. The proposed approach has been tested on Breast Cancer dataset from UCI repository and produced 99.04% prediction accuracy. A small comparative study is also carried out to justify that coupling of genetic algorithm and random forests followed by space reduction outperforms other wrapper-based approaches.

Keywords: Feature selection · Hybrid · Genetic algorithm · Random forest

1 Introduction

Data is too diverse. The diversity of data does not just exist in terms of dimensionality but also varied datatypes [1, 2]. Selecting and extracting relevant features from the dataset is a main challenge for researchers in data mining [3, 4].

Feature Selection from vast datasets without losing the relevancy and usefulness of information, while facilitating the predicting model to achieve higher accuracy and better performance is main objective of feature selection.

Different researchers have carried out different surveys regarding usefulness of varied techniques of feature selection filter, wrapper and embedded [1, 2]. And have shown with their work that wrapper FS technique outperforms other but its computationally expensive [5].

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 190–199, 2019.
https://doi.org/10.1007/978-3-030-22868-2_15

fridaya@uef.fi

This paper is addressing two problems:

1. Feature Selection
2. Performance of Heuristic search technique Genetic Algorithm.

Moreover, in this paper an un-attempted combination of wrapper, in which GA is wrapped with random forests for feature reduction and performance evaluation, is used and experimented.

In this paper, space reduction has been carried out using fast filter techniques Info gain, gain ratio, Gini index, and correlation. And this reduced space has been then presented as an initial population to GA and Random forests based wrapper approach.

The rest of the paper is organized as follow: Sect. 2 consists on literature review of related work that has been done in this regard. In Sect. 3, a framework for the proposed methodology is presented and explained. In Sect. 4, an experiment has been carried out on Breast Cancer Original Dataset. In Sect. 5, comparison with other wrapper approaches has been presented.

2 Related Work

Feature Selection serves as a preprocessing technique for classification problems. Feature Selection by removing noisy, irrelevant and redundant features improves the accuracy of predicting model [1, 2]. This Field has acquired attention of many researchers for many years. In literature many techniques have been proposed for feature selection.

Researches on different techniques of feature selection has been carried out [1, 2]. Some Researchers has influenced on filter methods [6] for feature selection because of their computational cost effectiveness but other has focused on wrapper methods [7] because it generates more effective and accurate feature subsets [5].

We have done comparison of four filter methods to come up with the best filter method to form a hybrid for feature selection [6]. Some previous Feature selection algorithms are surveyed [8] and some new [9] are proposed to make predictions more accurate [10]. This paper proposes a new feature selection algorithm LW-index with Sequence Forward Search algorithm (SFS-LW).

Some researchers have proposed different hybrid approaches for feature selection utilizing the goodness of both Filter and Wrapper based approaches [11]. This paper has used a filter for selecting candidate feature and then used SVM based wrapper to do further refinement on feature set [10]. This paper author has also suggested a hybrid Feature Selection Algorithm based on Particle Swarm optimization, FASTPSO and RAPID PSO and compared both algorithms. FastPSO outperformed RapidPSO as far as effectiveness is concerned but on the expense of computational cost.

Genetic Algorithm being an optimal tool, supportive for multi objective has been widely and successfully used for many domain like health industry, human activity classification, text mining, credit risk assessment where there exist optimality issues [12]. Many hybrids techniques containing genetic algorithm has been introduced.

In [13], the author has introduced a hybrid Genetic Algorithm with neural network (HGA-NN) for feature selection and classification. This paper has used GA for feature

fridaya@uef.fi

extraction and refinement (FEROM) [14]. GA has also been used for feature selection from Brain Image Dataset [15].

With the motivation of producing most relevant feature subset and better predicting accuracy this paper is introducing a kind of hybrid of filter and wrapper. A wrapper composed of Genetic Algorithm a heuristic search tool and Random Forest as a predicting model, keeping in view the optimality of Genetic Algorithm [16] and predictive accuracy of Random Forests [17] is used. And for the purpose of creating a reduced search space for GA-RF set of filters methods are used to generate a reduced subset of features by using weight assignment and threshold criteria. This search space reduction improves performance of GA-RF, by providing search area free from irrelevant features.

3 Framework for Proposed Methodology

The proposed methodology is comprised of two stages:

1. Space Reduction.
2. Feature Selection and Performance Evaluation by GA wrapped with Random Forests.

3.1 Space Reduction

The purpose of Space Reduction is to reduce the search space area for the Genetic Algorithm, thus avoiding irrelevant feature exploration.

In this stage set of fast filter techniques Info Gain, Gain Ratio, Gini Index, Correlation has been used, each of them takes complete feature set and assign weights on the basis of their relevancy to each attribute individually, then these weights are normalized and compared with threshold value. The attributes satisfying the set threshold criterion are kept while other are filtered out. The union of left attributes from each filter technique is carried out to get a reduced feature set containing only relevant features (Fig. 1).

3.2 Feature Selection and Performance Evaluation by GA Wrapped with Random Forests

This stage is carried out for the further refinement of the feature set by using an optimal and search heuristic approach Genetic Algorithm (Fig. 2). Genetic Algorithm by taking this reduced feature dataset as an initial population runs fitness function. The Random Forests is used as a predictive model to test the fitness of each individual. Genetic algorithm works by selecting the fittest individual from population for further reproduction. And keep on repeating this process and evolving new generation by applying mutation and crossover until it ends up with the generation with fittest individuals i.e. features in our case (Fig. 3).

fridaya@uef.fi

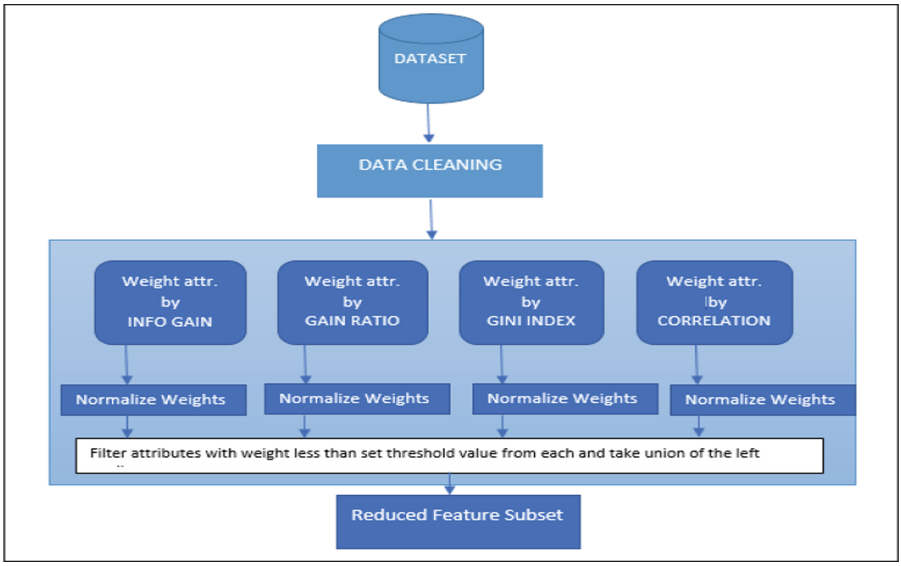


Fig. 1. Space reduction – first stage.

ALGORITHM

```
START
Generate the initial population
Compute fitness
REPEAT
    Selection
    Crossover
    Mutation
    Compute fitness
UNTIL population has converged
STOP
```

Fig. 2. Genetic algorithm.

fridaya@uef.fi

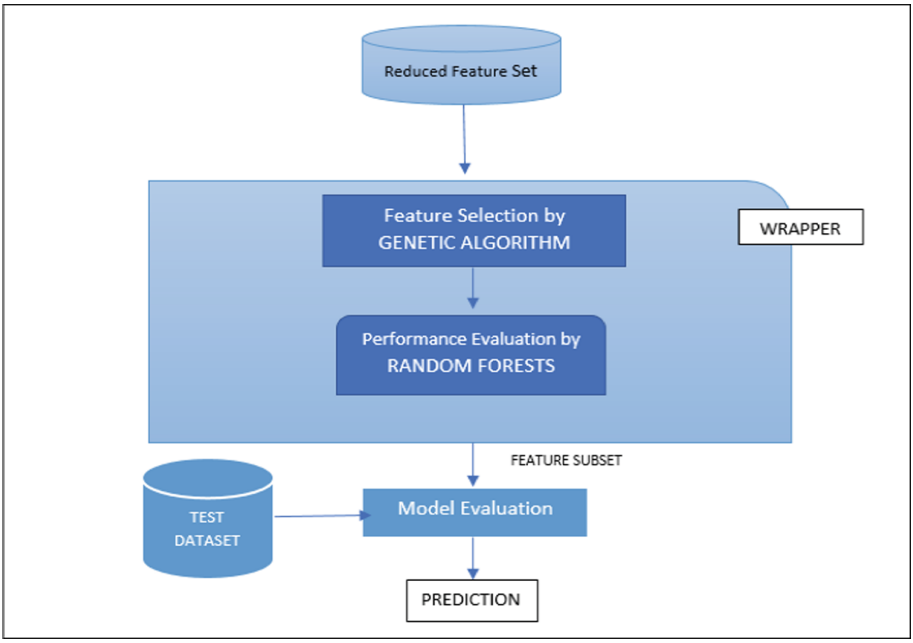


Fig. 3. Feature reduction and classification by GA - random forests wrapper – second stage.

4 Experiment

The experiment was carried out on different datasets to check how accurate is proposed approach? Experimentation on Breast Cancer (Original) Dataset is explained below.

4.1 Dataset

The Breast Cancer (Original) dataset is collected from UCI Data repository. The original Dataset was comprised of 699 data entries and 11 attributes, containing 10 regular attributes and 1 prediction attribute. The description of the attribute is in Table 1.

Table 1. Description of attributes of breast cancer dataset.

ID	Description	Values
att1	*Sample code number*	id number
att2	*Clump Thickness*	1–10
att3	*Uniformity of Cell Size*	1–10
att4	*Uniformity of Cell Shape*	1–10
att5	*Marginal Adhesion*	1–10

(*continued*)

fridaya@uef.fi

Table 1. (*continued*)

ID	Description	Values
att6	*Single Epithelial Cell Size*	1–10
att7	*Bare Nuclei*	1–10
att8	*Bland Chromatin*	1–10
att9	*Normal Nucleoli*	1–10
att10	*Mitoses*	1–10
att11	*Class*	2, 4 (2 for Benign, 4 for Malignant)

4.2 Tool

Rapid Miner is used to compute the results.

5 Process and Results

5.1 Data Cleaning

Data Cleaning is carried out and missing attributes are filtered out.

5.2 Data Splitting

Data is partitioned into two parts testing data and training data in 70, 30 ratios, respectively.

5.3 Space Reduction

Space Reduction is carried out on Breast Cancer dataset by following these steps:

Weight Assignment
The weights are assigned by each filter technique as shown in Table 2 below.

Table 2. Weights assigned by each filter.

Attributes	Info gain	Gain ratio	Gini index	Correlation
att1	0	0	0	0
att2	0.360	0.264	0.392	0.239
att3	0.671	0.952	0.705	0.878
att4	0.638	0.815	0.668	0.826
att5	0.338	0.518	0.392	0.666
att6	0.430	0.618	0.506	0.975
att7	0.556	1	0.624	0.579
att8	0.478	0.487	0.535	0.756
att9	0.375	0.655	0.455	1
att10	0	0.444	0	0.487

fridaya@uef.fi

Weight Normalization
Weights are normalized to convert the values within the range of 0 and 1.

Threshold Adjustment
Different threshold values are adjusted and tested to get the value of threshold that results in best feature subset. Table 3 shows different threshold values, corresponding resultant feature subsets and accuracies. For Breast Cancer Dataset, we opted 0.5 threshold value as it filtered out the best resultant feature subset.

Table 3. Different threshold values and resultant feature subsets

Threshold values	Resultant feature subset	Resultant accuracies
0.4	{att3, att4, att5, att6, att7, att8, att9, att10}	99.04%
0.5	{att3, att4, att5, att6, att7, att8, att9}	99.04%
0.7	{att3, att4, att6, att7, att9}	98.09%
0.9	{att7, att9}	96.65%

Attributes Filtration
Attributes with low weight value as compared to set threshold are filtered out (Table 4).

Table 4. Feature subset by each filter.

	att1	att2	att3	att4	att5	att6	att7	att8	att9	att10
Info gain	Out	Out	In	In	Out	Out	In	Out	Out	Out
Gain ratio	Out	Out	In	In	In	In	In	In	In	Out
Gini index	Out	Out	In	In	In	In	Out	In	In	Out
Correlation	Out	Out	In	In	Out	In	In	In	In	Out

Resultant Feature Subset
Union of left attributes is carried out to generate a feature subset.
Resultant Feature Set = {3, 4, 5, 6, 7, 8, 9}

5.4 Feature Selection and Performance Evaluation by GA Wrapped with Random Forests

This wrapper-based approach is applied on the reduced feature subset of Breast Cancer Dataset achieved from Step 3. This step resulted in Final Feature Set comprised of 4 features {att3, att4, att6, att8}.

5.5 Model Evaluation

Model Evaluation is carried out on 30% of dataset that was already separated for test purpose. 99.04% accuracy is obtained. Model Accuracy, Precision and Recall obtained can be seen in Table 5.

fridaya@uef.fi

Table 5. Performance vector

	True: Benign	True: Malignant	Class precision
Prediction: Benign	136	1	99.27%
Prediction: Malignant	1	71	98.61%
Class recall	99.27%	98.61%	

6 Comparison

6.1 Comparison of Different Wrapper Approaches for Feature Selection

Space reduction followed by two other Feature Selection Techniques like Forward Selection and Backward elimination in combination with random forest are tested for same Breast Cancer dataset, to justify that the accuracy of our proposed approach is better than other techniques (Fig. 4).

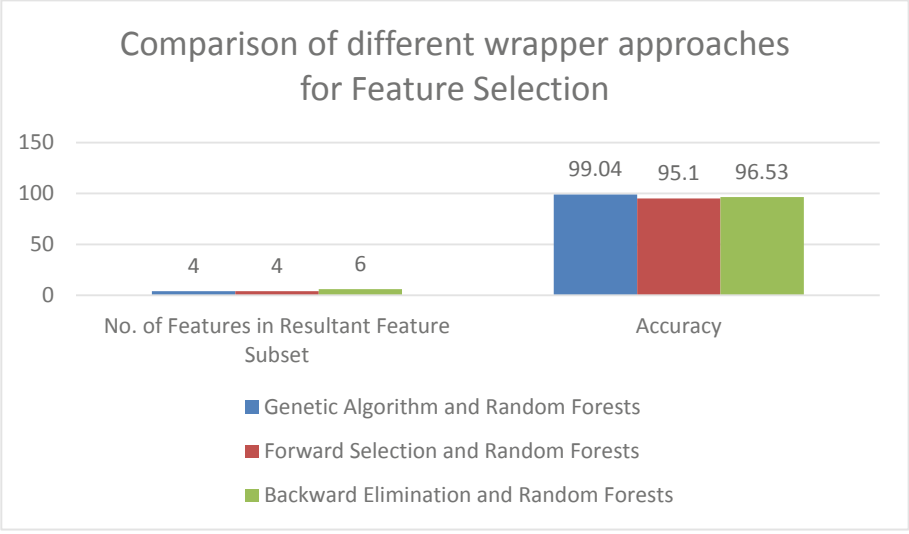


Fig. 4. Comparison of different wrapper approaches for feature selection.

6.2 Comparison with Some State of the Art Techniques

Comparison with some state of the art techniques in Table 6 shows that the accuracy of our proposed approach i.e. 99.04% is better than other techniques.

fridaya@uef.fi

Table 6. Comparison with some state of the art techniques

Techniques	Accuracy % for breast cancer dataset	Reference
Naive Bayesian + KNN	81.6	[18]
PSO based SVM	91.3	[19]
NN + adaptive Neuro Fuzzy	83.6	[20]
KNN + ICA	92.5	[21]
GA + RF	99.04	Proposed

7 Results and Discussion

The proposed methodology was tested on breast cancer dataset and its prediction accuracy was compared with other state of the art methods proposed by different researchers. Comparison in Table 6 justifies the effectiveness of our proposed method.

To further assess the effectiveness of our proposed methodology, we tested this model on other datasets from UCI repository. For small datasets with 10–33 attributes, this model outperformed many competitive proposed models.

Computational cost factor of wrapper methods has always been a challenge. Usage of filter techniques in proposed methodology to reduce search space of Genetic Algorithm reduces computation overhead of GA-RF wrapper but still there is a room for more research in this regard.

In this study, manual and comparative approach has been used to set appropriate threshold value for available dataset in pre-stage, which needs automation in future to come up with the most suitable threshold value hassle-free, with respect to each dataset provided.

8 Conclusion

In this study we proposed a novel hybrid of filters and GA-RF based to improve the feature selection and prediction accuracy. The proposed methodology has been tested on various datasets from UCI, and it has outperformed other feature selection techniques. The purpose of this methodology is to utilize the goodness of GA and Random Forests by coupling them together. And to improve the performance of this unattempted wrapper-based approach by reducing search space using Filter Feature Selection techniques in pre-stage. In future, proposed model will be experimented on large data sets and research work will be extended to minimize the computational expense of GA-RF based wrapper.

References

1. Miao, J., Niu, L.: A survey on feature selection. Procedia Comput. Sci. **91**, 919–926 (2016)
2. Chandrashekar, G., Sahin, F.: A survey on feature selection methods. Comput. Electr. Eng. **40**(1), 16–28 (2014)

fridaya@uef.fi

3. Li, J., Liu, H.: Challenges of feature selection for big data analytics. IEEE Intell. Syst. **32**(2), 9–15 (2017)
4. L'heureux, A., et al.: Machine learning with big data: challenges and approaches. IEEE Access **5**(5), 777–797 (2017)
5. Suto, J., Oniga, S., Sitar, P.P.: Comparison of wrapper and filter feature selection algorithms on human activity recognition. In: 2016 6th International Conference on Computers Communications and Control (ICCCC). IEEE (2016)
6. Sánchez-Maroño, N., Alonso-Betanzos, A., Tombilla-Sanromán, M.: Filter methods for feature selection – a comparative study. In: International Conference on Intelligent Data Engineering and Automated Learning. Springer, Heidelberg (2007)
7. Kohavi, R., John, G.H.: Wrappers for feature subset selection. Artif. Intell. **97**(1–2), 273–324 (1997)
8. Khalid, S., Khalil, T., Nasreen, S.: A survey of feature selection and feature extraction techniques in machine learning. In: Science and Information Conference (SAI). IEEE (2014)
9. Liu, C., et al.: A new feature selection method based on a validity index of feature subset. Pattern Recognit. Lett. **92**, 1–8 (2017)
10. Butler-Yeoman, T., Xue, B., Zhang, M.: Particle swarm optimisation for feature selection: a hybrid filter-wrapper approach. In: CEC (2015)
11. Hsu, H.-H., Hsieh, C.-W., Lu, M.-D.: Hybrid feature selection by combining filters and wrappers. Expert Syst. Appl. **38**(7), 8144–8150 (2011)
12. Bouaguel, W.: A New Approach for Wrapper Feature Selection Using Genetic Algorithm for Big Data. Intelligent and Evolutionary Systems, pp. 75–83. Springer, Cham (2016)
13. Oreski, S., Oreski, G.: Genetic algorithm-based heuristic for feature selection in credit risk assessment. Expert Syst. Appl. **41**(4), 2052–2064 (2014)
14. Mishra, R.: FEROM (Feature extraction and refinement) using genetic algorithm. In: 2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT). IEEE (2015)
15. Szenkovits, A., et al.: Feature selection with a genetic algorithm for classification of brain imaging data. In: Advances in Feature Selection for Data and Pattern Recognition, pp. 185–202. Springer, Cham (2018)
16. Lim, T.Y.: Structured population genetic algorithms: a literature survey. Artif. Intell. Rev. **41** (3), 385–399 (2014)
17. More, A.S., Rana, D.P.: Review of random forest classification techniques to resolve data imbalance. In: 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM). IEEE (2017)
18. Güzel, C., Kaya, M.M., Yıldız, O.: Breast cancer diagnosis based on Naïve Bayes machine learning classifier with KNN missing data imputation. In: 3rd World Conference on Innovation and Computer Sciences (2013)
19. Liu, X., Fu, H.: PSO-based support vector machine with Cuckoo search technique for clinical disease diagnoses. Sci. World J. **2014**, Article ID 548483, 7 pages (2014). https://doi.org/10.1155/2014/548483
20. Huang, M.-L., et al.: Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. J. Med. Syst. **36**(2), 407–414 (2012)
21. Mert, A., et al.: Breast cancer detection with reduced feature set. In: Computational and Mathematical Methods in Medicine 2015 (2015)

fridaya@uef.fi

A New Arabic Dataset for Emotion Recognition

Amer J. Almahdawi[1(✉)] and William J. Teahan[2]

[1] College of Science for Women, Computer Department, University of Baghdad, Baghdad, Iraq
amer.almahdawi@gmail.com
[2] College of Computer Science and Electronic Engineering, Computer Department, Bangor University, Bangor, UK
w.j.teahan@bangor.ac.uk

Abstract. In this study, we have created a new Arabic dataset annotated according to Ekman's basic emotions (Anger, Disgust, Fear, Happiness, Sadness and Surprise). This dataset is composed from Facebook posts written in the Iraqi dialect. We evaluated the quality of this dataset using four external judges which resulted in an average inter-annotation agreement of 0.751. Then we explored six different supervised machine learning methods to test the new dataset. We used Weka standard classifiers ZeroR, J48, Naïve Bayes, Multinomial Naïve Bayes for Text, and SMO. We also used a further compression-based classifier called PPM not included in Weka. Our study reveals that the PPM classifier significantly outperforms other classifiers such as SVM and Naïve Bayes achieving the highest results in terms of accuracy, precision, recall, and F-measure.

Keywords: Emotions recognition · Text categorization · Machine learning · PPM · Weka · Arabic corpora · Arabic NLP

1 Background and Motivation

A primary goal of the research described in this paper is to automatically recognising emotions in Arabic text (specifically, the Iraqi dialect) according to Ekman's [12] fine-grained emotion classification (Anger, Disgust, Fear, Happiness, Sadness, Surprise). To achieve this goal, a suitable 'gold standard' dataset of Arabic text for research experiments is needed where emotions in the text have been manually annotated. For evaluating any automatic learning system, annotated data is a prerequisite for performing a robust evaluation. However, our research in automatically recognising emotions in Arabic text is obstructed by the lack of publicly annotated data for written Arabic text.

Automatic text processing of Arabic language text has become a goal for many Natural Language Processing and text mining research [3,4,10]. However, despite the Arabic language being one of the top five spoken languages, there is

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 200–216, 2019.
https://doi.org/10.1007/978-3-030-22868-2_16

fridaya@uef.fi

a lack of emotion and sentiment datasets. This is one of the reasons for creating a new dataset for emotion recognition.

A study on available datasets shows none of the available Arabic corpora is suitable for research related to emotion recognition. We considered the Arabic Twitter corpus [16] which consists of 8,868 tweets which has been annotated with a particular positive, negative and neutral sentiment. But this corpus is inappropriate for this research as it does not support fine-grained emotion classification. Instead it supports only negative, positive, and neutral states. The Arabic tweets corpus composed by Abdulla et al. [2] consists of 1000 positive tweets and 1000 negative tweets [2] and also does not support Ekman's emotions. Instead, it supports the polarity of the tweet as positive and negative, so it is unsuitable for this research. Another corpus called AWATIF created by Abdul-Mageed et al. [1] is a multi-genre Modern Standard Arabic corpus for the purpose of sentiment analysis and subjectivity which is also not related to this research [1]. The corpus called LABR for sentiment analysis in the Arabic language consists of 63,000 books reviews by Aly [8]. One more available dataset called HAAD is composed of book reviews in the Arabic language but this dataset is again annotated just for sentiment analysis purposes, not for Ekman's emotions.

One of the closest corpora to our research is the corpus of micro tweets developed by Al-Aziz et al. [5]. It consists of 1552 tweets labelled with five emotions (Anger, Disgust, Fear, Happiness, and Sadness). Also it is composed of the Egyptian dialect and Modern Standard Arabic texts. Unfortunately, this corpus is not appropriate to our research as it supports just five emotions, not Ekman's six emotions. In addition, the text of this corpus is in the Egyptian regional dialect, whereas the goal of our research is to focus on another regional dialect—Iraqi.

According to the previous limitations of finding appropriate Arabic corpora to meet the requirements of this research, we decided to develop a new dataset. The most important consideration in choosing data in this research is the requirement that the text should often contain emotion-rich expressions. Another important consideration is the data should include many examples for all the emotion classes considered in this research. A question arises concerning where this type of data that expresses personal emotions can be obtained. A survey by Salem reveals that social media is the most appropriate place for 58% of Arab people to express their emotions toward their government's policies or services. 86% of these people who express their emotions in social media use Facebook as a platform, 28% use Twitter, and 28% use WhatsApp and other messaging applications, as shown in Fig. 1 [17].

Figure 2 shows the percentages of Arab people using social media platforms to express their emotions. The figure shows that Arab people prefer using Facebook as a platform compared to other platforms such as Twitter, Instagram and Snapchat.

Due to this study, Facebook has been chosen as the platform for collecting data for this research. Other considerations for dealing with public posts written by people are the misspelling of words and slang words included in the text. The classification system potentially needs to deal with these in some way.

fridaya@uef.fi

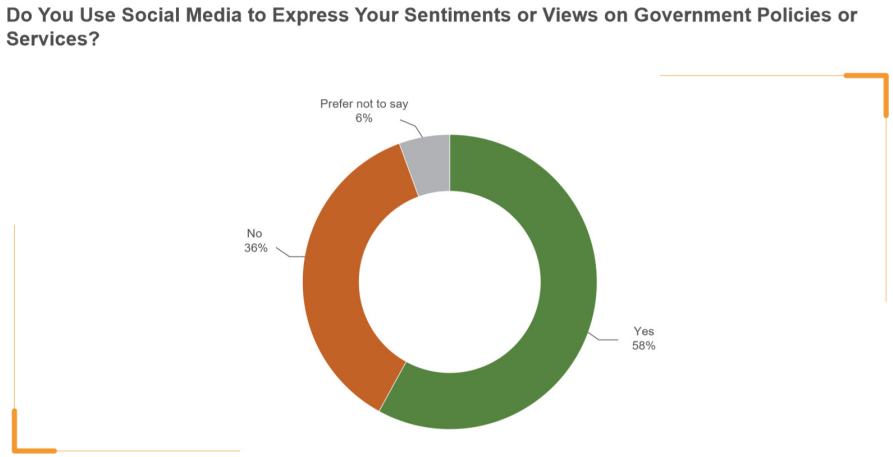


Fig. 1. The overall percentages of people who use social media to express their emotions [17].

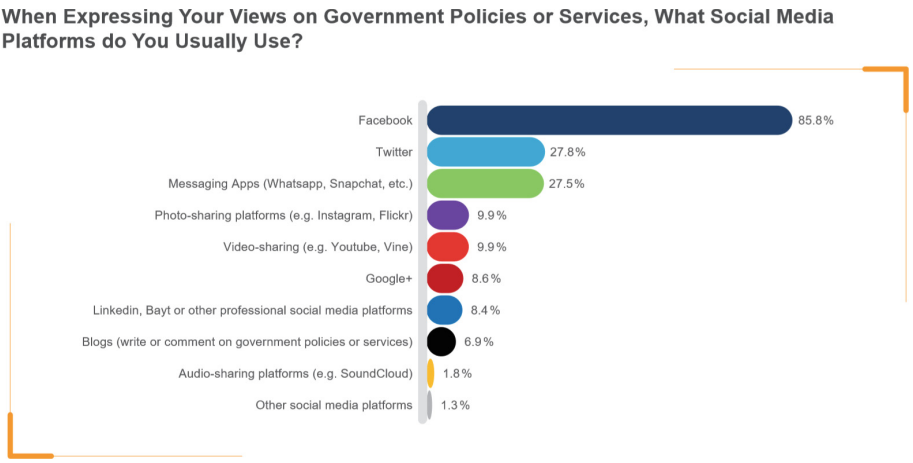


Fig. 2. The overall percentages of Arab people using social media to express their emotions [17].

This paper is organised as follows. Firstly, this paper creates an Arabic emotion dataset based on the six Ekman's emotions. Secondly, this paper evaluates this Arabic dataset using external judges to ascertain the quality of the dataset. Thirdly, this paper conducts an experimental evaluation using the new dataset by investigating how well various classifiers perform at identifying the emotions in the texts. The results are then discussed for each classifier and compared in the final section.

fridaya@uef.fi

2 Creating the New Arabic Dataset for Emotion Recognition

In this section, the process of creating a new Arabic dataset for emotion recognition will be described. We have named the new dataset IAEDS (which stands for Iraqi Arabic Emotions DataSet). As mentioned in the previous section, the data will be collected from the Facebook platform.

Facebook has features that help users searching for specific information. When a Facebook user is writing his/her post, Facebook supports users declaring his/her emotional state. This emotional state will appear in the post. This is shown in Fig. 3 which is an example of a Facebook post declaring the emotional state of the user (underlined in red at the top of the figure).

Fig. 3. The post of a user declaring his emotional state in Facebook.

As obvious in Fig. 3, the user in the post has declared their emotional state using the word "feeling" followed by the emotional state which is "happy". It is this "feeling" feature that we can use to help in searching posts for specific

fridaya@uef.fi

emotions. Specifically for the purposes of this research, we can use the search bar of Facebook to search for one of the Ekman's emotions. This is shown in Fig. 4 which shows the query for searching about posts that declare the angry emotion state.

Fig. 4. The query for searching Facebook for posts that have the angry emotional state.

After specifying this query, Facebook displays all posts that have the angry emotion. On the left side of the page, Facebook provides the user with a filter to search for more specific posts such as posts from friends or from all users, posts from users around the world or from a specific geographic place, posts from a specific date or not and so on. Figure 5 illustrates the filter that helps the user to search for more specific posts.

We used seed words defined by Aman [9] in the search bar of Facebook for searching for a specific emotion and its synonyms. We used a manual collection of data rather than automatic collection of data due to noise in the data such as links and images. As stated, we used the query "feeling happy" in the Facebook search bar when searching for happy posts and the same for synonyms of happy, and repeated this for the other emotions. Table 1 provides all the synonyms for each of Ekman's emotions used to collect posts for the dataset. These seed words were defined by [9], and we used these words in collecting the posts for the new Arabic emotion dataset.

Table 1. Seed words used to collect Facebook posts for the IAEDS dataset.

No.	Emotion	Synonyms
1	Anger	Angry, anger, annoyed, enraged, boiling, furious, mad, inflamed
2	Disgust	Disgusted, sucks, sickening, stupid, unpleasant, contempt, nauseating
3	Fear	Far, afraid, scared, frightened, insecure, nervous, horrified, panicked
4	Happiness	Happy, awesome, amused, excited, great, pleased, amazing, cheerful
5	Sadness	Sad, glooming, sorrowful, down, depressed, lonely, painful, guilty
6	Surprise	Surprised, confused, astonished, sudden, unexpected, shocked, perplexed

fridaya@uef.fi

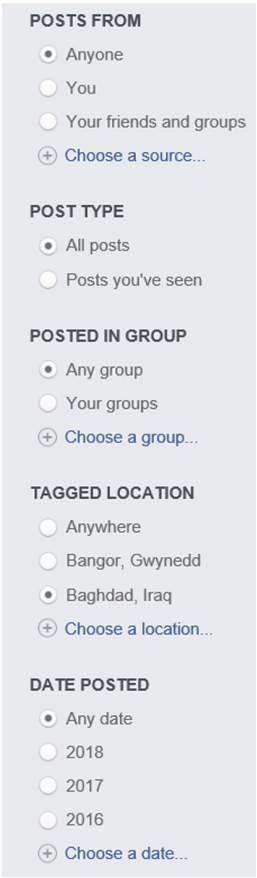


Fig. 5. Facebook's filter for searching for more specific posts.

Second, we used the following options in the search filter: we chose the option "anyone" from the filter "POSTS FROM"; the option "All posts" from the filter "POST TYPE"; the option "Any Group" from the filter "POSTED IN GROUP"; the option "Baghdad-Iraq" from the filter "TAGGED LOCATION"; and finally, the option "Any date" from the filter "DATE POSTED".

The reason for choosing the option "Baghdad-Iraq" for the filter "TAGGED LOCATION" is the lack of Arabic corpora specifically for the Iraqi dialect. There are a number of available corpora for Modern Standard Arabic or for regional Arabic dialects such as the Egyptian and Levante dialects, or for other regional dialects but there is less available for the Iraqi dialect [5]. Another reason for choosing "Baghdad-Iraq" as the "TAGGED LOCATION" is to focus on a specific Arabic dialect in order to see how difficult it is to recognise emotions for this dialect since research has shown that processing regional Arabic dialects can be significantly more problematic than processing Modern Standard Arabic, for

fridaya@uef.fi

example [5]. Variations between dialects can also pose problems—for example, some words mean one emotion in a certain Arabic dialect, but they can mean a very different emotion in another Arabic dialect. For example, the word منحبك in the Syrian dialect means "love you", but in the Iraqi dialect means the opposite "do not love you". Another example, in the Gulfian dialect is the idiom جبت العيد which means "you disappoint me" but in the Iraqi dialect means "you bring joy and happiness". Analysing emotion variations between Arabic regional dialects is therefore outside the scope of this research.

3 Description of the New Arabic Emotion Dataset

The text in the new Arabic emotion dataset consists of 1365 posts from Facebook. The posts were collected manually as mentioned in the previous section. Table 2 shows samples of these posts along with their English translations. We collected the data from December 2016 to August 2018.

Table 2. Samples of Facebook posts in the new Arabic dataset.

No.	Emotion	Post
1	Anger	طلعت روحي ابخسارت البرشا حرامات هيج ايفوز الريال والبرشه تطلع بره I nearly died when Barcelona lost the match, so Real Madrid won and Barcelona out.
2	Disgust	هذولة هم اكو امل يسوون شي ايجابي بالمستقبل Is there any hope they will do something positive in the future.
3	Fear	سترك ياربي شنو هاي صواريخ متت خوف بس انجي همزين بابا بمنه اذني كامت توجعني God save us, what are these rockets, I feared until death. I'm just crying. Thank God my father is here. My ears are hurting me.
4	Happiness	الف الف مبروك الله يسعدك و يهنئك بحياتك الزوجية Thousands of congratulations. God bless you in your marital life.
5	Sadness	الف رحمه ونور ع روحك الطاهرة يا بطل A mercy and light upon your pure spirit. You are such a hero.
6	Surprise	احلى مفاجأه من نور عيوني وعمري حمندشي ربي يحفظك الي The sweetest surprise from the light of my eyes and my life Hammandashi. God saves you to me.

Details of the IAEDS dataset are described in Table 3. The dataset consists of six sub-datasets. Each dataset consists of posts belonging to one of Ekman's emotions, i.e each sub-dataset represents one class. The table shows that the Anger class has the highest number of posts (309) with the fewest number of posts in the Fear (148) and Disgust (185) classes. This compares with the Happiness, Sadness and Surprise classes which have over 200 posts each. The total number of posts is 1,365, consisting of 22,438 words, and 286,775 characters.

We had some issues with collecting the data. Most Iraqi people were not including their feeling in their posts before 2013. After 2014, Iraqi people started

fridaya@uef.fi

Table 3. Number of posts, words, and characters in the IAEDS dataset.

No.	Emotion	#Posts	#Words	#Chars
1	Anger	309	6,960	71,028
2	Disgust	185	2,936	29,967
3	Fear	148	1,596	16,843
4	Happiness	256	2,514	27,886
5	Sadness	238	3,486	35,759
6	Surprise	229	4,946	52,533

using the feeling option on Facebook to state their emotional state in their posts. Even so, many users of Facebook in Iraq still do not use the feeling option to express their emotional state at the same time that they write their posts. This is due to them not being aware of this feature provided by Facebook, or they do not use any of the more advanced Facebook features at all. These issues led to a lack of suitable posts while we were collecting data for the dataset, and why it took longer than anticipated (about 18 months) to finish collecting the posts for the IAEDS dataset.

4 Dataset Evaluation

The goal of this comparison was to compare the emotion annotation between four annotators. The comparison was accomplished by measuring the inter-annotation agreement among the four annotators. This measurement can provide a valuable insight into the dataset usability and understandability. Four annotators (A, B, C, D) participated in the annotation process of the IAEDS dataset. Table 4 displays more details about the annotators who participated in the evaluation of IAEDS.

Table 4. Annotator details who participated in the IAEDS annotation process.

	A	B	C	D
Nationality	Iraqi	Iraqi	Iraqi	Iraqi
Qualifications	PhD. degree in Genetic Engineering and Bio-technology	MSc. degree in Science of Mathematics	MSc. degree in Electronics and Communication Engineering	MSc. degree in Mechanical Engineering
Experience in teaching	Teaching Genetics in the Disease Analysis Department at Alisraa University College	Teaching Mathematics in the Computer Science Department in the College of Science for Women at Baghdad University	Teaching in the Electronics Engineering College at Ninavah University	Teaching in the Mechanical Engineering Department at the University of Mosul
Experience in annotation	No previous experience	No previous experience	No previous experience	No previous experience

fridaya@uef.fi

One of the goals of this evaluation was to find out to what extent an untrained user could understand and use these posts in the dataset. It is known that variation in skills and the interest of the annotators, and the ambiguity in the annotation guidelines leads to disagreement among annotators [15]. The posts were written in Iraqi slang and many of these posts used idioms to express their emotions, unlike the corpora written in Classical Arabic or Modern Standard Arabic. It is hard to identify single words that declare or express a specific emotion. So we asked the annotators to evaluate each entire post as belonging to one of Ekman's emotions.

We used pairwise agreement to measure agreement between annotators, between A ↔ B, A ↔ C, A ↔ D, for each emotion in Ekman's emotions and we used the same pairwise analysis to evaluate the agreement for the whole dataset. Cohen's kappa co-efficient was used to calculate the pairwise agreement between annotators [11]. Commonly, the kappa co-efficient is used to calculate agreement between two annotators. Table 5 shows the pairwise kappa co-efficient between each pair of annotators for each emotion. The inter-annotator agreement value between A ↔ B was lowest for emotion Surprise with 0.721 and the highest value was for the Fear emotion with 0.785. On the other hand, the inter-agreement value between A ↔ C was lowest for the Fear emotion with value 0.706, but highest for the Sadness emotion with value 0.929. Finally, the inter-agreement value between A ↔ D was lowest for the Surprise emotion with 0.725; however, it was highest for the Fear emotion with 0.938.

Table 5. Kappa co-efficients for pairwise agreement among annotators per emotion.

Pair	Anger	Disgust	Fear	Happiness	Sadness	Surprise
A ↔ B	0.759	0.739	0.785	0.746	0.741	0.721
A ↔ C	0.840	0.826	0.706	0.796	0.929	0.710
A ↔ D	0.834	0.863	0.938	0.796	0.827	0.725

Table 6 displays the total inter-annotator agreement values among annotators, with the lowest agreement between A ↔ C (with a value of 0.679), and the best agreement between A ↔ D (0.824). The table also shows the average inter-annotator agreement for all annotators with an overall average value of 0.751 which is evidence of substantial agreement [14].

Table 6. Pairwise agreement amongst annotators.

	A ↔ B	A ↔ C	A ↔ D	Average
Kappa	0.749	0.679	0.824	0.751

fridaya@uef.fi

5 Experimental Results

In this section, various experiments were applied to classify Ekman's emotions from the Arabic text of IAEDS. The next two subsections report the results of applying different Weka classifiers [13] and the compression-based PPM classifier [18] to the Facebook posts in IAEDS.

5.1 Applying Weka Classifiers

We applied various classifiers supported by Weka [13] to find out the best classifier for Ekman's emotions. We used ten-fold cross-validation in our experiment. Table 7 lists results for five classifiers using Weka's `StringToWordVector` filter with `NGramTokenizer` to extract ngrams as features from the text, setting `NGramMaxSize` to 3, and `NGramMinSize` to 1. The worst text classifier performances were for ZeroR and Multinomial Naïve Bayes Text, which both achieving the same results with 74.2% accuracy, 0.04 precision, 0.17 recall, and 0.06 F-measure.

Table 7. Classification results using five classifiers supported by Weka.

Classifier	Accuracy	Precision	Recall	F-measure
J48	75.7	0.44	0.22	0.29
ZeroR	74.2	0.04	0.17	0.06
Naïve Bayes	75.9	0.49	0.23	0.31
Multinomial Naïve Bayes Text	74.2	0.04	0.17	0.06
SMO	75.9	0.49	0.22	0.31

Table 8 reports the confusion matrix for both ZeroR and Multinomial Naïve Bayes Text classifiers. (The numbers shown in bold font across the leading diagonal represent the number of correct classifications). The table shows that the classifiers simply labelled every post in the Anger class.

Table 8. Confusion matrix of Ekman's emotions classification for the IAEDS dataset using ZeroR and Naïve Bayes Multinomial text classifiers.

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	**309**	0	0	0	0	0
Disgust	185	**0**	0	0	0	0
Fear	148	0	**0**	0	0	0
Happiness	256	0	0	**0**	0	0
Sadness	237	0	0	0	**0**	0
Surprise	229	0	0	0	0	**0**

fridaya@uef.fi

On the other hand, both the SMO and Naïve Bayes classifiers achieved better results than the previous classifiers. Naïve Bayes was slightly better than SMO although they both achieved the same accuracy with 75.9%, with both Naïve Bayes and SMO achieving 0.49 precision, Naïve Bayes achieving 0.23 recall while SMO achieved 0.22, and both Naïve Bayes and SMO achieving a 0.31 F-measure value.

As shown for the confusion matrix for the SMO classifier in Table 9, there is still substantial mis-classifications with most posts being classified in the Anger class. The SMO classifier has 301 correctly classified posts and 8 misclassified posts in the Anger class. In contrast, all posts were misclassified in the Disgust class. There were 24 correctly classified posts and 124 misclassified posts in the Fear class, 12 correctly classified posts and 244 misclassified posts in the Happiness class, 36 correctly classified posts and 201 misclassified posts in the Sadness class, and 3 correctly classified posts and 226 misclassified posts in the Surprise class.

Table 9. Confusion matrix of Ekman's emotions classification for the IAEDS dataset using the SMO classifier.

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	**301**	1	1	4	0	2
Disgust	181	**0**	0	3	1	0
Fear	119	0	**24**	3	2	0
Happiness	238	3	0	**12**	2	1
Sadness	197	0	2	2	**36**	0
Surprise	223	0	0	3	0	**3**

Table 10 reports the confusion matrix of applying the Naïve Bayes classifier to classify Ekman's emotions in IAEDS. The only difference between the SMO and the Naïve Bayes confusion matrices is Naïve Bayes has 38 correctly classified posts for the Sadness class while SMO has 36 for the same class. This results in a slightly better recall value with 0.23 for Naïve Bayes compared with 0.22 for SMO.

The second experiment involved using the same Weka classifiers with the same `StringToWordVector` filter. This time the `CharacterNGramTokenizer` filter is used in order to find out whether this affects the classifiers' performance (and so that a character-based comparison can be made to the PPM classifier which is a character-based method rather than a word-based one). The results for this experiment showed that all classifiers achieved the same results as in the previous experiment for all measures.

fridaya@uef.fi

5.2 Applying the PPM Classifier

In this experiment, a PPM classifier [18] is used to classify Ekman's emotions in IAEDS. As far as we know, only Amer et al. [6,7] has used the PPM classifier to classify emotions for English text. The results for the PPM classifier is shown in Table 11. The table reports that the order 5 PPM classifier (PPMD5) has significantly outperformed all the other classifiers used in the previous two experiments in all measures. The PPM classifier achieved 86.9% accuracy, 0.63 precision, 0.59 recall, and 0.61 F-measure. Table 11 compares the PPM classifier result with the previous classifier results.

Table 10. Confusion matrix of Ekman's emotions classification for the IAEDS dataset using the Naïve Bayse classifier.

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	**301**	1	1	4	0	2
Disgust	181	**0**	0	3	1	0
Fear	119	0	**24**	3	2	0
Happiness	238	3	0	**12**	2	1
Sadness	197	0	0	2	**38**	0
Surprise	226	0	0	1	0	**2**

Table 11. Classification results using PPM classifier compared to classifiers supported by Weka for the IAEDS dataset.

Classifier	Accuracy	Precision	Recall	F-measure
J48	74.4	0.44	0.22	0.29
ZeroR	74.2	0.04	0.17	0.06
NaïveBayes	75.9	0.49	0.23	0.31
Multinomial NaïveBayes text	74.2	0.04	0.17	0.06
SMO	75.9	0.49	0.22	0.31
PPMD5	86.9	0.63	0.59	0.61

Analysing the confusion matrix shown in Table 12 for the PPMD5 classification provides an insight into why the PPM classifier outperformed the other classifiers. Here there is much less confusion concerning which posts should be in the Anger class. The PPM classifier has 241 correctly classified posts from a total of 309 posts in the Anger class, 76 correctly classified posts from 185 posts in the Disgust class, 81 correctly classified posts from 48 postsin the Fear class, 165 correctly classified posts from 256 posts in the Happiness class, 157 correctly

fridaya@uef.fi

classified posts from 238 posts in the Sadness class and 114 correctly classified posts from 229 posts in the Surprise class.

As stated, an order five PPM classifier (PPMD5) was used in this experiment. To find out which order of PPM model is most suitable for this type of classification for the Arabic language, a further experiment was used to check whether other orders (from order 2 up to order 12) give better results. Table 13 reports the results. Figure 6 also graphs the accuracy values shown in the table.

Table 12. Confusion matrix of Ekman's emotions classification for the IAEDS dataset using the PPMD5 classifier.

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	**241**	12	13	8	16	29
Disgust	64	**76**	5	5	9	26
Fear	30	6	**81**	7	9	15
Happiness	15	8	2	**165**	13	53
Sadness	40	10	13	8	**157**	10
Surprise	67	14	8	19	7	**114**

Table 13. Classification results of Ekman's emotions for the IAEDS dataset using different orders of the PPM classifier.

Classifier	Accuracy	Precision	Recall	F-measure
PPMD2	85.2	0.56	0.54	0.55
PPMD3	86.0	0.60	0.56	0.58
PPMD4	86.8	0.63	0.58	0.61
PPMD5	86.9	0.63	0.59	0.61
PPMD6	86.9	0.63	0.59	0.61
PPMD7	87.1	0.63	0.59	0.61
PPMD8	87.0	0.63	0.59	0.61
PPMD9	87.1	0.63	0.59	0.61
PPMD10	86.9	0.62	0.59	0.61
PPMD12	86.9	0.62	0.59	0.60

As is obvious from Fig. 6, the highest accuracy was achieved for orders 7 and 9 with value 87.1%, with order 2 producing the lowest accuracy of 85.2%. Figure 7 graphs the precision values achieved by the different order PPM classifiers. The figure shows that the highest precision value of 0.63 was achieved for orders 4, 5, 6, 7, 8 and 9. The figure also shows that the lowest precision value of 0.56 was achieved by the order 2 PPMD classifier.

fridaya@uef.fi

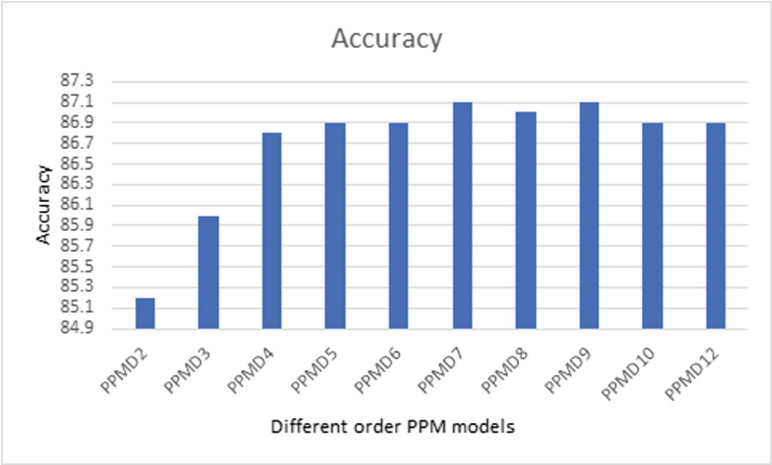


Fig. 6. Accuracies achieved by different orders of PPM.

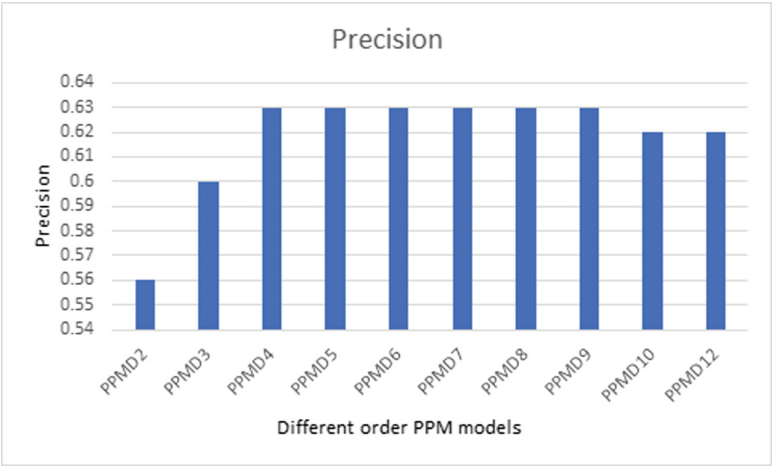


Fig. 7. Precision achieved by different orders of PPM.

Figure 8 graphs the recall values achieved by running the PPM classifier using different orders. The figure reports the highest recall value of 0.59 was achieved for orders 5 through 10 and 12 while the lowest recall value of 0.54 was achieved by PPMD2.

Figure 9 graphs the F-measures achieved by the different order PPM classifiers. The highest F-measure value of 0.61 was achieved by orders 4 through 10 whereas the lowest value of 0.55 was again achieved by PPMD2.

fridaya@uef.fi

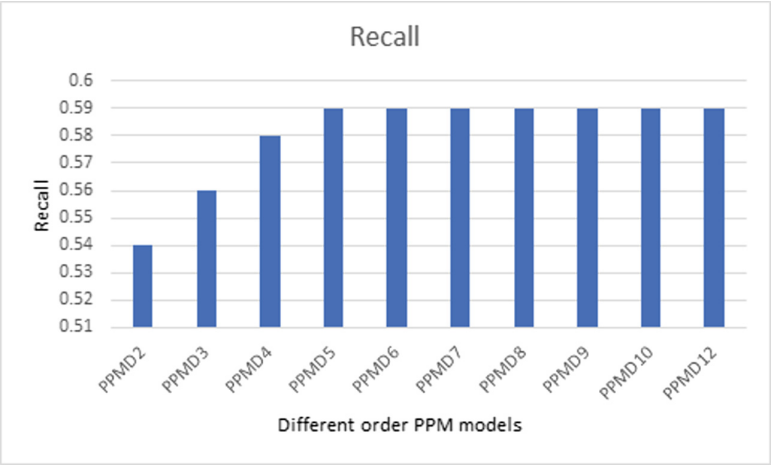


Fig. 8. Recall achieved by different orders of PPM.

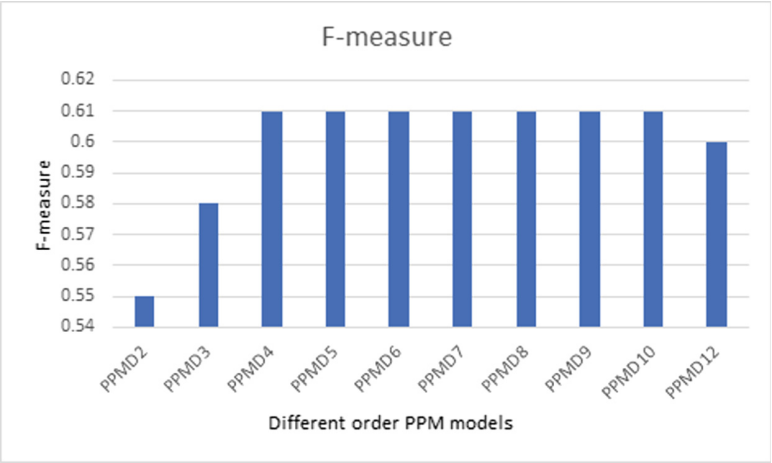


Fig. 9. F-measure achieved by different orders of PPM.

6 Conclusion

We have created a new emotion recognition dataset called IAEDS consisting of 1365 Iraqi Facebook posts labelled according to Ekman's six basic emotions. We performed several experiments using the IAEDS dataset. The first experiment was to test five classifiers supported by the Weka data analytic tool to classify Ekman's emotions from the IAEDS dataset. We found that the Naïve Bayes and the SMO classifiers were better than the J48, Multinomial Naïve Bayes for

fridaya@uef.fi

Text and ZeroR classifiers. The best performance was achieved by the Naïve Bayes and SMO classifiers with 75.9% accuracy, precision of 0.49, recall 0.23 and F-measure 0.31.

In the second experiment, we used the order 5 PPMD classifier based on the PPM character-based compression scheme to classify Ekman's emotions in IAEDS. Surprisingly, it significantly outperformed all the other classifiers in the first experiment with 86.9% accuracy, 0.63 precision, 0.59 recall, and 0.61 F-measure. Further experiments using PPM with different orders found that order 7 and 9 models achieved the highest accuracy of 87.1%.

Future work would benefit from collecting more data to train the classifiers supported by Weka and PPM to see whether classification performance can be improved further.

References

1. Abdul-Mageed, M., Diab, M.T.: AWATIF: a multi-genre corpus for modern standard arabic subjectivity and sentiment analysis. In: LREC, pp. 3907–3914. Citeseer (2012)
2. Abdulla, N.A., Ahmed, N.A., Shehab, M.A., Al-Ayyoub, M.: Arabic sentiment analysis: lexicon-based and corpus-based. In: 2013 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), pp. 1–6. IEEE (2013)
3. Ahmed, F., Nürnberger, A.: Evaluation of n-gram conflation approaches for arabic text retrieval. J. Am. Soc. Inform. Sci. Technol. **60**(7), 1448–1465 (2009)
4. Ahmed, F., Nürnberger, A.: A web statistics based conflation approach to improve arabic text retrieval. In: 2011 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 3–9. IEEE (2011)
5. Al-Aziz, A.M.A., Gheith, M., Eldin, A.S.: Lexicon based and multi-criteria decision making (MCDM) approach for detecting emotions from arabic microblog text. In: 2015 First International Conference on Arabic Computational Linguistics (ACLing), pp. 100–105. IEEE (2015)
6. Almahdawi, A., Teahan, W.J.: Emotion recognition in text using PPM. In: International Conference on Innovative Techniques and Applications of Artificial Intelligence, pp. 149–155. Springer (2017)
7. Almahdawi, A., Teahan, W.J.: Automatically recognizing emotions in text using prediction by partial matching (PPM) text compression method. In: International Conference on New Trends in Information and Communications Technology Applications, pp. 269–283. Springer (2018)
8. Aly, M., Atiya, A.: LABR: a large scale arabic book reviews dataset. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Volume 2: Short Papers, vol. 2, pp. 494–498 (2013)
9. Aman, S., Szpakowicz, S.: Identifying expressions of emotion in text. In: International Conference on Text, Speech and Dialogue, pp. 196–205. Springer (2007)
10. Azmi, A.M., Alzanin, S.M.: Aará-a system for mining the polarity of saudi public opinion through e-newspaper comments. J. Inform. Sci. **40**(3), 398–410 (2014)
11. Cohen, J.: A coefficient of agreement for nominal scales. Educ. Psychol. Measure. **20**(1), 37–46 (1960)
12. Ekman, P.: Facial expressions. Handb. Cogn. Emot. **16**(301), e320 (1999)

fridaya@uef.fi

13. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. ACM SIGKDD Explor. Newsl. **11**(1), 10–18 (2009)
14. Ku, L.-W., Lo, Y.-S., Chen, H.-H.: Test collection selection and gold standard generation for a multiply-annotated opinion corpus. In: Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions, pp. 89–92. Association for Computational Linguistics (2007)
15. Passonneau, R.: Measuring agreement on set-valued items (MASI) for semantic andpragmatic annotation. Lang. Resour. Eval. (2006)
16. Refaee, E., Rieser, V.: An arabic twitter corpus for subjectivity and sentiment analysis. In: LREC, pp. 2268–2273 (2014)
17. Salem, F.: Social media and the internet of things towards data-driven policymaking in the arab world: Potential, limits and concerns (2017)
18. Teahan, W.J., Harper, D.J.: Using compression-based language models for text categorization. In: Language Modeling for Information Retrieval, pp. 141–165. Springer (2003)

fridaya@uef.fi

Utilise Higher Modulation Formats with Heterogeneous Mobile Networks Increases Wireless Channel Transmission

Heba Haboobi(✉) and Mohammad R. Kadhum

Faculty of Arts, Science and Technology, University of Northampton, Northampton, UK
heba.haboobi@northampton.ac.uk

Abstract. In this paper, a higher modulation format with a heterogeneous mobile network (small cells, Macrocells) is proposed, explored and evaluated at a wireless transmission system. As such, study the effect of utilising developed schemes of modulation like the 256 Quadrature Amplitude Modulation (QAM) on the modulation/de-modulation level of the currently applied Orthogonal Frequency Division Multiplexing (OFDM). Since the higher bit-rate of transmission is one of the important topics for the forthcoming generation of mobile, the introduced system aims to regulate the trade-off relationship between the maximum achieved bit-rate and the minimum required level of the Signal-to-Noise Ratio (SNR). Hence, involve the small cell technology as a supportive tool for the higher schemes of modulation to increase the capacity of the channel at the accepted limit of error. Consequently, the presented system that combines both the higher modulation formats and the small cells can expand the transmission coverage with a higher bit-rate yet keeping a similar level of the received power. Moreover, the system performance in terms of the maximum bit-rate and the Bit Error Rate (BER) is investigated in the presence of the Additive White Gaussian Noise (AWGN) channel model. Also, the OFDM waveform is considered herein as an accommodating environment to examine the activity of the intended modulation techniques due to its' efficiency in using the available Bandwidth (BW). Furthermore, a MATLAB simulation is used to implement the promoted system and clarify the advantages and disadvantages of it in comparison with the currently applied 64 QAM.

Keywords: Orthogonal Frequency Division Multiplexing (OFDM) · Higher modulation format · Signal-to-Noise Ratio (SNR) · Bit Error Rate (BER) · Bandwidth (BW) · Small cells · Macrocells · Additive White Gaussian Noise (AWGN) · Bit-rate · Heterogonous mobile networks

1 Introduction

The increased growth of transmission data over mobile networks considers as an essential driving force for proposing higher modulation formats [1]. Hence, the

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 217–229, 2019.
https://doi.org/10.1007/978-3-030-22868-2_17

fridaya@uef.fi

growing demand for higher bit-rate is requesting a well exploiting for modulation formats in the utilised waveform. As such, fostering the required channel capacity of the data-hungry applications can be achieved by developing the applied transmission methods.

Consequently, many research studies have been induced to support a higher channel capacity of future market demand [2]. As such, achieve developed rates of channel capacity at accepted limits of errors get overwhelming interest in terms of R&D. Thus, recently, great efforts have been made to promote both bit-rate and Bandwidth (BW) efficiency of a transmitted signal by extensively employing higher modulation formats [3].

The diversity of the transmitted signals among users leads to expand the usage of the applied broadband where the mechanism of data transmission is developed from only working with emails to multimedia signals. Hence, nowadays, using smart phones, users are looking for downloading and uploading huge amounts of data rather than the conventional operations [4].

It's worth noting that, picking a suitable modulation format for future wireless networks of mobile depends on the intended performance in terms of the bit-rate, Bit Error Rate (BER) and the coverage of the transmission system. As such, compared to the lower schemes of modulation, the higher modulation formats significantly improve the BW efficiency and maximum bit-rate, whereas, the quality of the transmission service is typically measured by both the BER and the maximum gained rate of bits [5].

In addition, factors like the received power and probability of constellation table noise are major limitations for developers in the modulation format field of wireless communication [6]. Hence, such restrictions can affect the efficiency of spectrum and slow down the speed of transmitting information in an obtainable BW [7].

The main challenge of the future mobile networks is achieving a higher transmission rate for the mobile signals with lower levels of the BER. Hence, the possibility to employ the modulation process to promote the maximum rates of transmission yet keeping the minimum level of the Signal-to-Noise Ratio (SNR) at the accepted BER limit [8].

Very recently, the digital modulation is considered as one of the most significant issues in the wireless communications world. Hence, utilising an appropriate type of a modulation format for an advanced wireless transmission is a critical due to key limitations of the available transceiver power. As such, the performance in terms of maximum bit-rate and minimum BER of a modulation scheme depends on the efficiency of the BW and power usage [9].

Moreover, the predictably increased data rates for future mobile applications make the currently utilised large macrocells (base station towers) unable to provide efficiently higher bit-rate for the modern telecommunications [10]. Hence, the need for seeking a new approach is quite necessary.

As a result, the potentiality of deploying the small cells technology in the future wireless communication of mobile is highly encouraged [11]. Hence, introducing the small cells as an alternative or supportive technique to the current

fridaya@uef.fi

networks (macrocells) can play a more important role in the modern mobile networks. Thus, the expected topologies of the future networks can be at most as a heterogeneous network, where a mix of large and small cells are employed with variant sizes and power levels [12].

This scalable solution, however, comes up with increasing the computational complexity due to expanding the required number of utilised cells.

The key question of this paper is, can the higher modulation formats in cooperative with the heterogeneous networks be a good solution for a developed performance of future wireless mobile technology?

In this paper, the performance (bit-rate, BER) for the advanced digital modulation schemes, particularly, 128 and 256 QAM is investigated for the first time in diverse scales of wireless mobile networks. Hence, explore how the developed modulation formats can affect the performance of transmission in terms of channel capacity and the BER for the heterogeneous networks that combine both the macrocells and microcells.

Furthermore, the newly introduced modulation formats are examined utilising the most popular transmission air interface, Orthogonal Frequency Division Multiplexing (OFDM). Due to the orthogonality, the current waveform of wireless technology can deliver a better level of transmission service comparing with the old fashion Frequency Division Multiplexing (FDM) [13].

In addition, the channel model, Additive White Gaussian Noise (AWGN), is basically considered herein to explain how this uniform noise can degrade the performance of the transmission system [14].

The rest of the paper is organized as follows: Sect. 2 discusses theoretically the main concepts of the presented system, highlights the fundamentals beyond it physically and mathematically. Section 3 numerically simulates the performance of the system in terms of the channel capacity and BER. Section 4 summarises the outlines of the paper.

2 System Model

In this part, developed formats of modulation/de-modulation process are introduced. The proposed scheme can improve the capacity of the upcoming wireless channel in comparison with the conventional modulation formats. Thus, the advanced kinds of modulation which are higher than the existing 64 QAM are employed to promote the throughput of the transmission for the future mobile communication system (5G). As such, utilising higher modulation schemes like the 256 QAM, each sub-carrier in the forthcoming networks is eligible to carry 8 bits instead of 6 bits.

In addition, the introduced design is studied in terms of the maximum bit-rate and required SNR to explore how the newly proposed modulation formats can impact the performance of transmission in relative to the current wireless system.

As is shown in Fig. 1, the main environment of the transceiver at the physical layer is described. Thus, clarifying the relationship between the developed

fridaya@uef.fi

modulation formats and the other employed parts in the electrical back-to-back wireless transmission system. Hence, the basic phases involving the modulation/de-modulation operations of the current wireless mobile waveform (OFDM) are demonstrated herein.

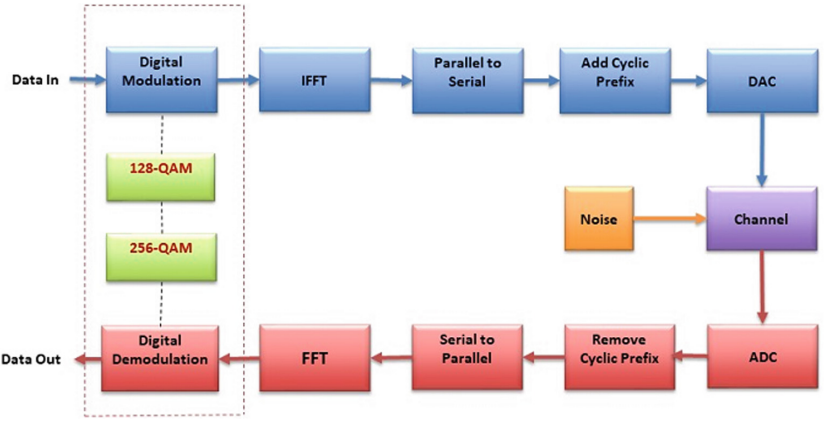


Fig. 1. Proposed design for a future wireless transmission system with higher modulation schemes (128 QAM & 256 QAM).

Regarding the first level of transmission part (modulation stage), the new design utilises a higher modulation format (128 QAM & 256 QAM) to convert each token of binary digits (0,1) to its' frequency domain shape.

Since the higher schemes of modulation are used, the supposed size for every token is larger than the currently available length (6 bits). As such, for a stream of binary data, it's much better, in terms of BW efficacy to carry 7 or 8 bits instead of 6 bits as a maximum.

Based on the input stream of digits, the modulation procedure fundamentally comes up with introducing an equivalent set of complex numbers. The new shape of data can work in the frequency domain, hence, specify each complex number to one of the available sub-carriers.

After that, employing the Inverse Fast Fourier Transform (IFFT), the complex number is converted to a new shape in the time domain which is formally called a sample. It's worth noting that, each a fixed group of samples (depending on the number of really offered sub-carriers) termed as a symbol.

In the time domain, where the symbol is prepared for transmission, an important part is attached to the original symbol (CP) providing an extra time protection. Thus, employing an optimal CP, the effect of probably occurred interference among symbols (ISI) is avoided.

Utilising an appropriate F_{DAC}, with a recommended high sampling frequency, the digital symbol is transferred to the analog domain. As such, the analog signal is ready now for transmitting by an antenna.

fridaya@uef.fi

During the transmission, the broadcasted signal may suffer from variant types of attenuation, which, as a result, can influence the performance of transmission in terms of the BER and maximum gained bit-rate. Herein, the impact of the AWGN channel is considered for exploring the transmitted signal under unwanted conditions.

On the other hand, at the receiver part, an opposite treatment is performed to retrieve the original signal. Thus, the received signal is converted back to digital domain using a similar sampler F_{ADC}.

Then after, the supplementary CP is removed from each symbol, producing the abstract symbol at the time domain. The processed symbol is then changed to the frequency domain using the Fast Fourier Transform (FFT).

The complex numbers, at this domain, are checked to ensure whether their phases and magnitudes were impacted by the channel response or not. The affected complex numbers are addressed utilising the Equalization process which is essentially applied to correct any probable change in both the phase and amplitude.

After the correction operation is completed, a developed de-modulation process is ultimately performed employing higher modulation formats (128 QAM & 256 QAM), to recover the original stream of binary digits.

As the modulation/de-modulation process represents a key stage of the proposed system, it is quite important to discuss, from a mathematical perspective, some related concepts that can affect directly the transmission operation.

In the transmitter side, particularly, in the modulation process, each token of the binary digits is converted to a complex number which is expressed in Cartesian forms as follows [15]:

$$C_k = I_k + jQ_k \tag{1}$$

where $j = \sqrt{-1}$, and I, Q represents the real and imaginary parts for k^{th} complex number respectively.

In this context, the mathematical relationship between the amplitude (A) and a produced complex number is clarified as follows [16]:

$$A_k = \sqrt{I_k^2 + Q_k^2} \tag{2}$$

Regarding the probable shapes number of complex numbers on the constellation map, the following equation achieves this purpose [17]:

$$Y = 2^X \tag{3}$$

where Y represents the total number of probabilities which might be assigned for each sample employing X of input bits.

Despite each spectrum of sub-carrier can coincide with the others, the ability to extract each sub-carrier is achieved over the digital signal processing. Hence, this overlapped property of sub-carriers increases the spectrum efficiency of the current OFDM in comparative with the previous multi-carrier design of

fridaya@uef.fi

waveform. Thus, the OFDM technique splits a wireless channel into smaller sub-carriers each one is modulated with an amount of data according to the applied modulation format [18].

The improved efficiency of the OFDM spectrum is gained due to applying the orthogonality between adjacent sub-carriers. Consequently, obtain a larger benefit for the same offered BW.

According to Shannon's theorem, the channel capacity represents the maximum achieved bit-rate with a vanished amount of errors as follows [19]:

$$Capacity = BW.\log(1 + SNR) \tag{4}$$

In addition, the required SNR to receive the signal is calculated as follows [19]:

$$SNR = P_{ReceivedSignal}/P_{Noise} \tag{5}$$

In this study, to improve the obtained channel capacity, an extra investigation is performed in the field of BW efficiency than going with expanding the offered BW itself since such expanded for resources requires a highly increased cost.

Besides, the modulation operation is explored in the presence of the AWGN channel. Hence, the received signal R that is composed of the transmitted signal T multiplied by the response of the channel E, where $E = 1$, is combined with the AWGN which is represented here as N, as follows [19]:

$$R = ET + N \tag{6}$$

3 Results and Discussion

In this part, the promoted wireless transmission system is numerically simulated to demonstrate the advantages and disadvantages of proposing higher modulation formats for the future mobile networks. Hence, the performance in terms of maximum bit-rate and the BER for the OFDM based 128 QAM and 256 QAM is explored.

In addition, the test is set up herein for 15 frequencies of sub-carriers where the developed modulation schemes are considered side by side with the conventional modulation formats. Besides, the optimally utilised parameters of the configured model are undertaken for a wireless electrical back-to-back system. Moreover, to examine the behaviour of wireless channel responses over variant modulation schemes, the MATLAB simulation is equipped.

Since the modulation system is the core of the proposed system, the number of specified bits for each applied sub-carrier is accurately fixed according to the corresponding level of the utilised modulation format. The expanded performance of transmission is achieved under the conditions shown in Table 1.

Regarding the relationship between gained channel capacity and the utilised modulation techniques, as is shown in Fig. 2, the newly proposed modulation formats, 128, and 256 QAM can increase the transmission bit-rate in compared with the currently applied 64 QAM to about 16% and 33% respectively.

fridaya@uef.fi

Table 1. System parameters for the OFDM

Parameter	Value
Modulation format	BPSK - 256 QAM
FFT size	40
FFT sampling frequency	20 MHz
Number of sub-carriers	15
Cyclic prefix	0.25

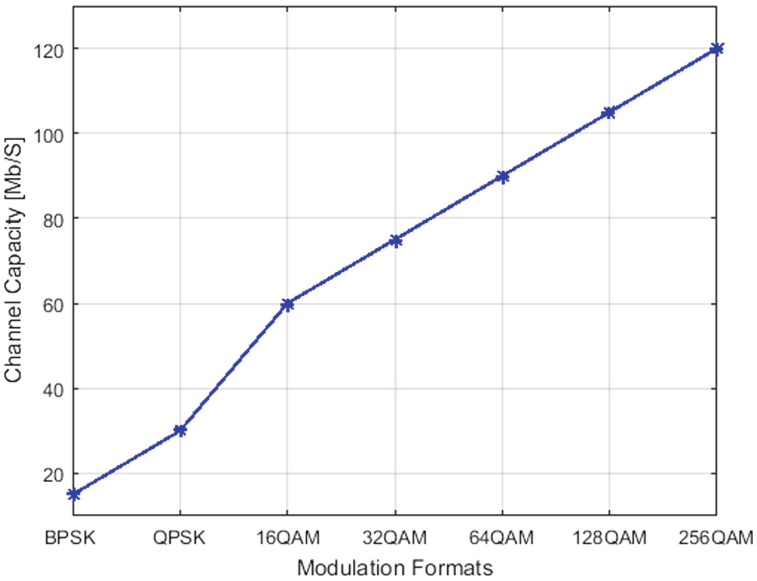


Fig. 2. Maximum achieved bits rates of different modulation formats including the 128 and 256 QAM.

Besides, in relative to the 32 QAM and 16 QAM, the higher modulation 256 QAM improves the overall channel capacity by around 60% and 100% sequentially. In addition, 3 and 7 times enhancement are recorded for both QPSK and BPSK respectively in related to the higher order modulation (256 QAM).

This, however, comes up with raising the level of errors for transmitted samples due to increased interference at the constellation map. Hence, a higher BER is introduced for each employed sub-carrier. Thus, raising a minimum limit of the received power at 10^{-3} BER.

The simulated work demonstrates, also, how the BER of the utilised sub-carriers are varied for diverse modulation schemes under a similar level of the SNR and over AWGN channel.

fridaya@uef.fi

As is seen in Fig. 3, the BERs of the applied frequencies of sub-carriers for the currently employed modulation format (64 QAM) are calculated. The measured BERs are achieved with a good level of the SNR which is equivalent here to 23 dB. Thus, the accepted limit of error is gained due to supplying a suitable standard of received power for the signal modulated with 64 QAM.

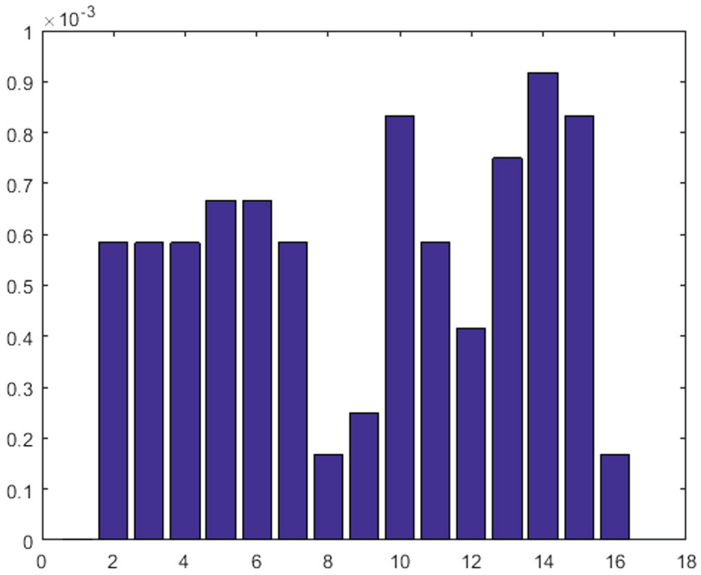


Fig. 3. BER bars of utilised sub-carriers with 64 QAM and SNR = 23.

This scenario, nevertheless, is not usually fit for all other modulation formats, particularly, higher order schemes of modulation like 256 QAM.

As is clear in Fig. 4, an extra rise of the BER limit is registered with additional expansion of the modulation format to get the 256 QAM and keep the SNR at the 23 dB. This, essentially, due to decrease the distances between the adjacent samples of the enlarged modulation schemes on the constellation table resulting in un-ability in recognizing the received signal.

To mitigate this issue, the SNR is gradually raised to be more suitable for the 256 QAM. Hence, whenever the received power of the signal is increased the BER is decreased until achieve the accepted criteria of tested modulation scheme.

As is shown in Fig. 5, the overall system performance in terms of the BER and SNR for the most common modulation format with the AWGN channel is presented. Hence, the trade-off relation between the gained BER and offered SNR is digitally processed for variant modulation schemes (QAM and PSK). Thus, compare the transmission performance of conventional modulation techniques as the BPSK, QPSK, 16 QAM, etc, with advanced configurations of modulation (128 and 256 QAM) over the AWGN channel.

fridaya@uef.fi

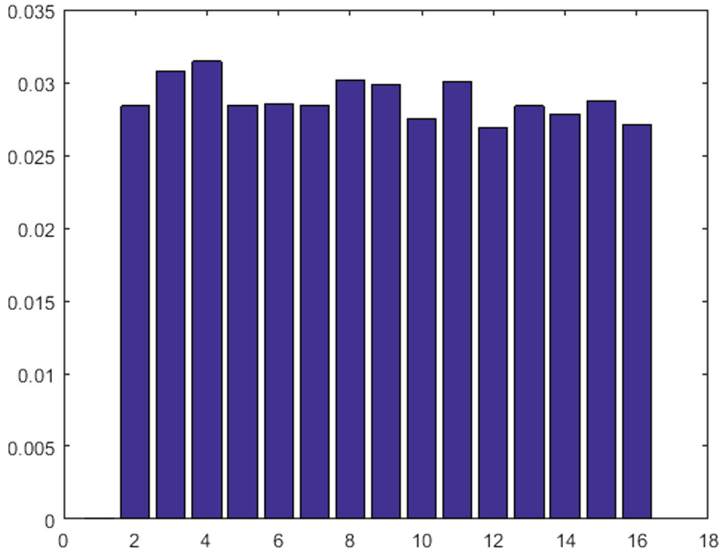


Fig. 4. BER bars of utilised sub-carriers with 256 QAM, SNR = 23.

It's worth noting that, among these 7 investigated modulation formats, the BPSK can be applied with a lower power requirement, while the 256 QAM requires a higher level of power due to the increment of bits number for each transmitted sample.

As a result, with the future generation of mobile, the higher modulation schemes are very recommended due to the increased need for a higher channel capacity.

Practically speaking, apply higher modulation schemes for an expanded coverage of the signal requires either raising the power of the receiver (higher SNR) or increasing the number of repeater towers (power injection).

Regarding the first option, raise the required power for receiving the signal with a higher modulation scheme can come up with increasing the cost of transmission (more expensive base stations).

As regards the second option, increase the number of repeaters is quite useful if and only if the kind of applied repeater can be offered with better attributes than the current big cells. Thus, the small cell with its' valuable features like simple installation, small size, and low cost can be the best option for this case. Hence, in the case of the higher modulation formats, the small cells repeaters (Femtocells, Picocells, Microcells) can be utilised with close distances to secure a larger transmission coverage.

However, this mix of large and small cells (heterogeneous network) comes up with raising the level of complexity due to the increased number of employed cells.

fridaya@uef.fi

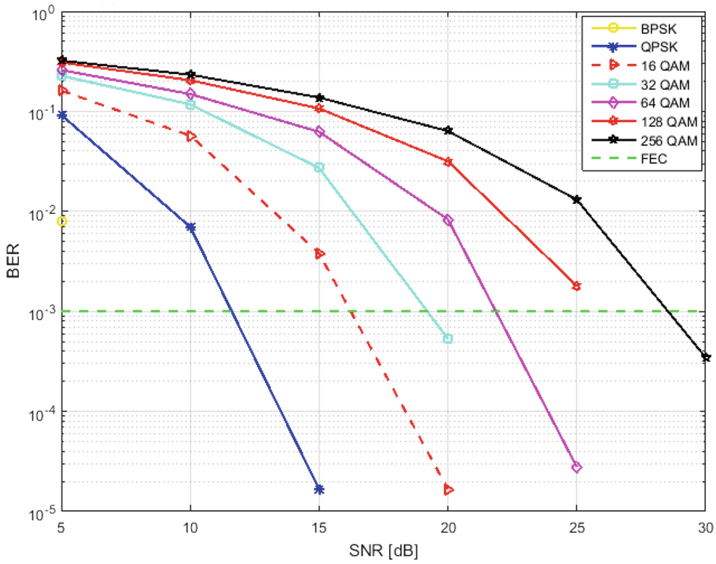


Fig. 5. System performance in terms of BER and SNR for most common modulation formats including both 128 and 256 QAM over AWGN channel.

To explain more about this, it is clear from Fig. 6 that, different levels of the SNR are required to achieve a similar coverage ($D1$) in a wireless transmission system with variant shapes of modulation, where $D1$ represents the supposed distance for transmitting a modulation scheme with a minimum required level of the SNR and accepted limit of the BER (10^{-3}).

It is noting that, the higher modulation formats suffer from a bigger amount of power consumption due to a raised level of the constellation map noise in comparison to the lower formats. As such, the relatively minor amounts of attenuation might significantly affect the signal with higher modulation order due to close space between samples on the table.

For instants, the minimum limit of received power for the higher order modulation (256 QAM) is larger than the lower order (BPSK) by about four times. Thus, the transmitted signal cannot be received unless this minimum level of power is achieved to ensure accepted limits of the BER.

As is illustrated in Fig. 7, the low modulation schemes can utilise this feature and travel farther distance ($D2$) with a good BER in comparison to the high modulation since the obtainable level of power is still supportive of for the low modulation. Hence, $D2 = D1 * M1$, where $M1$ denotes enhancement times of transmission distance due to improved level of the offered SNR with a transmission system that can basically support high modulation formats. This, however, reduces the bit rate of transmission to a minimum level to keep it constantly at accepted limits of errors.

fridaya@uef.fi

Fig. 6. Different modulation formats with different SNR levels at a wireless transceiver system.

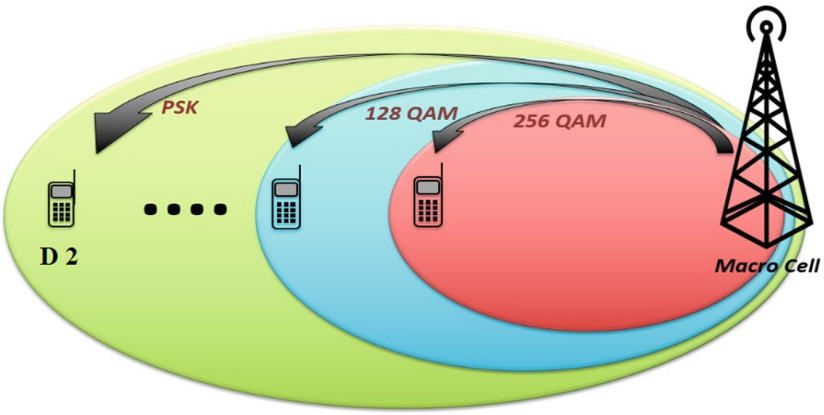


Fig. 7. Macrocell with similar SNR and different modulation formats involving the proposed 128 and 256 QAM.

On the other hand, as is seen in Fig. 8, after proposing the small cells for modern mobile networks, the range of transmission ($D1$) is quite expanded to ($D2$) for the higher modulation scheme (256 QAM) with a lower required limit of the SNR (29 dB). Hence, $D2 = D1 * M2$, where $M2$, denotes enhancement times of transmission distance due to employing the small cells. Thus, by employing both the high order modulation schemes and the small cells technology, the transmitted signal can move to a longer distance with higher channel capacity and acceptable limits the BER.

This is essentially occurred, since, the required levels of power are continuously offered by increasing the number of the utilised cells to sustain the wireless signal every time and everywhere depending on the small cells deployment.

fridaya@uef.fi

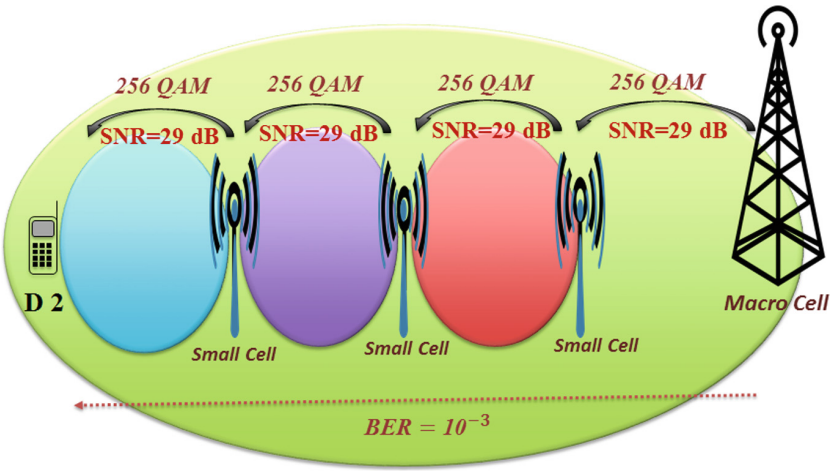


Fig. 8. Utilise both small cells and higher modulation format in one transmission system.

4 Conclusion

In this study, the higher modulation scheme (256 QAM) with the heterogeneous mobile network is introduced, investigated and evaluated in the PHY of an electrical back-to-back wireless transceiver system. As the higher channel capacity is one of the significant issues for the future generation of mobile, the currently applied macrocells are no longer be accepted with the higher schemes of modulation. The main reason beyond this is that the conventional network (big cell) which tests a higher modulation shape tends to increase the channel capacity with limited coverage of transmission due to the higher power consumption. Hence, according to the obtained results, about 4 times higher SNR (29 dB) is required with the 256 QAM to achieve the same short coverage of the BPSK in the macrocell networks. To address this issue, the high order modulation based small cells technique is proposed for the upcoming mobile technology (5G) improving both the BW efficiency and transmission coverage. The findings show that by employing the higher modulation scheme and the microcells together, the wireless signal can travel a long distance with a higher channel capacity and an acceptable BER limit yet keeping the same level of the SNR (29 dB). As such, regulate the trade-off relationship between the maximum achieved bit-rate and the minimum required level of SNR. Thus, the small cell with its' valuable features as simple installation, small size, and low cost can be the best option for the modern mobile networks. Nevertheless, this mix of large and small cells (heterogeneous network) comes up with raising the level of the computational complexity due to the increased number of employed cells.

fridaya@uef.fi

References

1. Nagarajan, K., Kumar, V.V., Sophia, S.: Analysis of OFDM systems for high bandwidth application, pp. 168–171 (2017)
2. Kadhum, M.R., Kanakis, T., Al-Sherbaz, A., Crockett, R.: Digital chunk processing with orthogonal GFDM doubles wireless channel capacity. 1–6 (2018). https://doi.org/10.1007/978-3-030-01177-2-53
3. Ndujiuba, C.U., Oni, O., Ibhaze, A.E.: Comparative analysis of digital modulation techniques in LTE 4G systems. J. Wirel. Netw. Commun. **5**, 60–66 (2015)
4. Barnela, M.: Digital modulation schemes employed in wireless communication: a literature review. Int. J. Wired Wireless Commun. **2**, 15–21 (2014)
5. Kadhum, M.R.: New Multi-Carrier Candidate Waveform For the 5G Physical Layer of Wireless Mobile Networks
6. Ghogho, M., McLernon, D., Alameda-Hernandez, E., Swami, A.: Channel estimation and symbol detection for block transmission using data-dependent superimposed training. IEEE Signal Process. Lett. **12**, 226–229 (2005)
7. Kadhum, M.R., Kanakis, T., Crockett, R.: Intra-channel interference avoidance with the OGFDM boosts channel capacity of future wireless mobile communication. In: Procedings of Computing Conference 2019, London (2019)
8. Chandran, I., Reddy, K.A.: Comparative analysis of various channel estimations under different modulation schemes, vol. 1, pp. 832–837 (2017)
9. Kadhum, M.R., Kanakis, T., Crockett, R.: Dynamic bit loading with the OGFDM waveform maximises bit-rate of future mobile communications. In: Procedings of Computing Conference 2019, London (2019)
10. Jiang, Z., Mao, S.: Energy delay tradeoff in cloud offloading for multi-core mobile devices. IEEE Access **3**, 2306–2316 (2015)
11. Reed, M.C., Wang, H., Reed, M.C., Wang, H.: Small cell deployments: system scenarios, performance, and analysis (2018)
12. 3Gpp. The Mobile Broadband Standard. 9, 1–6 (2014)
13. Jin, W., et al.: Improved performance robustness of DSP-enabled flexible ROADMs free from optical filters and O-E-O conversions. J. Opt. Commun. Netw. **8**, 521 (2016)
14. Haboobi, H., Kadhum, M.R.: Impact study and evaluation of higher modulation schemes on physical layer of upcoming wireless mobile networks. **10** (2019)
15. Ingle, V.K., Proakis, J.G.: Digital Signal Processing Using MATLAB (2012)
16. Tao, L., et al.: Experimental demonstration of 10 Gb/s multi-level carrier-less amplitude and phase modulation for short range optical communication systems. Opt. Express **21**, 6459 (2013)
17. Alexander, E., Poularikas, D.: The Handbook of Formulas and Tables for Signal Processing (1999)
18. Stern, S., Fischer, R.F.H.: OFDM vs single-carrier modulation: a new view on the PAR behavior, pp. 112–119 (2014)
19. Im, G.H., et al.: 51.84 Mb/s 16-CAP ATM LAN standard. IEEE J. Sel. Areas Commun. **13**, 620–632 (1995)

fridaya@uef.fi

Performance Analysis of Square-Shaped and Star-Shaped Hierarchically Modulated Signals Generated Using All-Optical Techniques

Anisa Qasim and Salman Ghafoor(✉)

School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad, Pakistan
salman.ghafoor@seecs.edu.pk
http://seecs.nust.edu.pk/faculty/drsalmanghafoor.html

Abstract. Millimeter-wave communications will play an important role in realizing the high-data rates of future 5G systems. In order to prioritize the data for different users, hierarchical modulation offers simultaneous transmission of multiple independent signals over a single electrical carrier. All-optical transmission and generation of millimeter-wave signals is a suitable solution to cope with the large number of base stations and high data rates required for future high frequency 5G networks. In this work, we have implemented an all-optical technique for the transmission and subsequent generation of hierarchically modulated millimeter-wave signals. We have compared the performance of square-shaped and star-shaped hierarchically modulated signals in terms of Bit-error-rate and power margin. It was observed that optical single sideband modulation is a suitable technique to overcome high frequency related impairments in optical fibers. Also, the high priority layer of a hierarchically modulated signal performs better than the low priority layer. Furthermore, star-shaped signal performs better than the square-shaped hierarchically modulated signal.

Keywords: Hierarchical modulation · Radio over fiber · Millimeter-waves · Optical communications

1 Introduction

Video streaming and multimedia services that are being run on smart devices demand high data rates and therefore, large bandwidths from the wireless access networks (WANs). For the next generation WANs, use of congestion-free millimeter-wave (mm-wave) spectrum and deployment of distributed antenna system (DAS) architectures are the two major solutions that can increase the capacity of the access network to the multi-gigabit per second range [1].

Figure 1 shows a coordinated multi-point (CoMP) architecture, where all the signal processing and resource management tasks of the radio access units

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 230–241, 2019.
https://doi.org/10.1007/978-3-030-22868-2_18

fridaya@uef.fi

(RAUs) can jointly be performed at the central unit (CU) [2]. In a DAS architecture, the fiber lengths from the CU to RAUs are variable, resulting in different transmission losses for each RAU. Therefore, different RAUs have different power margins for the same optical transmit power [3]. Power margin is defined as the difference between the received optical power and the required optical power to obtain an acceptable bit-error-rate (BER) performance. Similarly, the condition of the wireless link for each mobile user is different due to two main reasons that are different channel frequencies and different wireless link lengths [4]. To implement the generation and transmission of mm-wave signals over a DAS architecture, radio over fiber (RoF) is considered a promising solution due to its cost-efficiency and large bandwidth [1]. RoF offers the benefits of low loss, large bandwidth, transparency to different modulation schemes and cost-efficient deployment of RAUs [5,6].

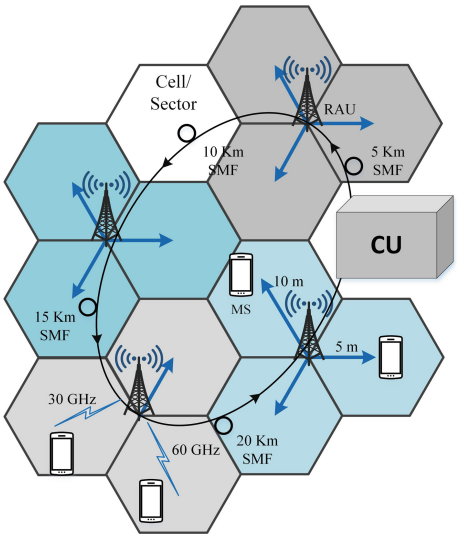


Fig. 1. RoF-powered coordinated multipoint (CoMP) architecture.

Hierarchical modulation (HM) is a layered modulation scheme with backward compatibility. For the fixed data rate systems, backward compatibility is a major issue. The user of already deployed system receives data of fixed modulation and bit rate. Therefore, it is cost inefficient to replace all these fixed rate receivers [7]. In HM, the already deployed receivers need not to be replaced. However, new receivers that are capable of receiving hierarchically modulated signals are required. The already deployed conventional receivers are still able to extract their information from the basic layer of the hierarchical signal. This improves system throughput without requiring extra bandwidth as compared to conventional modulation formats. The higher flexibility of HM allows major upgrade in the wireless communication network while maintaining low complexity [8]. Such advance modulation schemes help in minimizing the component count in RoF systems [9].

fridaya@uef.fi

HM was implemented experimentally for orthogonal frequency division multiple access (OFDMA) passive optical network (PON), showing improvement in power margin [10]. In [10], the optical network unit (ONU) located at a shorter distance is allocated lesser power compared to the ONU that is located at a longer distance. This resulted in an overall power margin improvement for the OFDMA-PON architecture. Power margin improvement for mm-wave signals carrying 16-quadrature amplitude modulation (QAM) data on OFDM carriers has been experimentally demonstrated in [11]. In this study, some of the power of the wireless channel having a shorter distance is transferred to the wireless channel having a longer distance. To quantify power margin improvement, both of these studies compared a conventional 16-QAM modulation with their proposed square QAM HM. In [12], some application examples of optical access network towards future digital PON are introduced. The study discusses higher efficiency of hierarchical PON compared to conventional PON in terms of transmitted and received power and loss budget. HM for single-carrier frequency division multiplexed system is proposed and experimentally demonstrated in [13]. It was shown that the hierarchically modulated system showed lesser power penalty compared to a conventional system.

Based on the previous discussion, the novel contributions of our work are summarized below:

- We propose various HM schemes to improve the power margin of the RoF systems. By employing star-QAM in combination with square-QAM HM, mm-wave signals for different users are mapped onto the basic and secondary layers of the electrical signal. It is shown that by employing this technique, power margins of the users receiving basic layer data are improved while they are degraded for the users receiving secondary data.
- The HM RF signals are transmitted by employing optical single sideband with carrier modulation to avoid dispersion induced power fading. High frequency HM signals are generated at the RAU through remote heterodyne detection between the carrier and the single sideband.

We have performed a simulation study by employing the commercial tool known as OptSim.

2 Hierarchical Modulation: Operating Principle

HM is a layered modulation, which was proposed to transmit different class of data to receivers having different reception conditions [14]. A receiver with suitable conditions such as close to the transmitter, having high antenna gain or an updated receiver, can access both layers of hierarchical data. Whereas, the receiver with poor reception conditions or legacy receivers only receive high priority information and are mostly employed in broadcast systems. HM based communication systems offer backward compatibility, which means that the updated receivers can accept both layers simultaneously while the legacy receivers receive only the basic layer [8]. In conventional QAM modulation, phase and quadrature

fridaya@uef.fi

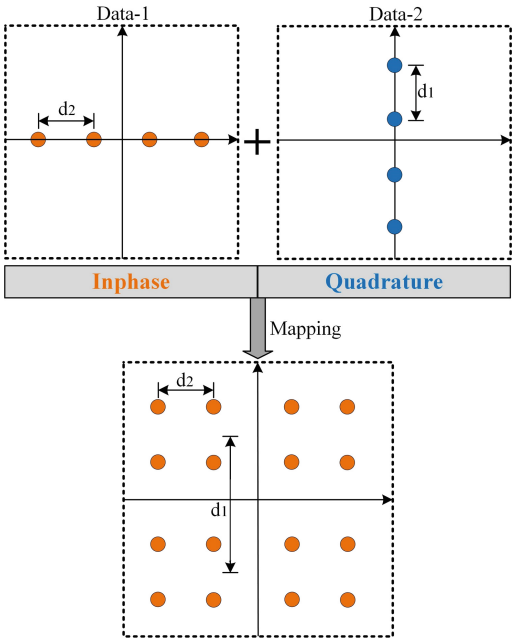


Fig. 2. Mapping of conventional 16-QAM

components are combined to generate a single data stream, as shown in Fig. 2. In HM, the low priority data is mapped on the high priority data by taking the constellation points of high priority data as a center point to LP layer constellation, as shown in Fig. 2. In this way, LP layer is mapped on each constellation point of HP layer. And both the data streams multiplex to form a single stream. The minimum distance between two consecutive points in high priority and low priority layers of HM are d_1 and d_2, respectively [11]. Since d_1 is greater than d_2, the BER performance of high priority data is better than that of low priority data in the same system. In this way we can achieve two different BER performances on a single data stream, also known as unequal error protection in broadcast systems [10].

The hierarchical parameter "α" is defined as the ratio of d_1 and d_2. Its value defines the power distribution in both layers. In uniform distribution where $(d_1 = d_2)$, HM adopts the exact conventional Square QAM shape, resulting in equal power distribution between high priority and low priority layer. With the increasing value of alpha, the distribution becomes non-uniform and power resides more in high priority (HP) layer as compared to low priory (LP) layer [15]. In this paper, we study two types of QAM constellations that are star QAM and square QAM. Figure 3 shows the conceptual diagram of star and square hierarchical QAM. We will consider different cases of constellations with different values of α. As the value of α increases, the low priority constellation squeezes and moves toward outer corners of the QAM constellation.

fridaya@uef.fi

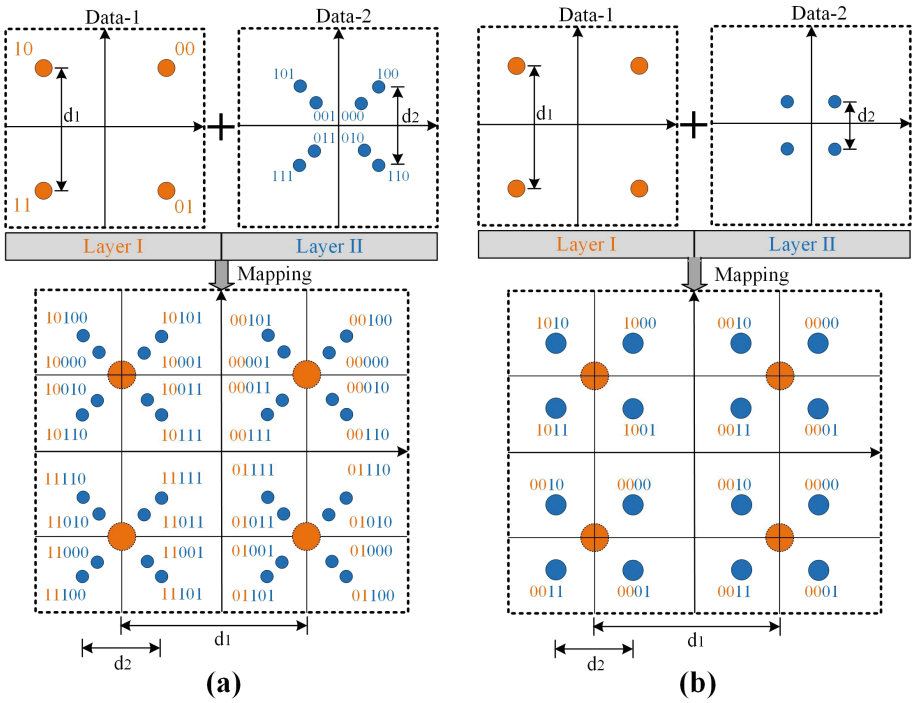


Fig. 3. Conceptual diagram of (a) Star hierarchical QAM modulation (b) Square hierarchical QAM modulation

Figure 4 shows the mapping of square QAM hierarchical modulation for $\alpha = 2$ and $\alpha = 4$ of HM square QAM. The points of low priority signal form a cloud in each of the four quadrants. For the originally designed conventional receivers, these clouds represent a single point in each quadrant of a QPSK signal. Therefore, the upgraded system after adding low priority signal to it, is still compatible with conventional/legacy receivers [16]. Legacy receivers receive the high priority bits but with the only difference that it may work with a slightly higher noise level. The additional noise due to low priority signal addition causes performance penalty in legacy receivers. Upgraded receivers of hierarchical environment have the ability to distinguish each point in the cloud. They can extract both the high priority as well as the low priority signal from the single data stream [7]. The penalty for this hierarchical system is expressed in terms of α. BER variation is also a function of α. When α increases, the constellation points in each quadrant/cloud shrinks and it is easy to extract high priority data with less BER and vice versa. As the alpha increases the high priority and low priority layers become more and less immune to noise, respectively. In this way, BER decreases and increases with increasing value of α for high priority and low priority layer, respectively. Therefore, α can be used to handle the tradeoff between the power penalties and robustness in HP and LP layer [10,11].

fridaya@uef.fi

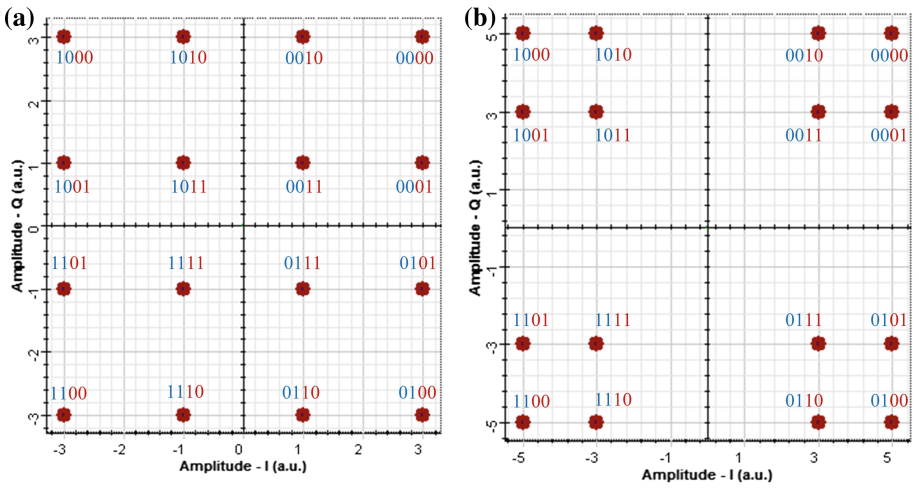


Fig. 4. Mapping of square QAM hierarchical modulation simulation for (a) $\alpha = 2$ and (b) $\alpha = 4$

3 The Proposed Architecture

Figure 5 shows the simulation setup for the proposed COMP architecture shown in Fig. 1. The optical field of the continuous-wave (CW) laser can be written as:

$$E_{laser} = \sqrt{2P_{laser}} \exp^{j\omega_c t} \tag{1}$$

Here P_{laser} and ω_c are the optical power and frequency of the laser, respectively. The output spectrum of the laser diode can be seen in Fig. 5(a). The output of the laser is split into two parts using a 3 dB optical splitter (OS) to get:

$$E_1(t) = E_2(t) = \sqrt{\frac{I_{OS}}{2}} E_{laser} \tag{2}$$

Here I_{OS} is the insertion loss of the splitter. The output electric field $E_1(t)$ is fed to the Single-Drive Mach-Zehnder Modulator (SD-MZM). The MZM is driven by the hierarchically modulated electrical drive signal $X_{HM}(t)$. The electric field of the optical signal at the output of the MZM $E_{MZM}(t)$ may be written as:

$$E_{MZM}(t) = \frac{E_1(t)}{2} \left[e^{j\pi\left(\frac{V_{bias}}{V_\pi} + \frac{X_{HM}(t)}{V_\pi}\right)} + 1 \right] \tag{3}$$

fridaya@uef.fi

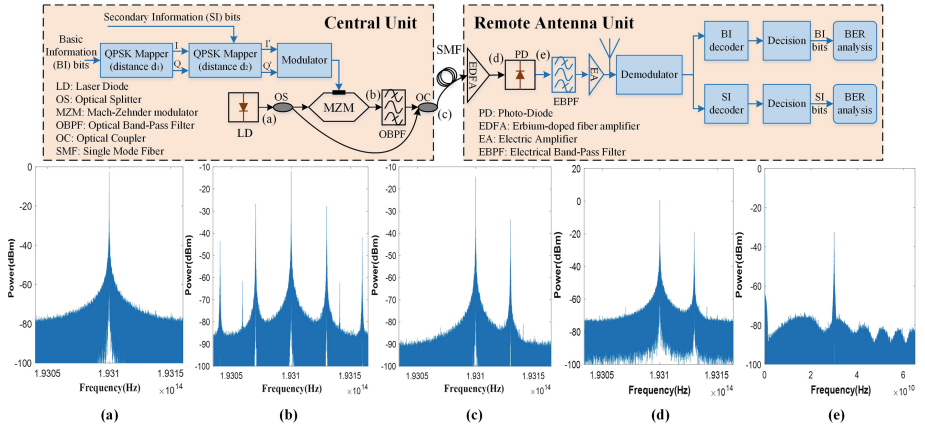


Fig. 5. Proposed setup for the hierarchical modulation based CoMP architecture, (a) Spectrum after continuous wave laser, (b) Spectrum after Mach-Zehnder modulator, (c) Spectrum after optical coupler, (d) Spectrum before photo detector, (e) Spectrum after photo detector.

Here V_π is the switching voltage of the MZM and V_{bias} is the DC bias voltage.

Table 1. Parameters of the components used in the proposed system

Parameter	Value
Optical power of laser P_{laser}	10 dBm
Linewidth of laser beam	10 MHz
Laser frequency f_c	193.1 THz
Insertion loss of OS I_{OS}	3 dBm
Insertion loss of OC I_{OC}	3 dBm
Dispersion D of SMF	16.75 ps/nm/km
Noise figure of EDFA	6 dB
Bandwidth B of OBPF	20 GHz
Insertion loss I_{filter} of OBPF	3 dB
Order N of OBPF	2
Central frequency of OBPF f_{filter}	193.13 THz
Length of SMF	10 Km
Responsivity of photo-diode	0.9 A/W

The hierarchically modulated signal $X_{HM}(t)$ has a frequency of 30 GHz and is obtained by mapping the low priority data stream on the high priority data stream. The output spectrum of the MZM is shown in Fig. 5(b), where it can

fridaya@uef.fi

be seen that sidebands are generated due to modulation of the optical signal. Transmitting such a signal over an optical fiber will induce dispersion dependent power fading [17]. In order to mitigate the effect of power fading due to fiber dispersion, only a single sideband is transmitted by employing an optical filter. To filter the required optical sideband at a frequency $f_c + 30\,GHz$, the output optical field $E_{MZM}(t)$ of the MZM is fed to a 3 dB Gaussian optical bandpass filter. The filter has a center frequency of $f_c + 30\,GHz$ and a transfer function that can be represented as:

$$H(f) = I_{filter} e^{-\ln(\sqrt{2})\left(\frac{f - f_{filter}}{B/2}\right)^{2N}} \tag{4}$$

Here B is the bandwidth (Hz), N is the order and I_{filter} is the insertion loss of the OBPF. The optical signal $E_{filter}(t)$ at the output of the Gaussian bandpass filter is coupled with $E_2(t)$ by employing a $3\,dB$ optical coupler and transmitted towards the RAU through a 10 Km single-mode fiber (SMF), as shown in Fig. 5. At the RAU, the received single sideband with carrier signal is amplified using an Erbium Doped Fiber Amplifier (EDFA) and then fed to a high bandwidth PIN photo-diode (PD). At the output of the PD, a 30 GHz mm-wave signal is generated as a consequence of heterodyne detection [18]. The spectrum of the mm-wave signal at the output of the PD is shown in Fig. 5 (e). The current generated at the output of the PD denoted as $I_{HD}(t)$, can be written as [19]:

$$I_{HD}(t) = R|E_s^{RAU}(t) + E_c^{RAU}(t)|^2 \tag{5a}$$

$$I_{HD}(t) = R|E_s^{RAU}(t)|^2 + R|E_c^{RAU}(t)|^2 + 2R|E_s^{RAU}(t)||E_c^{RAU}(t)|\cos(\omega_{diff}t + \phi_{diff}) \tag{5b}$$

Here $\omega_{diff} = \omega_s - \omega_c$, $\phi_{diff} = \phi_s - \phi_c$, R is the responsivity of the PD, $|E_s^{RAU}(t)|$ and $|E_c^{RAU}(t)|$ are the electric fields of the optical sideband and the optical carrier signal received at the RAU, respectively. The harmonics generated at the output of the PD are filtered out by employing a Gaussian electrical bandpass filter (EBPF) having a 3 dB bandwidth of 20 GHz. The output of the EBPF can be written as:

$$2R|E_s^{RAU}(t)||E_c^{RAU}(t)|\cos(\omega_{diff}t + \phi_{diff}) \tag{6}$$

It may be observed from the above equation that the power of the mm-wave signal at the output of the PD or EBPF depends upon the powers of the optical sideband, the carrier and the responsivity of the PD. At this stage, the mm-wave signal can be amplified and transmitted to the mobile users via an antenna. However, to observe the quality of the received basic and secondary layers of the HM signal, the mm-wave signal is first down-converted to the baseband signal using a local oscillator and a mixer. The baseband signal is decoded and a decision circuit determines the quadrant of the constellation points. Finally, BER measurements are performed on the bits received as a result of demodulation and decoding. Table 1 shows the parameters of the major components used in our simulation study.

fridaya@uef.fi

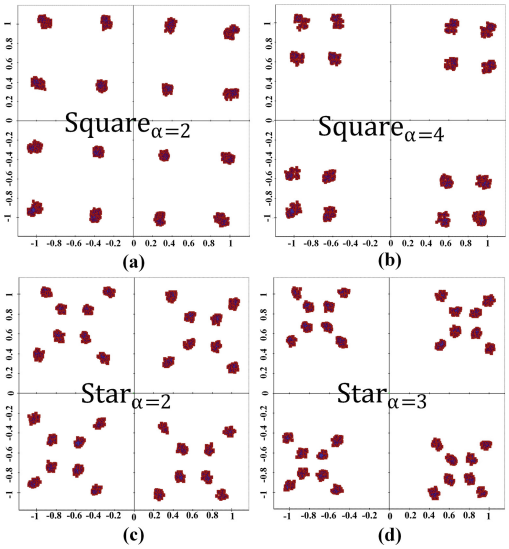


Fig. 6. Constellation of square QAM hierarchical modulation for (a) $\alpha = 2$ and (b) $\alpha = 4$. Constellation of star QAM hierarchical modulation for (c) $\alpha = 2$ and (d) $\alpha = 3$

4 Simulation Results and Discussion

The BER performance of the proposed architecture is shown in Fig. 7. The Conventional 16-QAM is used as a reference here. The high priority layer in Square$_\alpha = 2$ requires less power to maintain the BER of 10^{-3} as conventional QAM, but the difference is small. The reason is that for the value of $\alpha = 2$, the constellation adopts almost the same shape as the mapping of conventional QAM. The LP data forms a cloud around HP data. Therefore, when we receive only the HP data, the LP data points act as noise. This results in an increase in the BER. For a LP layer of square QAM having $\alpha = 2$, data points are distant apart. Therefore, they need less power to get the same BER as compared to other values of α. As the value of α increases for the Star QAM and Square QAM, the HP layer needs lesser power to maintain the BER of 10^{-3}. Whereas, LP layer needs higher power to maintain the same BER of 10^{-3}. The reason is that with an increase of α, the constellation points of low priority layer come closer and hence, needs higher power to maintain the same BER. On the other hand, for the high priority layer, when α increases the cloud of LP layer points shrink. The HP layer appears to be less noisy and results in the requirement of lesser power to maintain the BER of 10^{-3}. The shrinking of the cloud of LP layer points for Square QAM and Star QAM can be observed from Fig. 6. The BER for different values of α for Star and Square QAM are shown in Fig. 7.

We performed an analysis on the improvement of power margin due to HM. Figure 8 illustrates the power margin improvement based on HM. The signals on HP layer have better BER performance as compared to the signals on LP layer.

fridaya@uef.fi

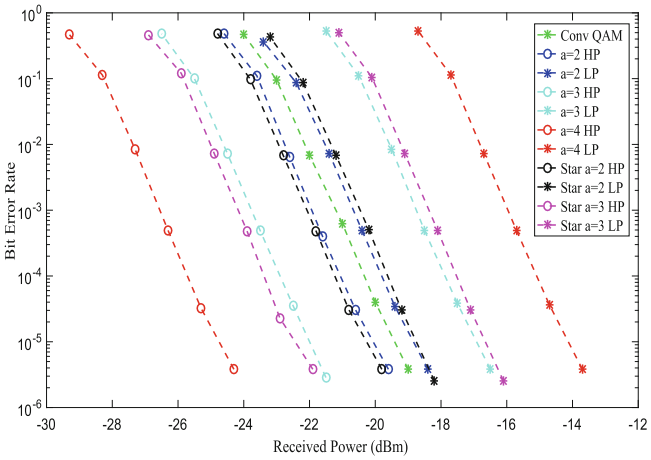


Fig. 7. BER vs received power for star and square hierarchical QAM with different value of α

This is because the HP layer only contains four constellation points which can be easily demodulated by detecting their respective quadrants. The power for the case of Square$_{\alpha=2}$ QAM has a variation of 0.75 dBm compared to conventional 16-QAM, as shown in Fig. 8. It may be observed from the figure that there is an improvement as well as degradation of 0.75 dBm in the HP layer of Square$_{\alpha=2}$ QAM and the LP layer of Square$_{\alpha=2}$ QAM, respectively. The reason for this power variation is that the constellation points of LP layer acts as a source of noise when we are demodulating the HP layer. Therefore, high power is needed to overcome this issue in HP layer. In HM, the HP layer needs higher power compared to the conventional QAM. The power distribution between the layers depends upon α. When the value of α increases, the power resides more in the HP layer than the LP layer. In Star$_{\alpha=2}$ QAM, the power difference increases to 0.8 dBm due to higher data rates. From Square$_{\alpha=2}$ QAM to Star$_{\alpha=4}$ QAM, power margin increases from 0.75 dBm to 5.5 dBm. The received optical power and sensitivities of the layers generate an inequality of $P_{HP} < P_Q$ and $P_{LP} > P_Q$. Here P_Q is the received optical power of conventional QAM and P_{HP} and P_{LP} are receiver sensitivities of both layers for the same value of BER. These inequalities confirm the re-allocation of power margin from one layer to the other layer and therefore, results in the balancing of the performance. The performance balancing is translated to an improvement in power margin by $P_M = min(P_Q - P_{HP}, P_{LP} - P_Q)$. Therefore, we performed an analysis to observe the variation in power margin for different values of α at different receiving conditions.

fridaya@uef.fi

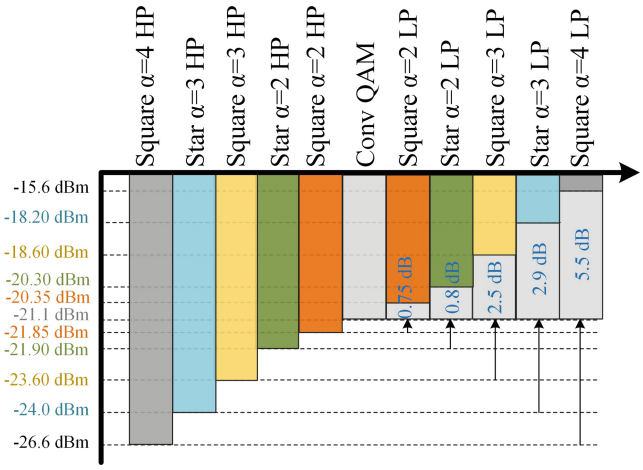


Fig. 8. Histogram for power margin improvement

5 Conclusion

We demonstrate a 30-GHz QAM mm-wave architecture where both conventional I-Q QAM and Hierarchical QAM are implemented and their power margin is analyzed along with the BER performance. HM improved the performance of the system by 0.75 dBm, 0.8 dBm, 2.5 dBm, 2.9 dBm and 5.5 dBm for $\text{Square}_{\alpha=2}$, $\text{Star}_{\alpha=2}$, $\text{Square}_{\alpha=3}$, $\text{Star}_{\alpha=3}$ and $\text{Square}_{\alpha=4}$ QAM, respectively than the conventional I-Q QAM. This performance improvement may lead to an increase in data rate or extend the transmission distance of a RoF system.

References

1. Thomas, V.A., Ghafoor, S., El-Hajjar, M., Hanzo, L.: The rap on ROF: radio over fiber using radio access point for high data rate wireless personal area networks. IEEE Microw. Mag. **16**(9), 64–78 (2015)
2. Irmer, R., Droste, H., Marsch, P., Grieger, M., Fettweis, G., Brueck, S., Mayer, H.-P., Thiele, L., Jungnickel, V.: Coordinated multipoint: concepts, performance, and field trial results. IEEE Commun. Mag. **49**(2), 102–111 (2011)
3. Lee, H.-K., Moon, J.-H., Mun, S.-G., Choi, K.-M., Lee, C.-H.: Decision threshold control method for the optical receiver of a WDM PON. J. Opt. Commun. Netw. **2**(6), 381–388 (2010)
4. Jiang, W.-J., Lin, C.-T., Shih, P.-T., He, L.-Y.W., Chen, J., Chi, S.: Simultaneous generation and transmission of 60-GHz wireless and baseband wireline signals with uplink transmission using an RSOA. IEEE Photonics Technol. Lett. **22**(15), 1099–1101 (2010)
5. Jia, Z., Yu, J., Ellinas, G., Chang, G.-K.: Key enabling technologies for optical wireless networks: optical millimeter-wave generation, wavelength reuse, and architecture. J. Lightwave Technol. **25**(11), 3452–3471 (2007)

fridaya@uef.fi

6. Thomas, V.A., El-Hajjar, M., Hanzo, L.: Performance improvement and cost reduction techniques for radio over fiber communications. IEEE Commun. Surv. Tutorials **17**(2), 627–670 (2015)
7. Tamgnoue, V., Moeyaert, V., Bette, S., Megret, P.: Performance analysis of DVB-H OFDM hierarchical modulation in impulse noise environment. In: 2007 14th IEEE Symposium on Communications and Vehicular Technology in the Benelux, pp. 1–4. IEEE (2007)
8. Sun, H., Dong, C., Ng, S.X., Hanzo, L.: Five decades of hierarchical modulation and its benefits in relay-aided networking. IEEE Access **3**, 2891–2921 (2015)
9. Tao, L., Ji, Y., Liu, J., Lau, A.P.T., Chi, N., Lu, C.: Advanced modulation formats for short reach optical communication systems. IEEE Netw. **27**(6), 6–13 (2013)
10. Cao, P., Hu, X., Zhuang, Z., Zhang, L., Chang, Q., Yang, Q., Hu, R., Su, Y.: Power margin improvement for OFDMA-PON using hierarchical modulation. Opt. Express **21**(7), 8261–8268 (2013)
11. Zhang, L., Cao, P., Hu, X., Liu, C., Zhu, M., Yi, A., Ye, C., Su, Y., Chang, G.-K.: Enhanced multicast performance for a 60-GHz gigabit wireless service over optical access network based on 16-QAM-OFDM hierarchical modulation. In: 2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC), pp. 1–3. IEEE (2013)
12. Iiyama, N., Shibata, N., Kani, J.-i., Terada, J., Kimura, H.: Application of hierarchical modulation to optical access network. In: 2014 12th International Conference on Optical Internet 2014 (COIN), pp. 1–2. IEEE (2014)
13. Zhang, L., Liu, B., Xin, X., Wang, Y.: Peak-to-average power ratio mitigation and adaptive bit assignment in single-carrier frequency division multiplexing access via hierarchical modulation. Opt. Eng. **53**(11), 116115 (2014)
14. Hausl, C., Hagenauer, J.: Relay communication with hierarchical modulation. IEEE Commun. Lett. **11**(1), 64–66 (2007)
15. Jiang, H., Wilford, P.A.: A hierarchical modulation for upgrading digital broadcast systems. IEEE Trans. Broadcast. **51**(2), 223–229 (2005)
16. Wang, S., Kwon, S., Byung, K.Y.: On enhancing hierarchical modulation. In: 2008 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, pp. 1–6. IEEE (2008)
17. Havstad, S.A., Sahin, A.B., Adamczyk O.H., Xie, Y., Willner, A.E.: Optical compensation for dispersion-induced power fading in optical transmission of double-sideband signals. US Patent 6,388,785, 14 May 2002
18. Kuri, T., Kitayama, K.-I.: Optical heterodyne detection technique for densely multiplexed millimeter-wave-band radio-on-fiber systems. J. Lightwave Technol. **21**(12), 3167 (2003)
19. Thomas, V.A., El-Hajjar, M., Hanzo, L.: Millimeter-wave radio over fiber optical upconversion techniques relying on link nonlinearity. IEEE Commun. Surv. Tutorials **18**(1), 29–53 (2016)

fridaya@uef.fi

Dynamic Bit Loading with the OGFDM Waveform Maximises Bit-Rate of Future Mobile Communications

Mohammad R. Kadhum(✉), Triantafyllos Kanakis, and Robin Crockett

Faculty of Arts, Science and Technology, University of Northampton, Northampton, UK
mohammad.kadhum@northampton.ac.uk

Abstract. A new Dynamic Bit Loading (DBL) scheme with the Orthogonal Generalized Frequency Division Multiplexing (OGFDM) is, for the first time, proposed, discussed and assessed. The key concept of this hybrid modulation format depends substantially on the adaptive distribution of the bit stream to be more compatible with the gained capacity of the realistic channel state. Due to the negative impact of employing the fixed schemes of digital modulation on the performance of the conventional telecommunications systems, the influence of using the multi-level modulation system is investigated for the future applications of mobile communications. Utilising the DBL in the physical layer (PHY), a flexible range of modulation formats can be optimally assigned for each applied frequency sub-carrier in accordance with wireless channel circumstances. In addition, depending on the supportive features of the proposed modulation system, the performance in terms of channel capacity can be maximised at the acceptable limit of the Bit Error Rate (BER). As such, an extra enhancement can be achieved in the spectrum efficiency (SE) of the adaptively modulated wireless signal. Thus, an adjustable boost of the transmission range of used modulation formats can be reached with the introduced adaptation system. The performance of the DBL system through a wireless mobile channel under the Additive White Gaussian Noise (AWGN) is evaluated according to a various level of the Signal to Noise Ratio (SNR). Ultimately, regarding the numerical simulation, a MATLAB code is employed to simulate the performance (channel capacity & BER) of the proposed DBL that is fundamentally accommodated by the recent candidate waveform of future mobile technology (OGFDM).

Keywords: Orthogonal Generalized Frequency Division Multiplexing (OGFDM) · Dynamic Bit Loading (DBL) · Modulation formats · Channel capacity · Signal to Noise Ratio (SNR) · Bit Error Rate (BER) · Physical layer (PHY) · Mobile communications · Spectrum Efficiency (SE)

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 242–252, 2019.
https://doi.org/10.1007/978-3-030-22868-2_19

fridaya@uef.fi

1 Introduction

The ever-growing demand of higher bit-rate and the predicated applications of the future mobile generation make researchers focus their attention on evolving the applied bit loading system from the conventionally fixed schemes towards a more dynamic, scalable, and smarter configuration. Thus, the flourishing development of mobile communications systems can help to offer improved data rates which can convoy with the fast change of traffic patterns from voice to video-based services in wireless access networks [1].

Despite the rapid progression in the digital signal processing (DSP) domain, optimisation of the transmission resources, particularly, that is directly related to the bandwidth (BW) efficiency, is yet a key issue for future wireless communications of mobile [2].

Its worth noting that the fixed modulation methods which are already employed in the traditional mobile communication system, are not quite matching with the developed requirements of the wireless transmission channel that experiences diverse types of complicated characteristics like dispersive fading and time-variant phenomena [3]. Thus, the BER performance of the system can be improved whenever the circumstance of the incompetent channel is enhanced [4].

Generally speaking, as the wireless channel suffers from frequency selective fading and time-varying phenomena, the system performance of channel capacity can be limited according to the worse channel condition [5]. Thus, with the fixed modulation, the above undesirable factors can principally impact the system performance in terms of maximum bit-rate.

To mitigate this issue, it is necessary to adopt a different scheme of modulation which is assigned adaptively for different frequencies of sub-carriers according to the characteristics of the applied channel [6]. This hybrid scheme of modulation/de-modulation which merges different kinds of formats in one appropriate modulation shape is called herein as Dynamic Bit Loading (DBL).

The orthogonal generalized frequency division multiplexing (OGFDM) which is considered recently as a promising waveform for the future generation of mobile (5G and beyond) [7], is applied here as an accommodating environment for the DBL technique.

Utilising a more flexible level of the digital modulation/de-modulation system with the future candidate waveform OGFDM, a further efficient spectrum allocation of the wireless BW resources is achieved. As such, the newly introduced DBL can be exploited for improving the performance of channel capacity for future mobile technology. Hence, depending on the physical layer (PHY) of the proposed DBL over the OGFDM system, the wireless channel capacity can be maximised under key transmission constraints like reception power and Bit Error Rate (BER).

Depending on the DBL scheme, the same frequency sub-carrier is possibly reused in various ways by allocating varied modulation schemes to each located user based on transmission conditions [8]. Thus, according to diverse channel circumstances, variable bit stream of employed sub-carriers are allocated dynamically [9].

fridaya@uef.fi

The proposed technique can back a variety of modulation formats for adjusting the key throughput parameters according to the channel quality. As a result, the flexible assignment of channel modulation can significantly enhance the bit-rate and Spectrum Efficiency (SE) by adjusting the bit distribution of the sub-carriers according to nature of subscribers' channels [10]. It is also worth pointing out that, under a varying selection of the modulation shapes, a more accurate bit-rate and a better spectral feature usage of transmission channel are obtained.

To adequately deal with the present channel conditions, signal propagation parameters are adaptively adjusted by the developed scheme of modulation format. Hence, the DBL enables the propagation system to optimise the performance based on the wireless link conditions [10]. Thus, via adaptively allocating the suitable modulation shape, the wireless transmission waveform can be maintained in all possible situations. Besides, exploiting the diversity of the sub-carrier quality in frequency domain comes up with allocating different levels of the quality of boosting [11].

It's worth noting that knowing state information of employed wireless channel can allocate efficiently the bits for each used sub-carrier in the OGFDM system.

Assuming the instantaneous states of the channel for all subscribers are recognized by the Base Station (BS), the DBL system can employ a higher modulation scheme with the frequency sub-carrier which physically has a large gain (high priority) to transfer more bits per sample and vice-versa [12]. Hence, a high order modulation can be chosen to support the good condition channel where a maximum limit of channel capacity and best SE are reached at acceptable limits of the BER. Thus, a preferable modulation format is allocated by the BS for each subscriber in accordance with the strength of the wireless signal interference [13].

The key principle of the flexible modulation is, its ability to bend to the real fading conditions in comparison with the fixed scheme which is essentially intended for the worst circumstances [14]. This, nevertheless, causes in the power diversity of constellation table which can be addressed by the normalisation process [15].

This paper is structured as follows: Sect. 1 is the introduction. Section 2 demonstrates the system model of the DBL that explains briefly the PHY of employing the proposed scheme in both the transmitter and the receiver sides. Section 3 presents the simulated work including the results and discussion. The conclusion is presented in Sect. 4.

2 System Model

Making use of the wireless channel conditions, the proposed adaptive format can expand the gained channel capacity at the intended limit of errors. To apply this, the dynamic bit loading system is utilised alternatively to mitigate the downside of the conventional loading of bits.

As is seen in Fig. 1, at the transmitter side, a wide range of the most common modulation formats can be employed for each used sub-carrier adaptively. Hence, depending on a variant size of bit token (N), assorted constellations can be

fridaya@uef.fi

introduced. Thus, varies average power limits for each dynamic constellation table is achieved. This is, as such, can come up with introducing a multi-level modulation scheme instead of the flat one. In addition, an extra number of bits can be obtained from this flexible shape of modulation increasing the bit-rate of transmission.

The hybrid scheme which is essentially used for transforming the assigned bits to its corresponding complex numbers can improve the freedom level of digital modulation. Hence, in the frequency domain, every sub-carrier can have a specifically different power consumption according to the modulation format employed. This, however, can result in an irregular power constellation map due to diverse schemes of bit loading. Nevertheless, utilising the normalisation process of scaling stage, the average energy of the combined Gray coded bit mapping is set to one.

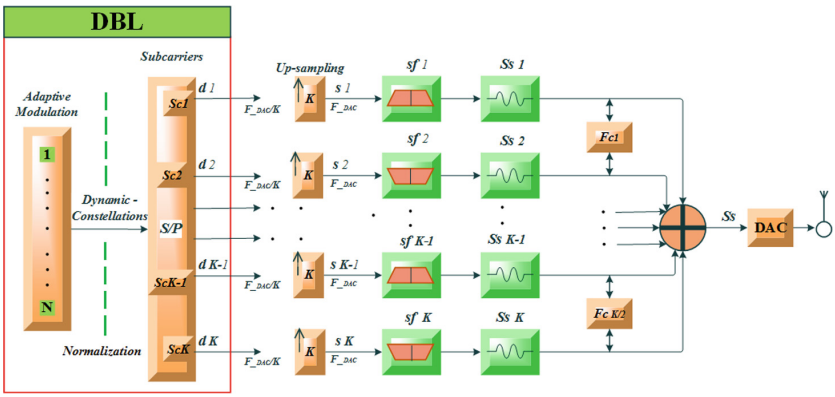


Fig. 1. Block diagram of the proposed DBL for the transmitter side of the OGFDM system.

In this paper, its worth noting that, all management operations of the bits loading scheme for both the transmitter and receiver are referred to as the DBL part.

From the time domain perspective, every generated complex number is equivalent to a sample, whereas, each group of samples with K length is equal to one OGFDM symbol. The normalized sub-carrier is subsequently up-sampled by a factor of K where $K - 1$ zeroes samples are inserted between any two adjacent points.

After that each up-sampled frequency sub-carrier is convoluted with one of shaping Hilbert filters (cosine or sine) producing an orthogonally filtered sub-carrier. In addition, the digital filters are employed for multiplexing the shaped frequencies of sub-carriers into an obtainable BW. Employing the electrical adder, the carried samples of all utilised frequencies of sub-carriers are

fridaya@uef.fi

ultimately collected in one digital signal. To put this signal in the transmission mode, the Digital to Analog Converter (DAC) is used outputting the analog signal. The delivered signal which is exponentially denoted as $(e^{j2\pi f_c t})$ is transmitted by the antenna.

As is shown in Fig. 2, the detected signal at the antenna of the receiver side is delivered to an Analog to Digital Converter (ADC). As such, inverse operations are launched to retrieve the original transmitted signal. The digital signal is then distributed into a various set of sub-carriers (frequencies), where, every two orthogonal sub-carriers are recognized by an identical frequency centre (f_c).

After that, the formerly convoluted sub-carriers are extracted utilising the matching filters of Hilbert pairs, where, each matching filter is corresponding to it shaping filter. Subsequently, the de-multiplexed sub-carriers are down-sampled by a factor of K, eliminating the $K-1$ zeroes of neighbouring samples.

In the frequency domain, where every sample is represented by a complex number, a flexible de-modulation system is employed to convert adaptively complex numbers to binary digits. Thus, according to the channel conditions, a dynamic range of the most popular de-modulation schemes can be used with an applied frequency sub-carrier. Hence, based on the diversity of de-modulation formats, multi-levels of the de-modulation process can be applied for frequencies of sub-carriers. As such, a specific level of power is consumed differently by the adaptive sub-carriers in accordance with the channel state.

As a result, the average power of the receiver constellations table can be unstable. Even though, a uniform level of the constellations map power can be obtained by the normalisation operation. Each complex number in this hybrid de-modulation shape is translated dynamically to N bit token, where 2^N indicates the order of utilised modulation/de-modulation scheme for every used sub-carrier.

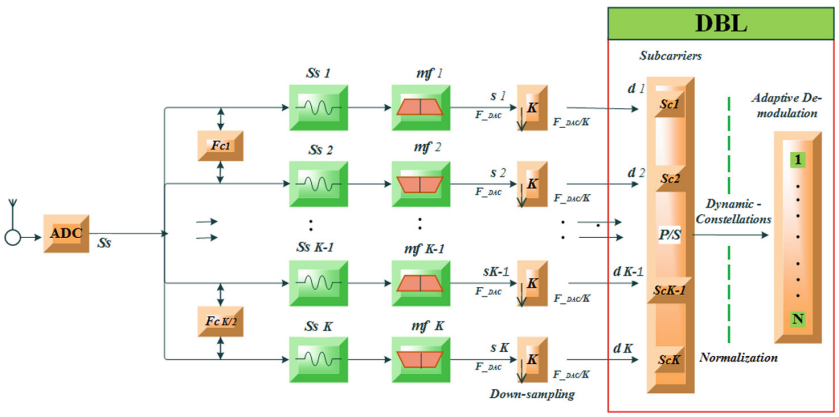


Fig. 2. Block diagram of the proposed DBL for the receiver side of the OGFDM system.

fridaya@uef.fi

The vital idea of this proposed system is utilising the DBL process to combine different types of modulation/de-modulation formats adaptively. Consequently, the major parameters of this method (complex numbers and bit stream) and their relational impacts are mathematically considered here. Accordingly, in terms of the complex numbers usage, the basic formula of the flexible modulus (M_i) for each recorded point at the mapping table can be defined as follows [16]:

$$M_i = \sqrt{X_i^2 + Y_i^2} \tag{1}$$

where X_i, Y_i the real and imaginary parts of i_{th} constellation point, respectively.

Also, the phase effect (ϕ) of intended constellations is calculated as the following [16]:

$$\phi_i = \arctan(Y_i / X_i) \tag{2}$$

In addition, depending on the gained value of modulus (M_i), the average power of the transmitted signal (AP) with length (L) is illustrated as follows [17]:

$$AP = \frac{1}{L} \sum_{i=1}^{L} M_i^2 \tag{3}$$

Moreover, the consumed power of the constellation table (P_{CT}) for a fixed format of modulation is represented by the following [18]:

$$P_{CT} = \frac{2}{3} * (Z - 1) \tag{4}$$

where Z indicates the order of modulation format.

It's worth mentioning that under the DBL scheme, multi-levels of power are obtained for the hybrid Gray mapping table. Hence, every utilised sub-carrier is treated differently based on the modulation shape used. This, however, can be unified to one level of consumption power via dividing every complex number ($X + Yj$) belong to an applied sub-carrier by the square root of its constellation table power. Thus, the normalisation process (N_P) is employed with each point to adjust the average power of the dynamic mapping table to one as follows [18]:

$$N_P = \frac{X + Yj}{\sqrt{P_{CT}}} \tag{5}$$

The system performance in terms of the maximum bit-rate of transmission (Br) for a symbol slot of time (t) at the accepted limits of the BER, is obtained as the following [19]:

$$Br = \frac{\sum_{i=1}^{K} b_i}{t} \tag{6}$$

where b_i denotes the number of transmitted bits through K available frequencies of sub-carriers.

fridaya@uef.fi

Based on this equation, the upgraded channel capacity can be reached under the satisfactory bounds of errors.

Moreover, the influence of assumed wireless channel (Es) on the received signal (Rx) can be illustrated as the following [19]:

$$Rx = Es * Tx + N \tag{7}$$

where Tx represents the impacted transmitted signal in the presence of a noise (N).

A related point to consider here is the relation between the mean value of the expected signal (μ) to the standard deviation of the undesirable signal (σ) can be shown as follows [19]:

$$SNR = \frac{\mu}{\sigma} \tag{8}$$

In terms of improved bit-rate, the DBL scheme can be more compatible with the dynamic range of the SNR than the fixed formats of modulation which is recommended with the static SNR (calculated for worst channel condition) [20].

3 Experimental Work

To explore the proposed DBL of the future candidate waveform (OGFDM), a numerical simulation is applied at the PHY of an electrical back-to-back (B2B) transmission system.

As is clear in Fig. 3, three key regions are essentially exploited based on the channel transmission circumstances. Hence, for a dynamic SNR range, several important spots are mainly introduced for every two adjacent shapes of the fixed modulation. Thus, various power thresholds are manipulated between the minimum levels of required SNR for both the first and second proposed styles of modulation (128 QAM & 256 QAM). Besides, the emerging improvements of the SNR which basically result from the enhanced status of the transmission channel are optimally employed by the DBL system.

Regarding the investigated cases, the first one is the "Low Boost" (LB) case, which is counted for the SNR threshold that is typically a higher than the first fixed threshold with the 128 QAM but less than the central point of both decided modulation formats. The LB case can come from different scenarios of Bit Loading Map (BLM). As such, depending on the ratio of enhancements in the utilised frequency sub-carriers, the BLM is obtained. For example, for eight sub-carriers, when the rate of improvement is equivalent to 25%, the BLM is equal to one of the following probabilities, [8, 7, 7, 7, 7, 7, 7, 8], [7, 8, 7, 7, 7, 7, 8, 7], [8, 8, 7, 7, 7, 7, 7, 7], etc. This, as a result, can come up with a slight increase in the capacity of the channel due to the growing ability of about 25% of the sub-carriers to carry added bits.

Regarding the second case "Medium Boost" (MB), the SNR threshold herein is acquired in the middle distance between the first and the second threshold of both 128 and 256 QAM respectively ($SNR_{128QAM} < SNR_{MB} < SNR_{256QAM}$).

fridaya@uef.fi

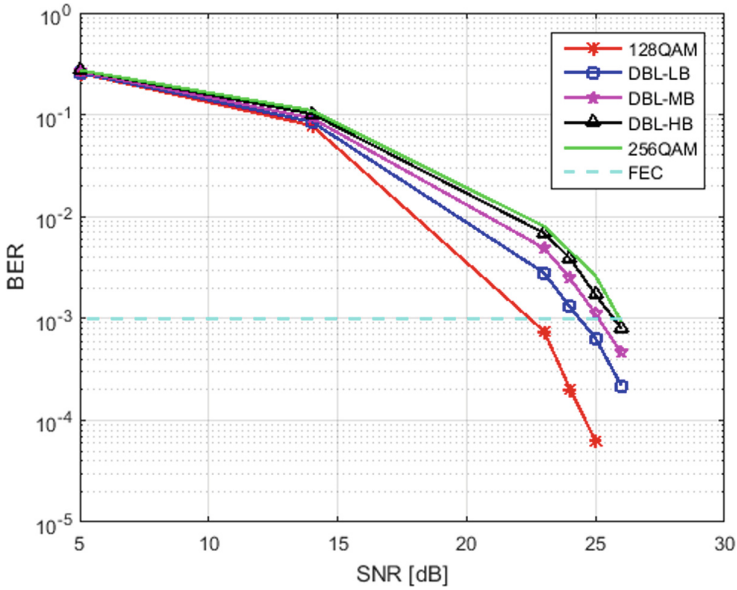


Fig. 3. Apply the DBL system for induced SNR between two successive modulation formats.

Moreover, a fair rise in the channel capacity results from the amended ability for approximately 50% of the utilised sub-carriers. The BLM of this situation, where the proportion of increment is averaged, equivalents to one of the next prospects, [8, 8, 7, 7, 7, 7, 8, 8], [8, 8, 8, 8, 7, 7, 7, 7], [7, 7, 8, 8, 8, 8, 7, 7], etc. Thus, according to the intermediate improvement of the channel circumstance, half of the sub-carriers are capable to carry further bits. Hence, the gained channel capacity of case two is higher than case one due to the increased number of promoted sub-carriers.

Regarding the third case "High Boost" (HB), due to an extremely improved circumstance of the channel, the SNR threshold is recorded herein near to the second fixed threshold with the 256 QAM. Thus, most of the employed sub-carriers (around 75%) are furthered giving an extra ability to carry more bits. As such, the BLM for this advanced state counterparts to one of the following diagnoses, [8, 8, 8, 7, 7, 8, 8, 8], [7, 8, 8, 8, 8, 8, 8, 7], [8, 8, 8, 8, 8, 8, 7, 7], etc. This, as a result, comes up with a higher channel capacity in relative to the formerly mentioned cases (MB, LB).

It's worth noting that all these achieved thresholds of the SNR are allocated for upgraded statuses of the channel condition that are fundamentally bounded by two successive modulation formats ($SNR_{MF} < SNR_{DBL} < SNR_{MF+1}$).

As is seen in Fig. 4, graduated amounts of channel capacity can be obtained between any consecutively applied schemes of fixed modulation. Thus, three channel capacities are elected mainly to give a good example of the diversity in transmission bit-rate for enhanced statuses of channel condition. The gained

fridaya@uef.fi

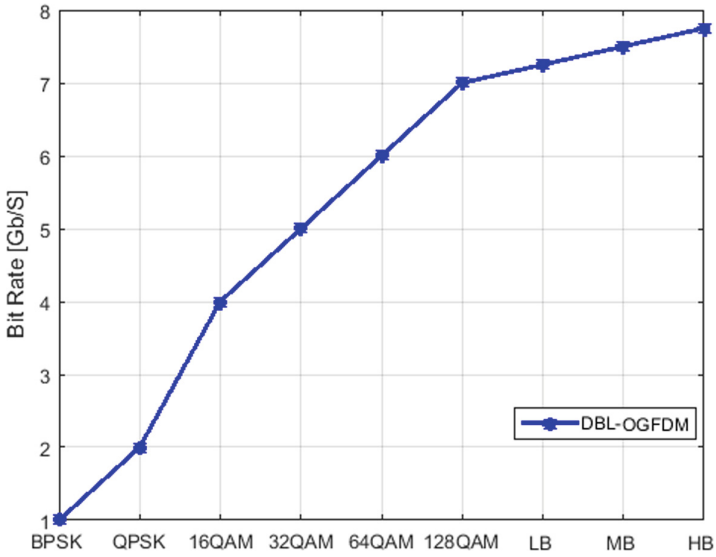


Fig. 4. Maximised bit-rate with the DBL system (LB, MB, HB) and BER equals to (10^{-3}).

channel capacities are varied herein from low increment with the LB, to medium increment with the MB and to high increment with the HB.

In addition, with the DBL system, the bit-rate of transmission can be maximised in accordance with the grade of the channel condition improvement. Thus, herein, the raised level of the SNR can play a big role in promoting the capacity of the channel rather than the BER performance. Hence, when the provided SNR is developed, only the achieved bit-rate can be enlarged keeping a stable case of the BER (10^{-3}) for all dynamic gains between two adjacent shapes of fixed modulation.

As a result, comparing with the fixed modulation schemes, the improved channel capacities are obtained using the DBL with a changeable channel status. Thus, except for the worst channel condition, achieved channel capacities

Table 1. System parameters for the DBL-OGFDM

Parameter	Value
No. of frequency centres	4
$F_{DAC/ADC}$	4 GHz
SNR	Static & dynamic
Modulation format	Fixed & adaptive
OGFDM symbols	2000
Filter type	Hilbert filter

fridaya@uef.fi

with the DBL system can outweigh the gained channel capacity of a fixed SNR threshold. Hence, in contrast with the fixed modulation, the transmission bit-rate can be maximised up to an extra 14%, i.e. 114% of the original channel capacity with a similar level of error. The performance of transmission, in terms of the channel capacity is examined under the condition stated in Table 1.

4 Conclusion

In this paper, a novel way of mapping bits among utilised sub-carriers of frequencies in the OGFDM system is introduced, explored and evaluated. The proposed DBL scheme comes from substituting the conventionally fixed bit loading system of the OGFDM with another flexible design. The main feature of this optimised technique is the preferably adaptive allocation of the applied stream of bits in accordance with the instantaneous channel status. The multi-levels adjustment comes up with maximising the performance capacity of the wireless channel without the need for employing a higher fixed modulation scheme. Thus, the hybrid modulation/de-modulation mechanism of future mobile generation leads to an extra enhancement in the efficiency of BW usage with maintaining a suitable level of the BER. The implemented work clarified that utilising the DBL with changeable channel cases, several better channel capacities can be achieved in comparison with the conventional fixed modulation format. Hence, depending on the typically supposed circumstances of the transmission channel, three key exploiting regions (low, medium, high) can be declared for any dynamic SNR area of each two adjacent fixed modulation shapes. Thus, apart from the worst condition, all gained channel capacities of the DBL configuration exceed the threshold channel capacity of the used fixed scheme of modulation.

Acknowledgment. This research was funded by the Ministry of Higher Education and Scientific Research, Republic of Iraq, Scholarship 2633.

References

1. Xiao, Y.U.E., Haas, H., Member, S.: Index modulation techniques for next-generation wireless networks. IEEE Access **5**, 16693–16746 (2017)
2. Kadhum, M. R., Kanakis, T., Crockett, R.: Intra-channel interference avoidance with the OGFDM boosts channel capacity of future wireless mobile communication. In: Proceedings of Computing Conference 2019, London (2019)
3. Ishikawa, H., Furudate, M., Ohseki, T.: Performance analysis of adaptive downlink modulation using OFDM and MC-CDMA for future mobile communications system, pp. 194–198 (2004)
4. Zhang, H., Bi, G., Zhang, L.: Adaptive subcarrier allocation and bit loading for multiuser OFDM systems (2009)
5. Haboobi, H., Kadhum, M.R., Al-sherbaz, A.: Utilize higher modulation formats with heterogeneous mobile networks increases wireless channel transmission. In: Proceedings of Computing Conference 2019, London (2019)

fridaya@uef.fi

6. Hao, W., Hongwen, Y., Jun, T.: A novel adaptive bit loading algorithm for BICM-OFDM system, pp. 311–315 (2012)
7. Kadhum, M.R., Kanakis, T., Al-sherbaz, A., Crockett, R.: Digital chunk processing with orthogonal GFDM doubles wireless channel capacity, pp. 1–6 (2018). https://doi.org/10.1007/978-3-030-01177-2-53
8. Yu, M.: Adaptive bit loading algorithm of shortwave broadband OFDM system, pp. 49–52 (2011)
9. Vo, T.N., Amis, K., Chonavel, T., Siohan, P.: A low-complexity bit-loading algorithm for OFDM systems under spectral mask constraint. IEEE Commun. Lett. **20**, 1076–1079 (2016)
10. Elattar, H., Dahab, M.A.A., Ashour, A.F.: Efficiency improvement using adaptive hybrid modulation/coding/frequency selection scheme for future 5G wireless network, pp. 1392–1396 (2017)
11. Junhui, Z., Guan, S., Gong, Y.: Performance analysis of adaptive modulation system over mobile satellite channels, pp. 3–6 (2011)
12. Jun, W., Yong, W., Xia, Y., Chonglong, W.: Adaptive subcarrier and bit allocation for multiuser OFDM networks (2006)
13. Torrea-duran, R., Desset, C., Pollin, S., Dejonghe, A.: Algorithm for LTE pico base stations, pp. 1–8 (2012)
14. Al-mawali, K.S., Sadik, A.Z., Hussain, Z.M.: Low complexity discrete bit-loading for OFDM systems with application in power line communications. Int. J. Commun. Netw. Syst. Sci. **2011**, 372–376 (2011)
15. Nadal, L., Fàbrega, J.M., Vílchez, J., Moreolo, M.S., Member, S.: Experimental analysis of 8-QAM constellations for adaptive optical OFDM systems. IEEE Photonics Technol. Lett. **28**, 445–448 (2016)
16. Alexander, E. Poularikas, D.: The Handbook of Formulas and Tables for Signal Processing (1999)
17. Ghogho, M., McLernon, D., Alameda-Hernandez, E., Swami, A.: Channel estimation and symbol detection for block transmission using data-dependent superimposed training. IEEE Signal Process. Lett. **12**, 226–229 (2005)
18. Barry, E.J.R., Lee, E.A., Messerschmitt, D.G.: Digital Communication. 3rd edn. (2004)
19. Im, G.H., et al.: 51.84 Mb/s 16-CAP ATM LAN Standard. IEEE J. Sel. Areas Commun. **13**, 620–632 (1995)
20. Kadhum, M.R.: New multi-carrier candidate waveform for the 5G physical layer of wireless mobile networks. In: Proceedings of IEEE 2019 Wireless Days, Manchester (2019)

fridaya@uef.fi

The Two Separate Optical Fibres Approach in Computing with 3NLSE–Domain Optical Solitons

Anastasios G. Bakaoukas(✉)

Faculty of Arts, Science & Technology, Computing & Immersive Technologies Department, University of Northampton, Avenue Campus, St. Georges Avenue, Northampton NN2 6JB, UK
Anastasios.Bakaoukas@northampton.ac.uk

Abstract. While in the past research was focused on how best to treat computational arrangements when all optical solitons were propagating down a single optical fibre, has been increasingly become apparent that using two separate optical fibres to propagate solitons involved in the same computational arrangement possesses a number of unique properties and offers a number of unique possibilities. In this paper we present this alternative approach to computing involving both first-order and second-order optical solitons in the 3NLS–domain to construct optical soliton gates composed of two individual physical optical fibres and solitonic arrangements propating individually and in parallel, yet forming collectively a single solitonic all-optical logic gate fulfilling all requirements for solitonic gateless computing. More specifically, the focus is on investigating fundamental properties of collisions between first order and second order solitons and on providing proof-of-principle for the feasibility of using collisions between such solitons in the 3NLS–domain that propagate in separate optical fibres for computing.

Keywords: Unconventional computing · Solitons · 3NLS–domain · All-optical computing · Schrödinger equation · Soliton collisions · Soliton computational schemes

1 Introduction

The use of soliton optical pulses in the 3NLS–domain for the purpose of carrying out computation has been presented and adequately verified in past studies [1–4]. A number of logic functions and the fundamental digital signal processing related Fast Fourier Transform (FFT) computation were implemented and demonstrated based on first order soliton collisions in the 3NLS–domain. In addition to the common $NAND$ and AND logic gates, first order soliton collisions were proven to be capable of imitating any other possible logic gate with the overall computational system proposed to be computationally universal (Turing Universal), meaning that any computation can be implemented within the

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 253–280, 2019.
https://doi.org/10.1007/978-3-030-22868-2_20

fridaya@uef.fi

boundaries of the system using the appropriate initial arrangement of solitonic pulses. With reversibility being an inherited property of the system, operations will not generate waste heat as a result of the collisions.

Involving a balance between Kerr type non-linearities and dispersive effects in glass fibres, temporal solitons are governing the 3NLS–domain. This is the very type of solitons used to define computational arrangements in the domain, with all their types of interactions being a relatively long-range phenomenon because of the Kerr non-linearity exerting a rather weak effect on the solitonic pulses as they propagate down the optical fibre [5].

Neglecting higher order effects (higher order dispersive and non-linear effects) due to optimum pulse widths selected, temporal solitons in optical fibres are solutions of the integrable cubic non-linear Schrödinger equation (3NLSE):

$$j\frac{\partial u}{\partial \xi} sgn(b_2)\frac{1}{2}\frac{\partial^2 u}{\partial r^2} - N^2|u|^2 u = 0 \tag{1}$$

where b_2 is the second order dispersion parameter and $sgn(b_2) = \pm b_2$. A positive value for the dispersion describes the formation of bright optical solitons whilst a negative value leads to the formation of dark solitons. The 3NLS equation in general, describes a modulated wave packet propagating through a non-linear dispersive medium with a constant velocity [6,7]. A solution of the integrable 3NLSE applicable to pulse propagation in optical fibres (for any number of initial pulses) is the Hyperbolic Secant:

$$\begin{aligned} u(0,\tau) = r\ sech[r(\tau - q_0)]e^{j\theta}e^{jv\tau} + \\ + r\ sech[r(\tau + q_0)]e^{j\theta}e^{jv\tau} + \ldots \end{aligned} \tag{2}$$

where, r represents the amplitude of the solitons, θ is the relative phase value, and q_0 is the initial displacement between the two, or more, solitonic pulses. This function represents the envelope of optical solitons in an optical fibre.

This paper is a numerical study of collisions between first order and second order solitons in optical fibres using four different numerical techniques for carrying soliton propagation and collision simulations, the: (a) "Finite Difference Runge–Kutta Technique" (FDRKT), (b) "Split–step Fourier Transform" (SSFT), (c) "Fourier Series Analysis Technique" (FSAT), (d) "Fuzzy Mesh Analysis Technique" (FMAT) [4].

Unlike previous studies, in this paper is presented an alternative approach to the computational system in the 3NLS–domain, with solitons propagating down the length of two independent optical fibres (the optical logical gate), instead of a single optical fibre, and provided that the output of the two optical fibres (the gate) is combined in the end by a coupler, appropriate practical arrangements – computationally universal systems – based on collisions between first order and second order solitons become possible using logical gates based on the "controlled" gates paradigm originally proposed by Toffoli [8,9].

Collisions of this type of solitons and their controlled distribution over to independent optical fibres, have the appropriate properties to allow, in principle, useful computation to be carried out. Solitons can be made to collide

fridaya@uef.fi

by propagating them with different velocities, which can be achieved by altering the frequency of the modulated wave packet. The work presented in this paper demonstrates the development of a collection of digital gates (i.e. AND, $NAND$, OR etc.) that if taken together form a Boolean complete set, meaning that by using only these digital gates one can implement any arbitrary Boolean function (circuit). Then arranging these Boolean functions in different combinations makes it possible to implement any conventional computing machinery organisation.

The performance of the soliton-based logic gates is estimated by assuming the computational complexity of a single soliton collision. As is the case with other solitonic computational arrangements [10], in the least restrictive model the maximum number of collisions possible is simply counted and used as the absolute upper bound on the complexity of the computational operation taking place in a single collision. As a result, one of the most fundamental properties of the computational system presented here is the fact that, in general, the computational system is realised as a collision system with an n number of 3NLSE–domain solitons entering and exiting the collision system. The collisions between solitons internal to the collision system realise a computation with the input data encoded in the states of the input solitons and the output data encoded in the states of the output solitons. While all computation is internal to the optical medium and is in the form of soliton collisions, not every input soliton necessarily carries input data, some are present only for controlling the computation carried out (Control Solitons) and not every output soliton carries output data as some solitons can now be considered as by-products of the computation (Garbage Solitons).

The material in this paper is presented in a total of five sections. In Sect. 2, an essential discussion is provided about the dynamics of first order and second order solitons for a better understanding of the solitonic logic gate arrangements presented in this paper. Section 3 presents and analytically discusses the foundation concepts involved in collisions between first order and second order solitons. The fundamental concept and realisation parameters of computing with first and second order solitons in the 3NLSE–domain using independent optical fibres are introduced and discussed in Sect. 4. In particular, complete solitonic arrangements are presented for the realisation of an $Inverter\ NOT$ logic gate, an OR logic gate, a NOR logic gate and the very important for every computational system $NAND$ logic gate. Finally, Sect. 5 includes the conclusions of the paper.

2 First Order and Second Order Solitons

As discussed in the previous section, two are the dominant effects on the optical fibre, those of Dispersion and Non-linearity. A balance between them, necessary for the successful and stable formation of solitonic pulses, is achieved in the case that the value for the fibre length parameter L is comparable to both the Dispersion parameter L_D value and the Non-linearity parameter L_{NL} value. When the values for L_D and L_{NL} are such that the two effects cancel each other out, the solitonic pulse propagates with absolutely no change in its initial profile (Fig. 1).

fridaya@uef.fi

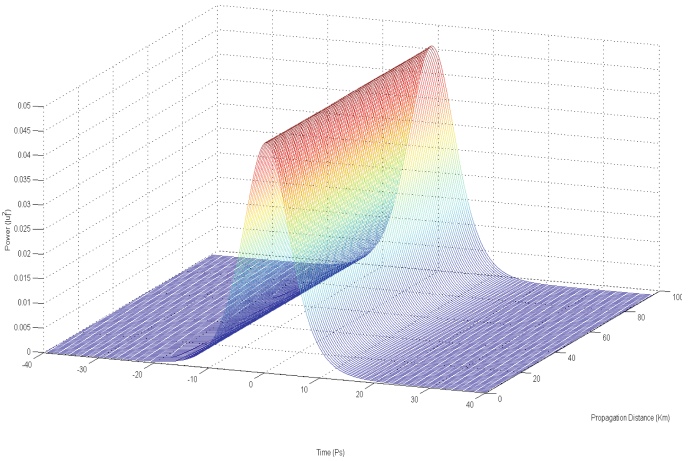


Fig. 1. A Soliton initial pulse propagating in the balance regime (this graph, and all subsequent ones, are plotted relative to the Mean Group Velocity).

For the computational system we present in this paper two kinds of first order solitons are distinguished, one with a phase value of π and one with a phase value of 0. These two types of first order solitons give four possible combinations of phase dependent collisions when taken in pairs. Two first order solitons are colliding and both are having a phase value of *pi* or both have a phase value of 0, while the remaining two refer to cases where one of the solitons has a phase value of π and the other a phase value of 0 and vice versa.

The four combinations are correspondingly defined in general two types: of collisions between the solitons. If the two solitons are in phase, then an "In-phase" collision envelope occurs and secondly, if the two solitons are out of phase, then an "Out-of-phase" collision envelope results (Figs. 2 and 3). These characteristic for the nature of the collision envelopes are due to the attractive force between the two solitons present when they are in-phase and the repulsive force present when they are out-of-phase.

The peak power necessary to launch the N^{th} order soliton is N^2 times that required to launch the fundamental (1^{st} order) soliton for which $N = 1$. So, for the case of a second order soliton now, its propagation behaviour can be interpreted as the combined effect of two first order solitons oscillating backwards and forwards through each other with the actual evolution of the pulse depending on the number of eigenvalues found from the associated scattering problem and this too, depends directly on the value of N.

A second order soliton is plotted in Fig. 4. The input pulse shape is restored periodically at a characteristic distance of z_0, defined as the soliton period. Is exactly this characteristic behaviour of the second order soliton that can be interpreted as the combined effect of two first order solitons oscillating backwards and forwards through each other. The values of N are not rigidly integer.

fridaya@uef.fi

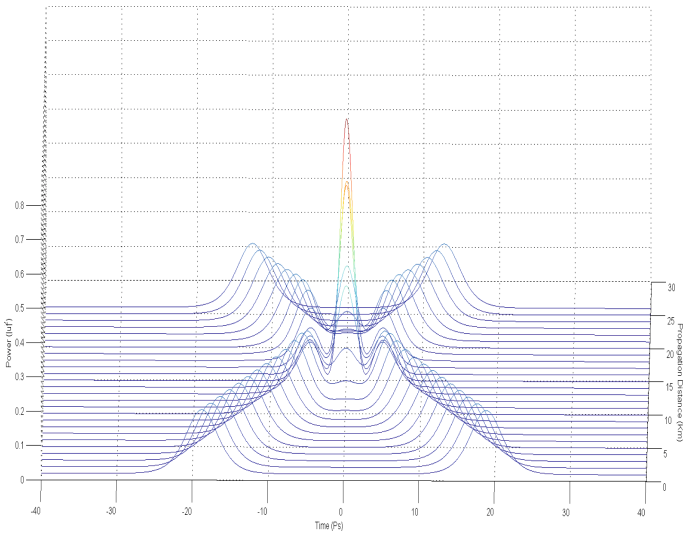


Fig. 2. First order soliton collision. Phase values of 0 & 0 and π & π for the two solitons involved, both produce this characteristic for the nature of the collision envelope (5ps solitons with velocities 0.4 & -0.4, respectively).

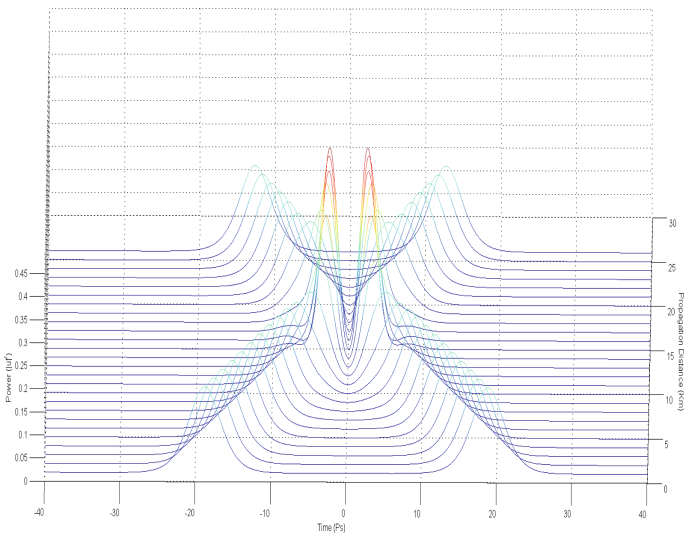


Fig. 3. First order soliton collision. Phase values of π & 0 and 0 & π for the two solitons involved, both produce this characteristic for the nature of the collision envelope (5ps solitons with velocities 0.4 & -0.4, respectively).

fridaya@uef.fi

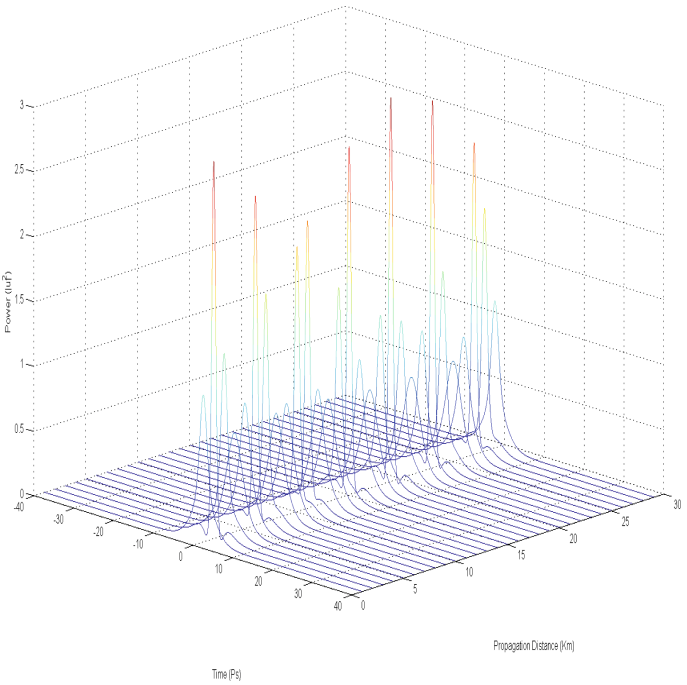


Fig. 4. Evolution over two soliton periods for a second-order soliton.

Fundamental solitons exist for values of N from 1 to 2 and second order solitons exist for values of N from 2 to 3. If however, a soliton is launched with a value of N intermediate between integer values, the resulting soliton radiates away the excess energy and settles to the nearest lower integer. Thus for example, a soliton launched with $N = 2.3$ will radiate light as it propagates and eventually settle as a $N = 2$ soliton at which point the radiation of energy ceases. The soliton period z_0 is then defined as:

$$z_0 = \frac{\pi}{2} L_D = \frac{\pi}{2} \frac{T_0^2}{|k_2|} = 0.322 \frac{\pi T_{FWHM}^2}{2|k_2|} \tag{3}$$

3 Collisions Between First and Second Order Solitons

In Figs. 5 and 6 a numerical simulation of the collision between first and second order solitons is presented, where each of the solitons is propagating at equal and opposite angles. As becomes apparent from these plots, irrespective of the phase difference between the solitons, they pass through each other with all their characteristics preserved apart from a small phase shift.

The special property of this kind of collisions that can be computationally exploited arises from the fact that in the case that the second order soliton propagates at zero angle and the first order soliton propagates at an angle of

fridaya@uef.fi

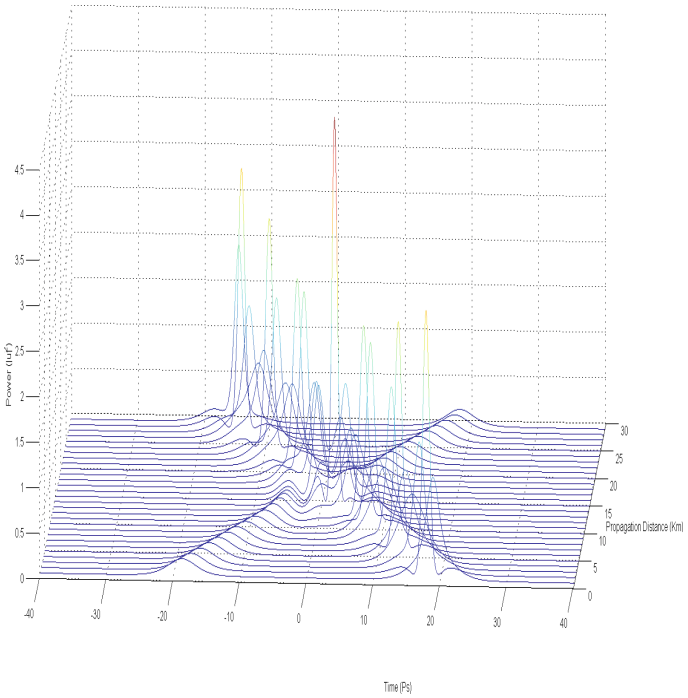


Fig. 5. Collision between a second order and a first order soliton when both are propagating at angles equal and opposite (±0.4). In this simulation the two solitons are in phase (phase values, π and π for the second and the first order solitons respectively).

±0.4, the collision between the two solitons is not perfectly elastic with the special property that a second first order soliton is generated which propagates along side the original second order soliton.

In Figs. 7 and 8 the results of inelastic collisions are presented and the generation of the second first order soliton is shown clearly for the case that the two solitons are in phase and for the case that the two solitons are out of phase. The original second order soliton propagating at zero angle continues to be a second order soliton for the entire length although now possessing a lower energy level and this is true for both phase relationships between the original solitons.

The second first order soliton starts propagating exactly at the collision point and its energy level and phase characteristics are the same as those of the original first order soliton, which takes part in the collision. For different collision angles, slightly different results emerge but the basic principle remains the same for all phase and angle combinations.

Observation shows that: (a) the displacement experienced by the first order soliton generated after the collision depends on the angle of the colliding soliton and, (b) a soliton collision with propagation angle for the colliding first order soliton higher than −0.5 appears to be always perfectly or nearly perfectly elastic.

fridaya@uef.fi

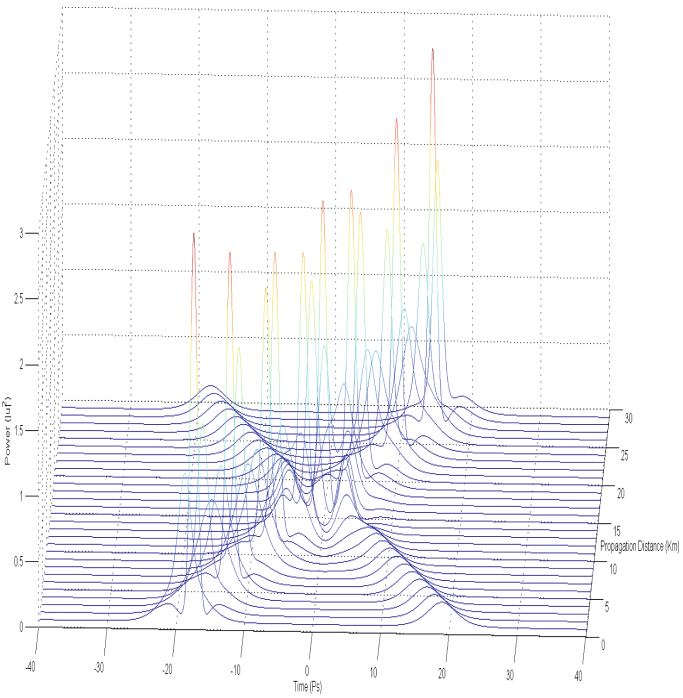


Fig. 6. Collision between a second order and a first order soliton when both are propagating at angles equal and opposite (±0.4). In this simulation the two solitons are out of phase (phase values, 0 and π for the second and the first order solitons, respectively).

For all the computational collisions presented in this paper propagation angles associated with the relative velocities ±0.4 are used since these were found to give potentially useful computational results (inelastic collisions). Thus there will always be at least two solitons at the input end, namely, the injected first and injected second order solitons, while, three solitons will appear at the output end after the collision, that is, the initial first order soliton, the initial second order soliton and the first order soliton generated after the collision. This is a remarkable property possessed only by collisions between first and higher order solitons with the higher order solitons propagating at a 0 angle.

4 Computing with Solitons in Independent Optical Fibres

As shown in the previous section, a first order soliton can be generated and propagate stably alongside an initial second order soliton as a result of that second order soliton colliding with an initial first order soliton. The requirements for achieving this type of collision are: (a) for the two initial solitons to collide at the correct angle, and (b) for the initial first order soliton to start propagating

fridaya@uef.fi

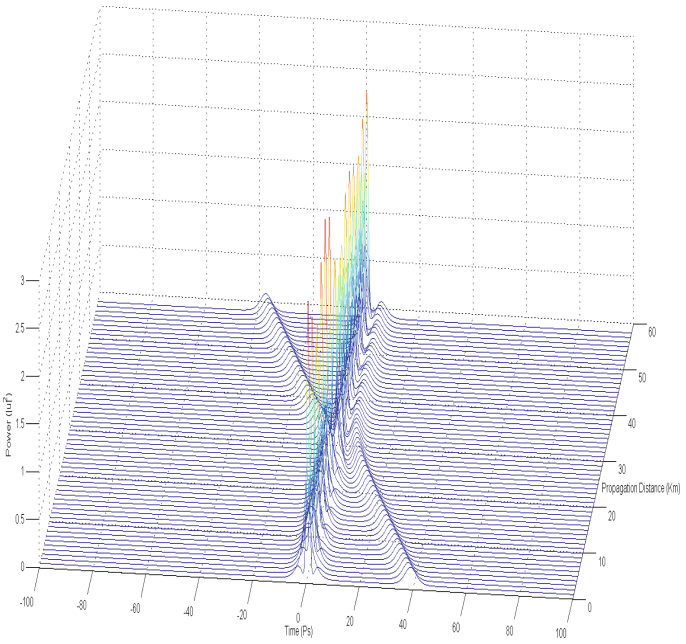


Fig. 7. Collision between a second order and a first order soliton with the second order soliton propagating at a 0 angle. The collision becomes inelastic with a second first order soliton generated as a result of this inelasticity. The two solitons are in phase (phase values, 0 and 0 for the second and the first order soliton respectively).

down the optical fibre to the right or to the left of the initial second order soliton. That way, considerable control over the amount of displacement the generated first order soliton experiences is maintained with it emerging either to the right or to the left of the initial second order soliton, something that encodes a digit 1 or a digit 0 in the overall framework of the computational system presented here (Figs. 9 and 10).

Complementary to the two digits' creation arrangements, is the arrangement at which the collision that generates a digit is balanced by another counter-collision (or, as for the case of the logic gates discussed in subsequent sections, even the overall arrangement of a logic gate). Then, no generated first order soliton is found to propagate along the side of the initial second order soliton either to the right or to the left at any of the output time-slots and no digit is provided by the arrangement at the output end. This is a situation that in the overall computational framework of the system presented here is to be translated as "no output provided" from the specific arrangement at the output end. In other words the situation can be considered equivalent to that of an arrangement capable of "Destroying" the input solitonic pulse and, as a result, of providing no output. The concept is graphically illustrated in Fig. 11.

fridaya@uef.fi

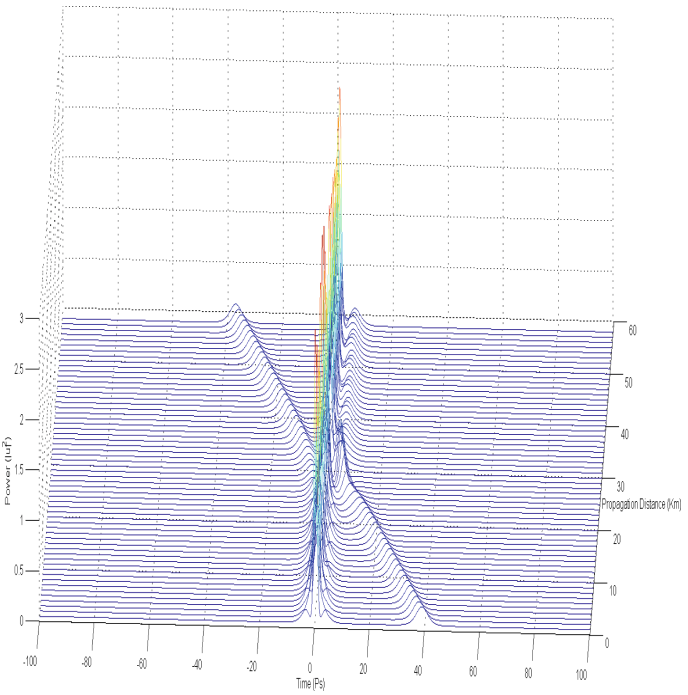


Fig. 8. Collision between a second order and a first order soliton with the second order soliton propagating at a 0 angle. The collision becomes inelastic with a second first order soliton generated as a result of the inelasticity. The two solitons are out of phase (phase values are π and 0 for the second and the first order soliton respectively).

The generation of a first order soliton to the right or to the left of the initial second order soliton is not the only remarkable property of this system. Propagation of the generated first order soliton is also fully steerable in the sense that by subsequent collisions with other first order solitons (control solitons) can be forced to switch propagation time-slot. This remarkable property of the 3NLSE–domain solitons is illustrated in Fig. 12.

Further, Fig. 12 also demonstrates the fact that any displacement of the generated first order soliton, at any level, is completely reversible. That being the case, and because the amount of displacement of the generated soliton with every other collision remains constant, any displacement of a generated first order soliton to the right or to the left of the second order soliton can be completely compensated by an appropriate number of subsequent calculated collisions to the extent where the second order soliton be fully restored to its original state, if this is required by the internal arrangement of a particular computation. The total fibre length required by the system to restore the second order soliton is in every case exactly double the length required for the introduction in the arrangement and the initial displacement of the generated first order soliton. This property of

fridaya@uef.fi

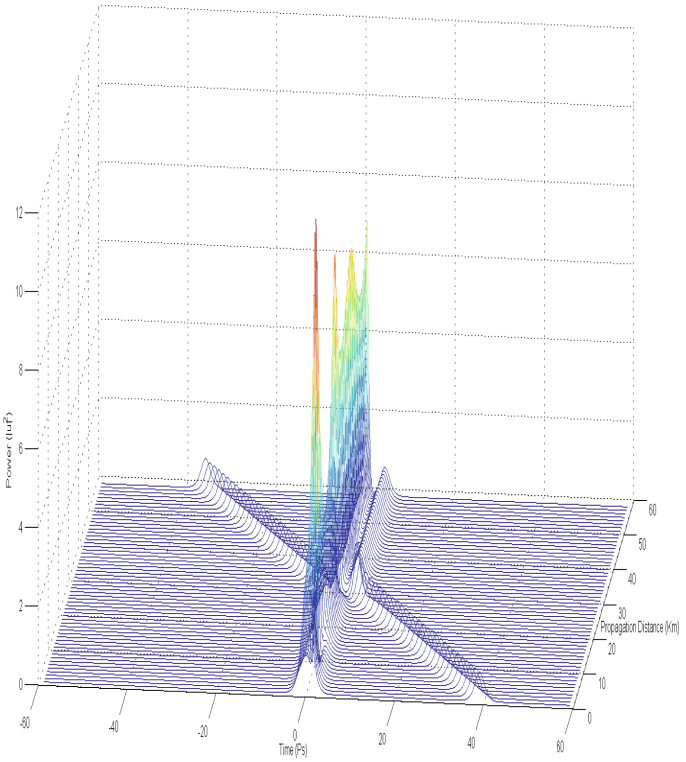


Fig. 9. Generation of the second first order soliton on the right of the initial second order soliton. In the computational system this type of collision represents digit "1".

the system adds a new dimension to the generally accepted idea of reversibility in logic gates, considering that the Toffoli definition of a reversible logic gate is that of an output digit fed back into the output end of a logic gate causing at the input end the generation of the original input digit. As required, this property must be assessed under the light of the undeniable fact that is applicable only to a single digital operation, while the original Toffoli concept of reversibility was intended to apply not only to a single digital operation but also to an entire calculation.

4.1 The Inverter (*NOT* Gate)

It is well known that one of the most important units of any computational system is the *NOT* logic gate. In this section is shown how displacement in time of the generated first order soliton (by taking advantage of the steering property of the system) can effectively represent the two logic states involved in the realisation of any *NOT* logic gate.

The core process of the *NOT* logic gate is based on an arrangement that has been introduced already. A generated first order soliton when starts propagating

fridaya@uef.fi

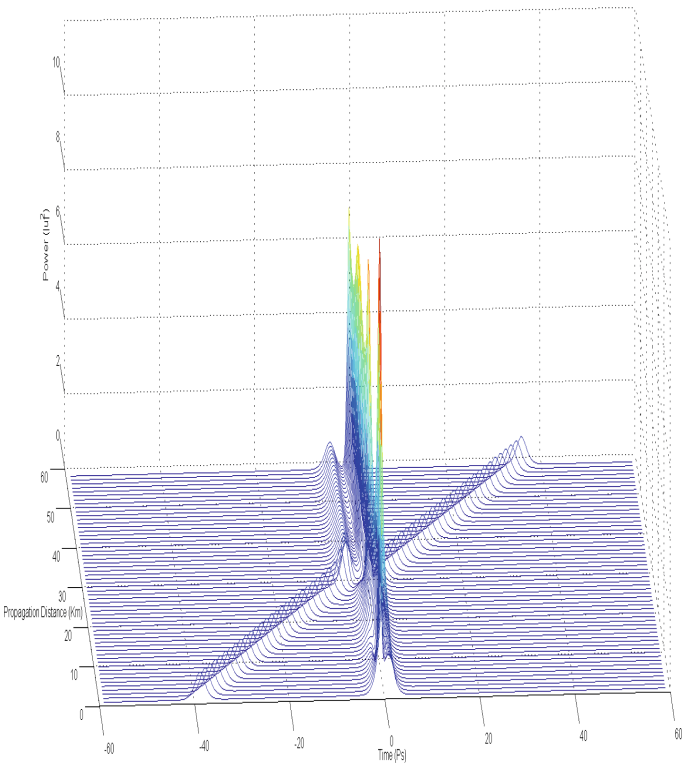


Fig. 10. Generation of a second first order soliton on the left of the initial second order soliton. In the computational system this type of collision represents digit "0".

on the right hand side of the initial second order soliton, as already demonstrated, represents digit 1 in the computational system and the opposite arrangement, with the generated first order soliton starting propagating on the left hand side of the initial second order soliton, represents digit 0. By considering these two primary arrangements and the steering property of the system, as discussed in the previous section, we end up with a computational set similar to the one presented in Figs. 12 and 13. More specifically, the arrangement in Fig. 12 is representing in fact the realisation in this system of a fully enabled NOT logic gate which accepts a logic digit 1 at its input and produces a logic digit 0 at its output. Correspondingly, the arrangement in Fig. 13 represents in this system the realisation of a fully enabled NOT logic gate which accepts at its input a logic digit 0 and produces a logic digit 1 at its output.

As required, the arrangement holds true not only for the initial case but throughout a complete computation as well. This means that an appropriate number of initial logic digits must be injected to the medium before any computation can be carried out and as a result of that any calculation of the total propagation distance required by a particular computation must necessarily include

fridaya@uef.fi

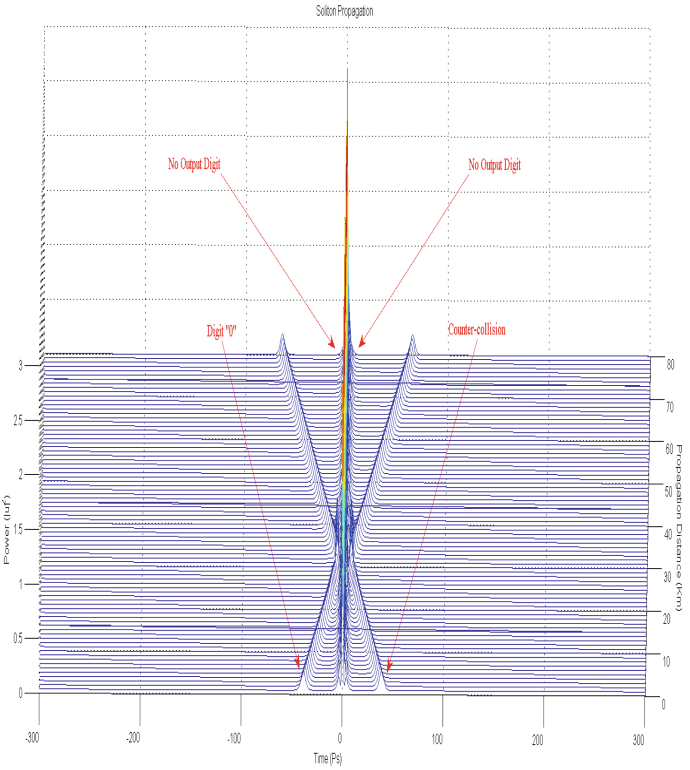


Fig. 11. The input digit "0" along with a counter-collision. The result of the balance in collisions for the overall arrangement is no digit at the output end.

the propagation distance required for the generation of the initial solitons (input digits) at the input end. For convenience we will be referring to this process from now on as the "Bit Generation Gate".

Using this terminology, the arrangement presented in Fig. 10 is the "Bit Generation Gate" for digit 1 and the arrangement presented in Fig. 11 is the "Bit Generation Gate" for digit 0. For a velocity of ± 0.4, the propagation distance needed for the generation of a single digit is constant under all circumstances. In addition to all other properties of the solitonic arrangements discussed so far, it now becomes apparent that they perfectly comply with the requirements set by the classic electronic logic gates as well. The solitonic NOT logic gate presented here is a "one bit input/one bit output" type of a logic gate as is the case with its classic electronic equivalent. Also, the solitonic NOT logic gate is fully cascadable to any level provided the appropriate amount of control solitons for achieving this are present in the arrangement.

fridaya@uef.fi

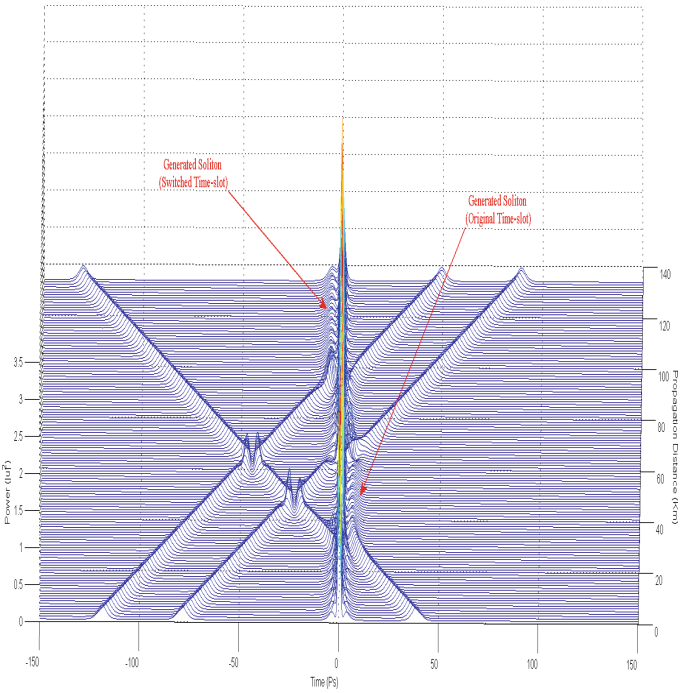


Fig. 12. Generation of a first order soliton to the right of the initial second order soliton and its successive displacement (time-slot switch) by means of using subsequent collisions with another two first order control solitons.

4.2 The *OR* Logic Gate

From the discussion in the previous sections must have now become apparent that two-bit logic gates is not possible to be easily accommodated in the overall computational system if we consider they are implemented on a single optical fibre. This is the case because of the way in which a single bit is represented, that is, as involving one second order soliton per bit generation (it is, however, possible in theory to have two second order solitons within a single fibre simultaneously generating two bits but, as extensive simulations have proven, computational arrangements in this case can have the tendency of becoming rapidly chaotic). For this reason a more straightforward approach would be to simply use one optical fibre per bit, and this is the very way by which we will be arranging solitons for all the logic gates to be presented from now on. In particular, solitons injected into two separate and parallel optical fibres will eventually produce two separate outputs at the end of the arrangement, which when combined together will provide us with the logic gate's overall output.

In order to achieve an *OR* logic gate functionality, the internal workings of the logic gate are simulated be means of using three additional first order control solitons. This combination of first order solitons, along with the initial

fridaya@uef.fi

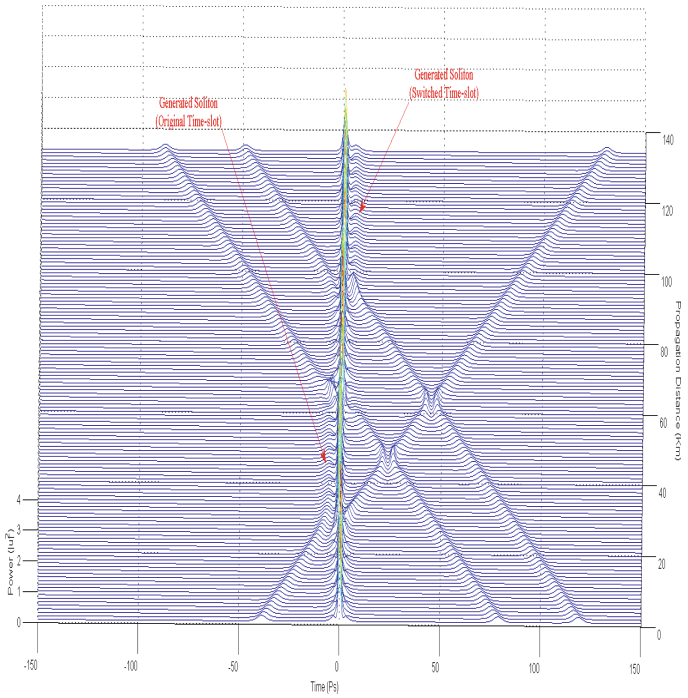


Fig. 13. Generation of a first order soliton at the left of the initial second order soliton and its successive displacement (time-slot switch) by means of using subsequent collisions with another two first order control solitons.

second order soliton and the "Bit Generation Gate", is capable of producing an overall output identical to that of an OR logic gate. Figures 14 and 15 present simulation results for the OR logic gate when the input to the logic gate bits are 1 and 0. The two solitonic arrangements propagate on separate optical fibres and the overall output is considered as produced out of the combination of the two individual outputs.

After all collisions have taken place, and by assessing the two outputs from the two separate optical fibres the system is capable of identifying a generated soliton in the second time slot next to the second order soliton in the first optical fibre (Fig. 14), something that is equivalent to digit 1 at the output. Respectively, no pulse is detected at any time slot around the second order soliton in the second optical fibre (Fig. 15), something that is equivalent to digit 0 at the output. According to the computational scheme the two results must be interpreted by the system as an overall output digit 1. All possible results for an OR gate are summarised in Table 1.

The input/output bit combinations presented in Table 1 can be clearly seen as an OR logic function. In contrast to the control solitons involved in the general realisation of the computational system [2] where different combinations of phase

fridaya@uef.fi

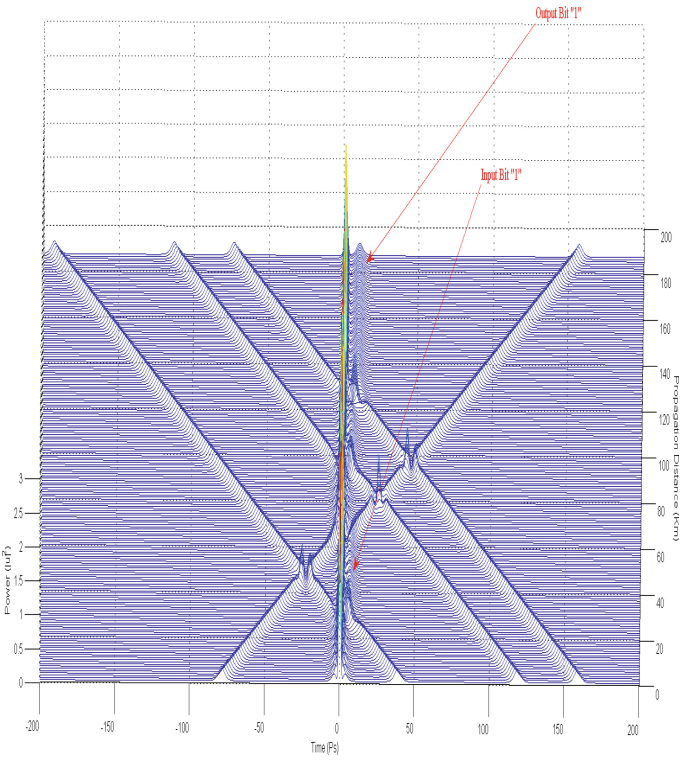


Fig. 14. The *OR* logic gate. The input bit is "1", with the output bit being "1" as well. This arrangement propagates in the first optical fibre.

values have no effect on the final output, the control solitons that had to be used in the *OR* logic gate arrangement were found to require possession of specific phase values. The phase values for digit 1 at the input end are as given in Table 2 and the equivalent phase values for digit 0 are as given in Table 3.

The control solitons have exactly the same propagation arrangement for all the different input combinations as those given in Tables 2 and 3, that is, one control soliton on the left and two control solitons on the right hand side of the central second order soliton.

In relation to all the above stated, one point that requires particular attention is an apparent inconsistency now introduced between the outputs provided from the logic *OR* gate arrangement as presented in Figs. 14 and 15 and those provided from the inverter *NOT* logic gate arrangement as presented in Figs. 12 and 13. Specifically, the inconsistency is in respect to the time slot the generated soliton is recorded at the output end of the logic gates arrangements. While for the inverter *NOT* logic gate the generated first order soliton appears either at the first time slot to the left or to the right of the second order soliton (bits 0 & 1), in the case of the *OR* logic gate the generated first order soliton appears instead

fridaya@uef.fi

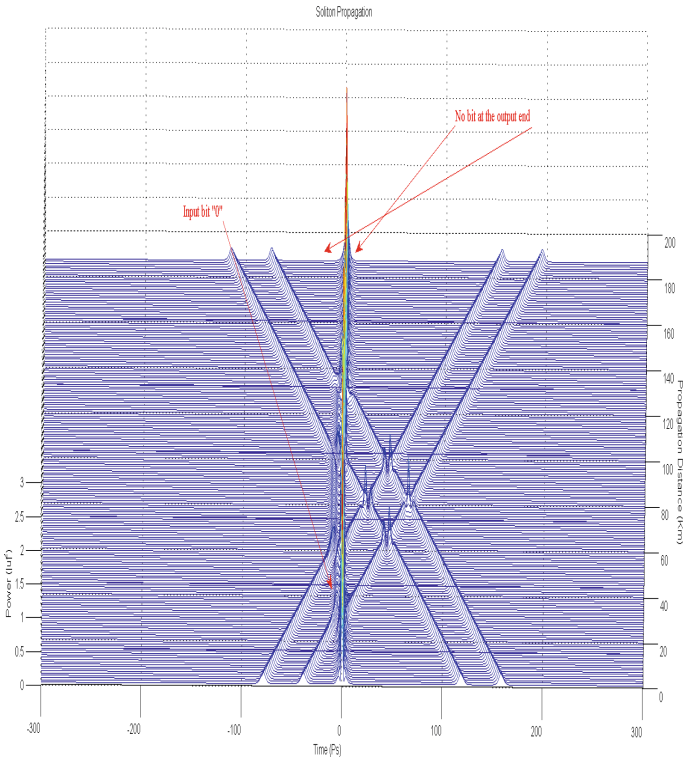


Fig. 15. The *OR* logic gate. The input bit is "0", with the logic gate arrangement providing no bit at the output end. This arrangement propagates in the second optical fibre.

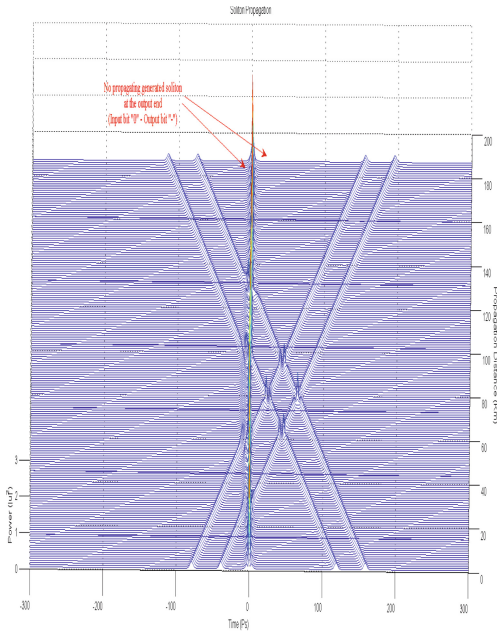
Table 1. All possible input combinations and their corresponding outputs for logic gate *OR*, along with the overall output of the logic gate after combining the two output results. All arrangements are considered propagating in two separate optical fibres.

Logic gate *OR*				
Input I	Input II	Output I	Output II	Overall output
0	0	–	–	0
0	1	–	1	1
1	0	1	–	1
1	1	1	1	1

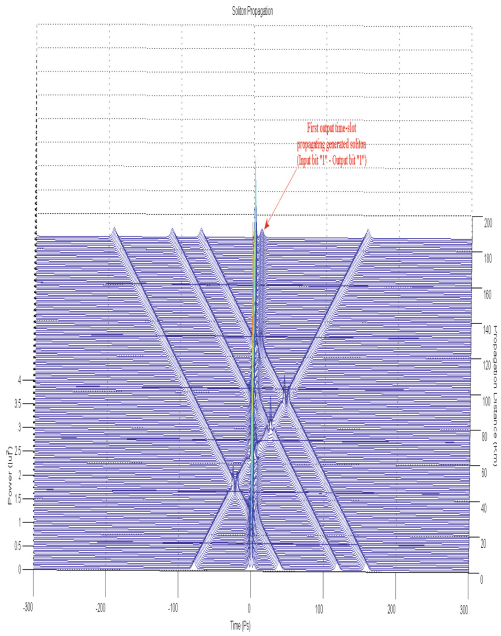
Table 2. The phase values for the control solitons and the second order soliton involved in the *OR* logic gate arrangement for digit "1".

Logic gate *OR* (Digit 1) (Phase values)			
Control Soliton I	Second Order Soliton	Control Soliton II	Control Soliton III
0	0	π	π

fridaya@uef.fi



(a) First Fibre



(b) Second Fibre

Fig. 16. The complete *OR* logic gate arrangement. The input bits are "0" and "1", with the output bits being "–" and "1" respectively. The overall output of the logic gate is "1".

fridaya@uef.fi

Table 3. The phase values for the control solitons and the second order soliton involved in the OR logic gate arrangement for digit "0".

Logic gate OR (Digit 0) (Phase values)			
Control Soliton I	Second Order Soliton	Control Soliton II	Control Soliton III
0	0	0	π

at the second time slot to the left or to the right of the second order soliton (bits 0 & 1). The difference was introduced in the simulations for clarity purposes, acknowledging at the same time the fact that there is also a possibility, for when the computational system is materialised, to be decided for this difference to be deliberately and permanently introduced in the system for the purpose of providing detecting equipment with the capability of a more rapid and error free detection of the output digit. In any case, consistency can be easily restored, if desired, by means of using an extra "correction" first order control soliton.

4.3 The *NAND* and *NOR* Logic Gates

For clearly demonstrating possession of computational universality by a computational system, the system is required to be able to realise $NAND$ and/or NOR logic functionality. In this respect, the $NAND$ logic gate is indeed an important one for every computational system since it can form the undeniable basis of its computational universality. For the computational system presented here, the $NAND$ logic gate can be constructed as a combination of the logic gates discussed so far and by means of using the following well-known property derived directly from De Morgan's law:

$$\overline{A \bigwedge b} \Leftrightarrow \overline{A} \bigvee \overline{b} \tag{4}$$

Thus the $NAND$ logic gate can be realised by using the OR logic gate on input digits that have initially being inverted (individually passed through a NOT logic gate). The overall arrangement for this gate would then be the arrangement of two NOT logic gates followed by the arrangement of an OR logic gate (Fig. 17). For illustration purposes a characteristic output of the $NAND$ logic gate is provided in Fig. 18. Taken together, these results can be summarised as in Table 4.

As was the case with the OR logic gate, if now the output bits are combined into a single optical fibre using a coupler, then the overall result is that of a $NAND$ logic gate. One arm of the coupler would need to introduce a small optical delay since the output pulses to be combined are separated in time by a few picoseconds.

4.4 Complex Computations

As a last remark on the computational scheme considered here, this section discusses the arrangement of a complex computation involving two $NAND$ logic

fridaya@uef.fi

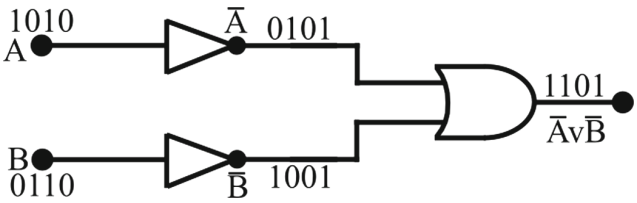


Fig. 17. The $NAND$ logic gate.

Table 4. All possible combinations and results for logic gate $NAND$ with arrangements propagating in two separate optical fibres.

Logic gate $NAND$				
Input I	Input II	Output I	Output II	Overall output
0	0	1	1	1
0	1	1	–	1
1	0	–	1	1
1	1	–	–	0

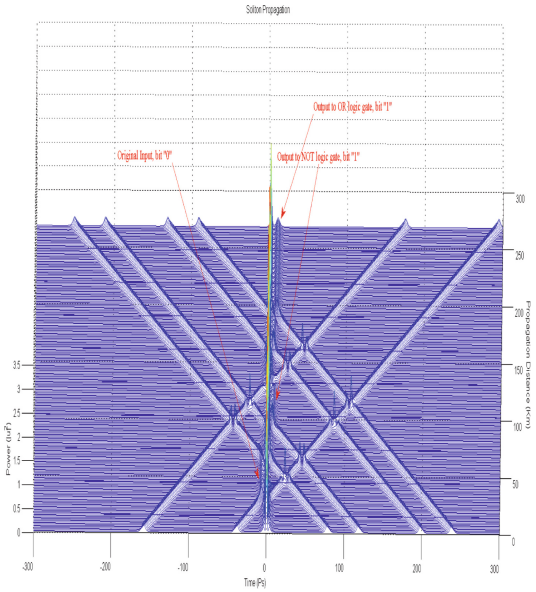
gates. Through investigating and reflecting upon this complex computation the true potential of the computational scheme will be revealed as well as an understanding achieved on how some of the purely theoretical elements of the system can turn practical achieving full realisation of the computational arrangements necessary. The conventional schematic diagram for the complex computation is as shown in Fig. 19.

The soliton arrangement required and the output for the first $NAND$ logic gate ($NAND$ logic gate (A) in the conventional diagram of the computation) is as shown in Fig. 20, with the soliton arrangement and the output for the second $NAND$ logic gate ($NAND$ gate (B) in the conventional diagram of the computation) presented in Fig. 21. More specifically, in Fig. 21a the second order soliton and the soliton arrangement around it represent the generated state as a result of the output provided by the first branch of the first $NAND$ logic gate and the same is true for Fig. 21b in respect to the output provided by the second branch of the first $NAND$ logic gate.

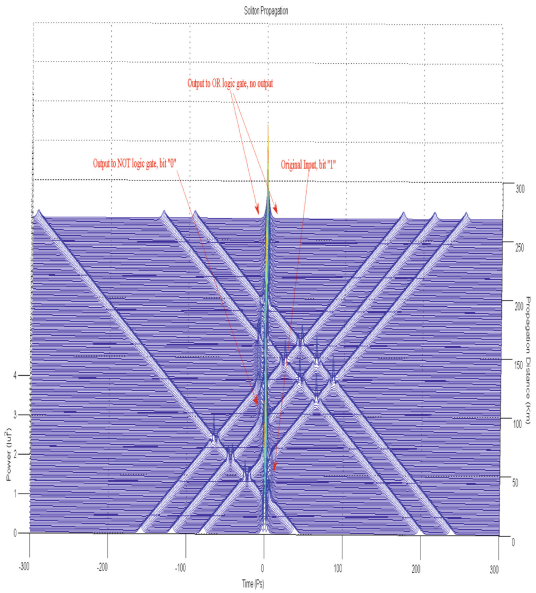
For clarity, the schematic diagram of the first branch (first optical fibre) for the two cascaded $NAND$ logic gates is presented in Fig. 22. The initial input bit values for $NAND$ logic gate (A) in the arrangement are 1 & 1, with the overall output of the two cascaded logic gates being 1, as expected, since effectively the second $NAND$ logic gate in the arrangement reverses the output of the first. In Fig. 23 the schematic diagram of the complete arrangement is presented including both optical fibres along with the input and output bit values.

For when the input bit at point (C) in the arrangement is digit 1 instead, the only variation occurs at the second $NAND$ logic gate. The technique used in this complex arrangement may seem to be a rather strange one at first but

fridaya@uef.fi



(a) First Fibre



(b) Second Fibre

Fig. 18. A characteristic output of the $NAND$ logic gate arrangement. The output digits are "1" and "0", with the input digits being "0" and "1" respectively. For the first fibre arrangement the time-slots for the solitons involved are (from left to right): 160, 40, 0, −80, −120, −200, −240 and the solitons' phase values are: pi, 0, pi, pi, 0, 0, 0. For the second fibre arrangement the time-slots for the solitons involved are (again, from left to right): 160, 120, 80, 0, −40, −200, −240 and the solitons' phase values are: pi, pi, pi, 0, 0, 0, 0.

fridaya@uef.fi

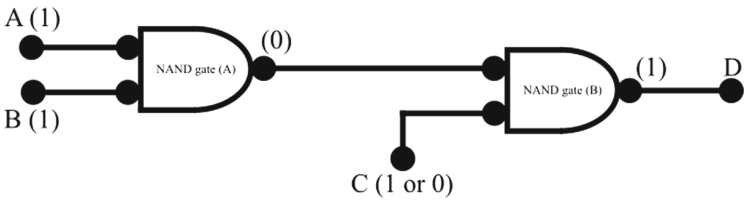


Fig. 19. The conventional schematic diagram for the complex computation involving two cascaded $NAND$ logic gates.

if the way Toffoli gates are cascaded is considered then the similarities become apparent. As a matter of fact, the two techniques are totally equivalent. For comparison reasons the arrangement diagram of two cascaded $NAND$ Toffoli logic gates is presented in Fig. 24.

5 Conclusion

In this paper the framework has been developed and presented for the investigation of non-linear collisions between first order and second order solitons in the 3NLSE–domain with the purpose of realising a computationally universal all-optical computational system using two separate standard communications fibres as the medium for soliton propagation.

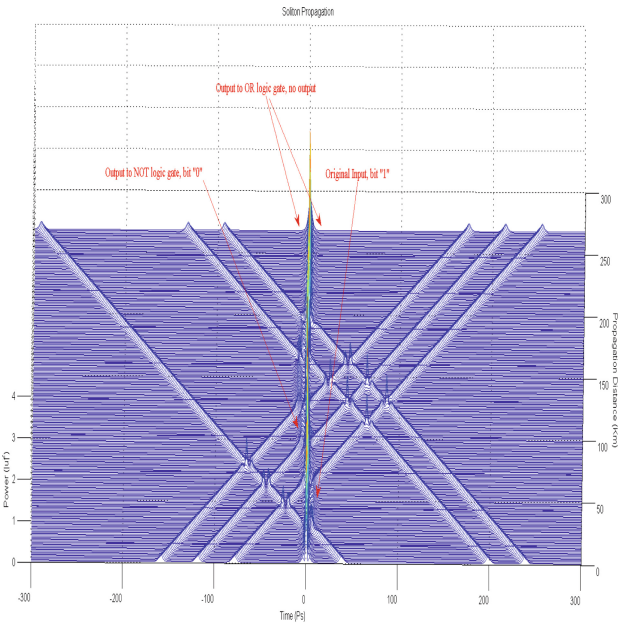
The study has identified an optimum relative velocity difference between solitons for exploiting the use of soliton collisions for computational purposes. In turn, this framework has been applied to the study of novel optical logic devices based on collisions between first order and second order solitons in the 3NLSE–domain. A domain previously considered to be one which could not offer the possibility of computationally useful schemes because of the "oblivious", elastic, nature of collisions between solitons in this domain.

In contrast, it has been shown that the oblivious nature of the collisions can be viewed as an advantage of this domain because the fact that 3NLSE–solitons are unaffected by collisions actually means that solitons representing data and control states alike can be reused in successive computational collisions, offering the possibility of cascading through several computational operations within the same fibre or even achieving outright parallel computations.

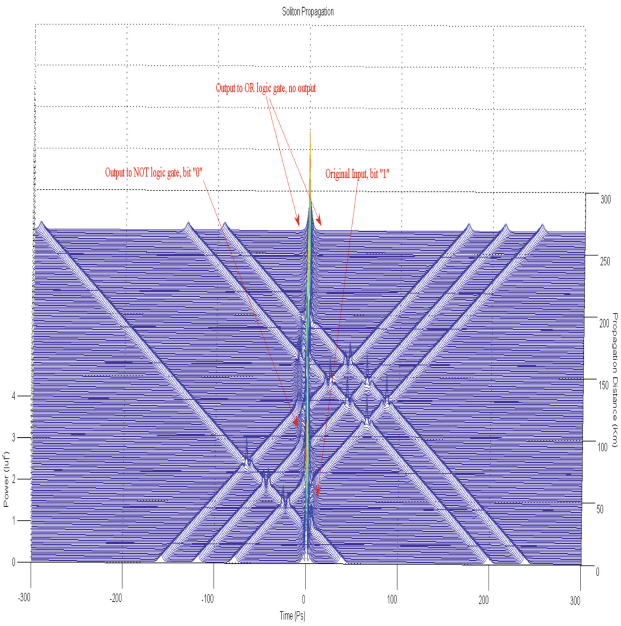
More specifically, this paper suggests that the 3NLSE–domain, can, with suitable hardware arrangements and the application of suitable computation rules, be used to achieve useful computation using collisions between first order and second order solitons. The work presented in this paper also clearly demonstrates that soliton displacement and steering, or soliton "generation" and "destruction" as a state can be used to build useful computation arrangements when second order solitons are involved, something considered previously to be impossible.

The intrinsic limitations imposed by oblivious collisions in the 3NLSE–domain can be overcome by using pulse detection devices to read and identify

fridaya@uef.fi



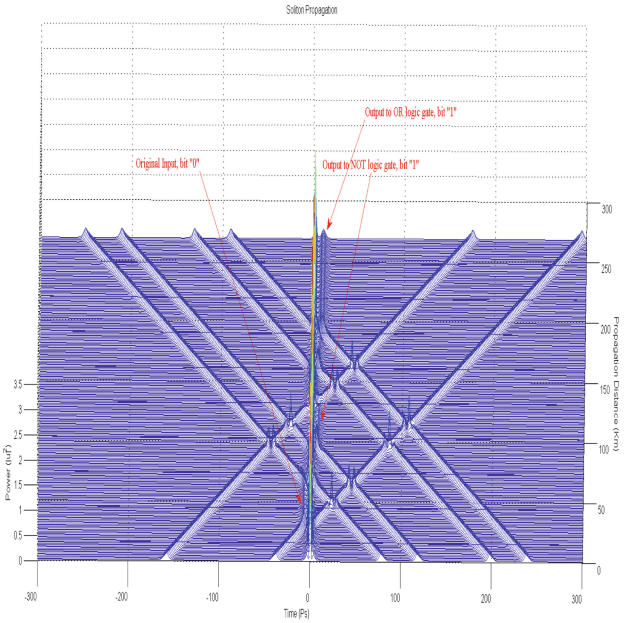
(a) First Fibre



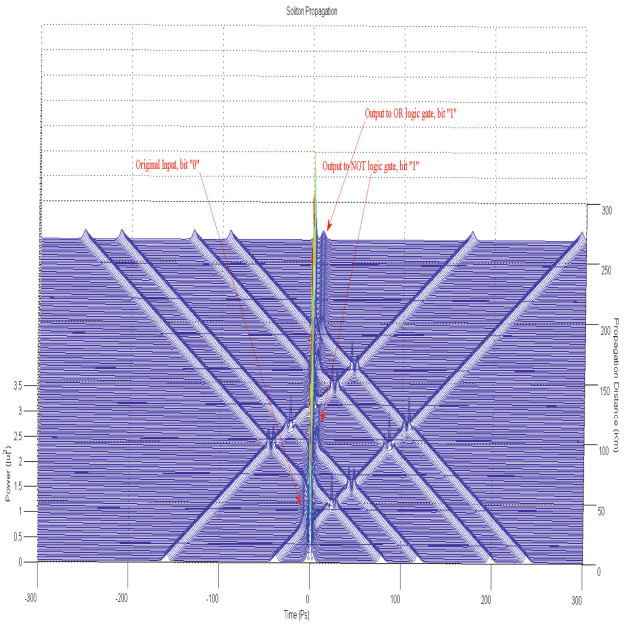
(b) Second Fibre

Fig. 20. A characteristic output of the $NAND$ logic gate arrangement ($NAND$ gate (A) in the conventional diagram). The output digits are "−" and "−", with the input digits being "1" and "1" respectively. The time-slots for the solitons involved are (from left to right): 160, 120, 80, 0, −40, −200, −240 and the solitons' phase values are: pi, pi, pi, 0, 0, 0, 0.

fridaya@uef.fi



(a) First Fibre



(b) Second Fibre

Fig. 21. A characteristic output of the $NAND$ logic gate arrangement ($NAND$ gate (B) in the conventional diagram). The output digits are 1 and 1, with the input digits being 0 and 0 respectively. The time-slots for the solitons involved are (from left to right): 160, 40, 0, −80, −120, −200, −240 and the solitons' phase values are: pi, 0, pi, pi, 0, 0, 0.

fridaya@uef.fi

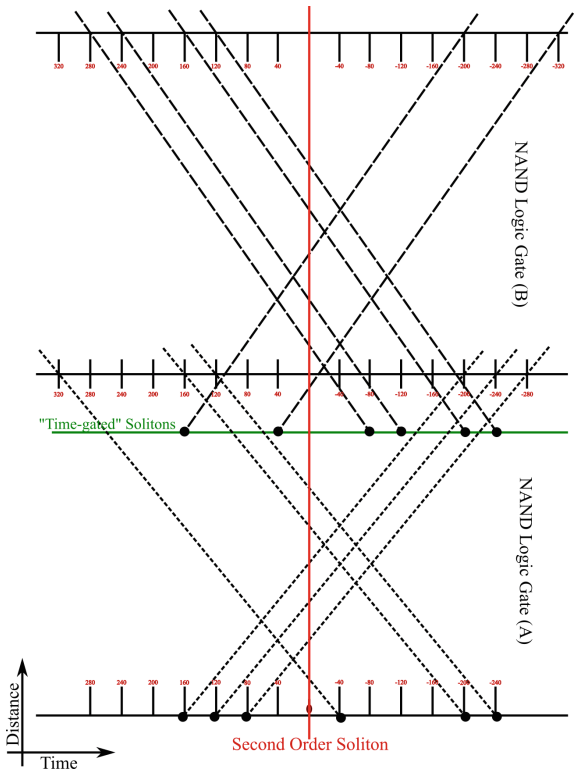


Fig. 22. The conventional schematic diagram for the first branch (first fibre) of the two cascaded $NAND$ logic gates.

pulse envelopes at well-defined points along the fibre. A small fraction of these envelopes would need to be coupled out of the fibre for detection and identification purposes.

The optimum relative velocity for logical gate construction is determined as a result of the simulations carried out; an outcome that helps to build optimum computation arrangements to achieve useful computation. A scheme for achieving an *Inverter* NOT logic gate, an OR logic gate, a NOR logic gate and the very important for every computational system $NAND$ logic gate is described using optimum collisions between first and second order solitons. All these schemes are based on the theoretical Toffoli CN and CCN logic gates and preserve the reversibility of these logic gates. The logic gates are simulated using the physical parameters of currently available optical fibres only. Every complete computation takes place over a distance of several soliton periods and the coupling of pulses into and out of the fibres at various points along their length is required without restricting the capabilities of the overall computation system.

fridaya@uef.fi

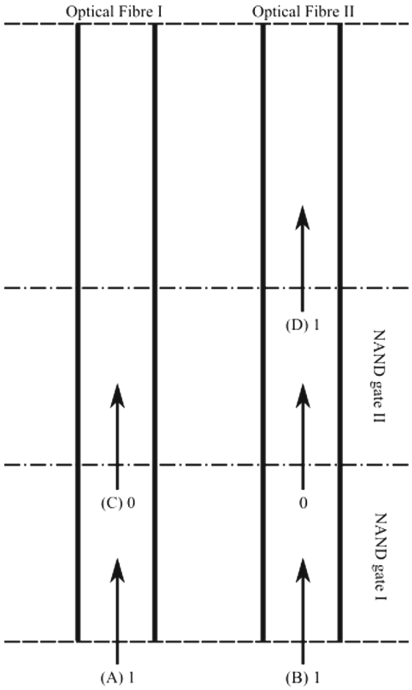


Fig. 23. The schematic diagram of the two cascaded $NAND$ logic gates with the input and output digits shown.

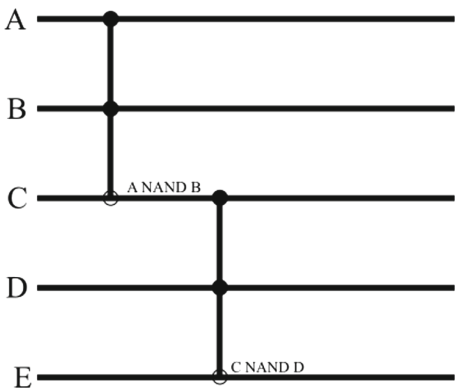


Fig. 24. Two cascaded Toffoli $NAND$ logic gates. The way these logic gates are cascaded is totally equivalent to the way used to cascade the solitonic gates by means of using two independent optical fibres.

fridaya@uef.fi

That the 3NLSE–domain can be used for all-optical computing is also a fundamental result of the work presented in this paper. The extensive investigation of collisions involving higher (second) order solitons from the standpoint of their potential for useful computation constitutes another major contribution to the understanding of the behaviour of solitons under collision situations. This work demonstrates that using higher order solitons, with the computational scheme proposed, it is possible to perform three different logical operations simultaneously using only one logic gate. The development of these logic gates on a physical medium and the investigation of collisions between third, second and first order solitons constitutes a wide area for future investigation.

Acknowledgment. The author would like to thank editors and anonymous reviewers for their valuable and constructive suggestions on this paper.

References

1. Bakaoukas, A.G., Edwards, J.: Computing in the 3NLS domain using first order solitons. Int. J. Unconventional Comput. **5**(6), 489–522 (2009). ISSN 1548-7199
2. Bakaoukas, A.G., Edwards, J.: Computation in the 3NLS domain using first and second order solitons. Int. J. Unconventional Comput. **5**(6), 523–545 (2009). ISSN 1548-7199
3. Bakaoukas, A.G.: An all–optical soliton FFT computational arrangement in the 3NLSE-domain. In: Unconventional Computation and Natural Computation Conference, UCNC 2016, Manchester, UK, 11–15 July 2016 (Conference Proceedings). Lecture Notes In Computer Science. Springer, Cham (2016). ISBN 978-3-319-41311-2
4. Bakaoukas, A.G.: An all-optical soliton FFT computational arrangement in the 3NLSE-domain (Extended). Natural Comput. J. (2017). https://doi.org/10.1007/s11047-017-9642-1. Print ISSN 1567-7818, Online ISSN 1572-9796
5. Agrawal, G.P.: Non-Linear Fiber Optics. In: Quantum Electronics–Principles & Applications. Academic Press (1989). ISBN 0-12-045140-9
6. Ablowitz, M.J., Segur, H.: Solitons and the Inverse Scattering Transform. SIAM, Philadephia (1981)
7. Hasagawa, A., Kodama, Y.: Solitons in Optical Communications. Oxford University Press, Oxford (1995)
8. Toffoli, T.: Reversible computing. In: de Bakker, J. (ed.) Automata, Languages, and Programming. Springer, New York (1980)
9. Fredkin, E., Toffoli, T.: Conservative logic. Int. J. Theor. Phys. **21**, 219–253 (1981)
10. Steiglitz, K.: Time-gated Manakov spatial solitons are computationally universal. Phys. Rev. E **63**, 016608 (2000)
11. Zhang, M., Wang, L., Ye, P.: All optical XOR logic gates: technologies and experiment demonstrations. IEEE Opt. Commun. (2005)
12. Froberg, S.R.: Soliton fusion and steering by the simultaneous launch of two different-colour solitons. Opt. Lett. **16**, 1484–1486 (1991)
13. Hasegawa, A., Tappert, F.D.: Transmision of stationery non-linear optical pulses in dispersive dielectric fibres. I. Anomalous dispersion. Appl. Phys. Lett. **23**, 142–144 (1973)

fridaya@uef.fi

14. Hasegawa, A., Tappert, F.D.: Transmision of stationery non-linear optical pulses in dispersive dielectric fibres. II. Normal dispersion. Appl. Phys. Lett. **23**, 171–172 (1973)
15. Zabusky, N.J., Kruskal, M.D.: Interaction of 'Solitons' in a collisionless plasma and the recurrence of initial states. Phys. Rev. Lett. **6–9**(15), 240–243 (1965)
16. Herman, R.L.: Soliton propagation in optical fibres (1992). Article that appeared in American Scientist July–August 1992
17. Lax, P.D.: Integrals of non-linear equations of evolution and solitary waves. Pure Appl. Math. **21**, 467–490 (1968)
18. Gardner, C.S., Greene, C.S., Kruskal, M.D., Miura, R.M.: Method for solving the Korteweg-De Vries equation. Phys. Rev. Lett. **19**, 1095–1097 (1967)
19. Zakharov, V.E., Shabat, A.B.: Exact theory of two dimensional self-focusing and one dimensional self-modulation of waves in non-linear media. Zh. Eksp. Teor. Fiz. **61**, 118–134 (1972). Sov. Phys. JETP, 34, pp. 62–69
20. Haus, H.A., William, W.: Solitons in optical communications. Rev. Mod. Phys. **68**(2), 423–444 (1996)
21. Snyder, A.W., John Mitchell, D.: Big incoherent solitons. Phys. Rev. Lett. **80**, 1422–1424 (1998)
22. Mollenauer, L.F., Stolen, R.H., Gordon, J.P.: Experimental observation of picosecond pulse narrowing and solitons in optical fibres. Phys. Rev. Lett. **45**, 1095–1098 (1980)
23. Kivshar, Y.S., Luther-Davies, B.: Dark optical solitons: physics and applications. Physics Reports **298**, 81–197 (1998)
24. Preitschopf, C., Thorn, C.B.: The Backlund transform for the Liouville field in a curved background. Phys. Lett. **250B**, 79–83 (1990)
25. Ghafouri-Shiraz, H., Shum, P., Nagata, M.: A novel method for analysis of soliton propagation in optical fibres. IEEE J. Quantum Electron. **31**(1), 190–200 (1995)
26. Franken, P.A., Hill, A.E., Peters, C.W., Weinreich, G.: Phys. Rev. Lett. **7**, 118 (1961)
27. Benney, D.J., Newell, A.C.: The propagation of non-linear envelopes. J. Math. Phys. **46**, 133–139 (1967). (Name changed to: Studies in Applied Mathematics)
28. Goodman, J.W., Leonberger, F.I., Kung, S.Y., Athale, R.A.: Optical interconnections for VLSI systems. In: Proceedings of the IEEE, vol. 72, pp. 850–866 (1984)

fridaya@uef.fi

Intra-channel Interference Avoidance with the OGFDM Boosts Channel Capacity of Future Wireless Mobile Communication

Mohammad R. Kadhum(✉), Triantafyllos Kanakis, and Robin Crockett

Faculty of Arts, Science and Technology, University of Northampton, Northampton, UK
mohammad.kadhum@northampton.ac.uk

Abstract. The Orthogonal Generalized Frequency Division Multiplexing (OGFDM) with Intra-Channel Interference Avoidance (ICIA) approach is, for the first time, proposed, explored and evaluated. Since the interference manipulation currently represents a hot topic for wireless mobile communication, the conventional approach of mitigating the interference is no longer acceptable. As a result, a novel method for addressing the interference between adjacent filtered sub-carriers (in-phase/out-phase) is comprehensively investigated herein. The proposed approach utilises the oversampling factor to effectively avoid interference and improve the quality of service of affected filters under bad transmission states. Thus, this supportive method which is essentially aware of propagation conditions is employed for removing the roll-off (α) impact yet improving the level of Bandwidth (BW) efficiency for applied filters of the OGFDM waveform. Besides, in terms of the system performance, the trade-off relation between the channel capacity and the key Hilbert filter parameter is theoretically and practically discussed. This requires investigation of the influence of α factor on the maximum achieved bit-rate at the acceptable limit of the Bit Error Rate (BER). A MATLAB simulation was introduced to test the performance characteristics of the proposed system in the physical layer (PHY) of an electrical back-to-back wireless transmission system.

Keywords: Orthogonal Generalized Frequency Division Multiplexing (OGFDM) · Intra-Channel Interference Avoidance (ICIA) · Bit Error Rate (BER) · Hilbert filter · Roll-off factor · Oversampling factor · Wireless mobile communication · Channel capacity · Physical layer (PHY)

1 Introduction

The rapid growth in mobile networks and big limitation of the current mobile waveform have pushed the cellular networks developers to explore new transmission technologies that can improve the channel capacity and deliver a more reliable service to the user in comparison to the present 4G standard [1]. Thus, due to

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 281–293, 2019.
https://doi.org/10.1007/978-3-030-22868-2_21

fridaya@uef.fi

the key drawbacks of the Orthogonal Frequency Division Multiplexing (OFDM) [2], new candidate waveforms like the Filter Bank Multi-Carrier (FBMC) [3], the Universal Filter Multi-Carrier (UFMC) [4], the Filtered OFDM (F-OFDM) [5], and the Generalized Frequency Division Multiplexing (GFDM) [6] are proposed to the future scenarios of the next mobile generation (5G).

Moreover, lately, a new candidate waveform termed as Orthogonal Generalized Frequency Division Multiplexing (OGFDM) [7] is presented with extra supportive facilities to mobile applications of the 5G and beyond. Hence, the recently introduced waveform can achieve a dual channel capacity yet keeping the same level of error relative to the most candidate waveform GFDM.

It's worth noting that the need for a higher channel capacity pushes the networks operators to upgrade the usage of the Bandwidth (BW) by increasing the number of employed frequency sub-carriers. On the other hand, the increased number of frequency sub-carriers can raise the level of interference, particularly, with bad transmission conditions [8].

Due to the predicted scenarios of the future generation of mobile, particularly, higher channel capacity (Gb/s), the conventional management for the induced interference of wireless mobile networks is no longer applicable [9]. The main reason beyond this that the traditional approach can mitigate the interference with wasting the BW efficiency of affected frequency sub-carriers [10]. Hence, depending on the traditional approach, the delivered amounts of channel capacity are decreased severely after mitigating the affected sub-carriers.

As a result, the new candidate waveforms, especially, the best candidate one "OGFDM" needs to a modern manipulation for the interference makes it able to cope with the future market demands of mobile communications [11].

In this paper, a new approach called Intra-Channel Interference Avoidance (ICIA) for removing the interference between the adjacent filtered sub-carriers of the OGFDM yet improving the efficiency of the BW is investigated.

The proposed scheme employs the oversampling operation to efficiently avoid any probable bad behaviour of the α factor for utilised filters. Thus, by controlling the oversampling factor, new guard intervals are generated between the filtered sub-carriers to accommodate optimally the expanded cases of the α factor.

In addition, the organized distances between frequency centres ($f_c s$) of the employed filters are fixed optimally. Hence, assure, primarily, that all Nyquist intervals of sub-carriers and their filters are mostly secured and followed a similar allocation procedure for variant $f_c s$ [12].

Also, under a standard level of the transmission circumstances (good signal power), the transmitted data are distributed and carried by different $f_c s$, then to be received sequentially by the counterparts of them [13].

Furthermore, the introduced approach is basically achieved in the physical layer (PHY), where the system can manage the assigned value of the oversampling factor in accordance with the α parameter of the filter.

Regarding the changeable conditions of the transmission, the trade-off relationship between the overall channel capacity and the α factor of filter is investigated to identify the optimal guidelines of filter design under various transmission

fridaya@uef.fi

status. Hence, explore the system performance of the transmitted signal which may suffer from power loss during its travel to the receiver to ensure a secured transmission for all situations.

The level of power attenuation is varied to optimal, acceptable and severe due to the conditions of transmission. As such, the possibility of interference occurrence between rolled-off filters is increased if the level of filter expansion is enlarged due to the decreased average of transceiver power [14].

The rest of the paper is structured as follows: Sect. 2 discusses theoretically the main advantages and disadvantages of the proposed system in an electrical back-to-back wireless mobile system. Section 3 evaluates the transmission performance in terms of the channel capacity and Bit Error Rate (BER) of the OGFDM transceiver utilising a MATLAB simulation code. Section 4 summarises the key outlines of the paper.

2 System Model

As is shown in Fig. 1, in the transmission operation, firstly, each sub-carrier in the OGFDM system is digitally encoded. Then, the complex numbers are produced by employing one of the most popular modulation formats, like BPSK, 16 QAM, 128 QAM, etc.

After that, in the ICIA part, where important manipulations that directly related to the oversampling operation are achieved, each modulated sub-carrier is up-sampled dynamically either by K or $2K$ factor depending on the applied system (conventional/dual). Hence, insert a set of zeros, equivalents to the utilised factor of oversampling, between every two adjacent samples of original data.

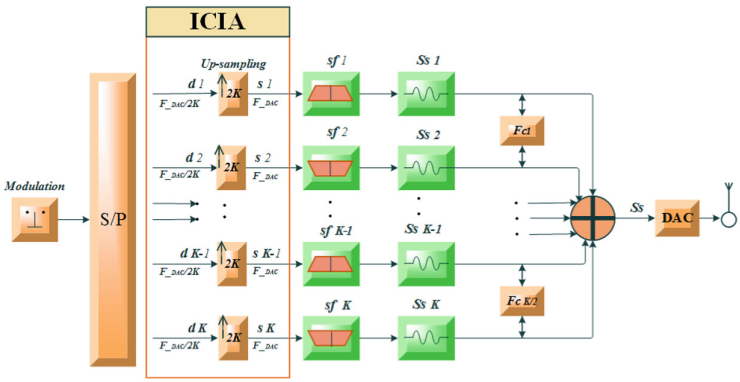


Fig. 1. Transmitter side of the OGFDM with ICIA.

Subsequently, the up-sampled sub-carriers are passed through Hilbert shaping filters. As a result, every two adjacent sub-carriers are filtered orthogonally

fridaya@uef.fi

and assigned to a similar fc. Hence, every single fc, can be shared between two different sub-carriers simultaneously. Thus, accommodate an extra number of filtered sub-carriers (double) for the same BW.

In addition, to ensure optimum performance for the filtering process, key shaping filter parameters like the coefficients set are adjusted precisely in the PHY.

Having been filtered digitally, different convoluted sub-carriers from various $f_c s$ are then combined utilising the adder in the digital domain. The out coming sequence of data is then pushed through a digital-to-analog converter (DAC) generating the analog signal. Finally, the electrical analog signal is propagated through the antenna.

As is clear in Fig. 2, in the reception operation, the received analog electrical signal is firstly digitized employing an analog-to-digital converter (ADC). Thereafter, the stream of data corresponding to each sub-carrier in the OGFDM system are de-multiplexed using equivalent matching filters.

To achieve a successful filtering operation for all employed matching filters, the matching coefficients are similarly adjusted to their shaping counterparts.

After matching and de-multiplexing process of the received sub-carriers via the digital filters, the up-sampled sub-carriers are then down-sampled in the ICIA part, where, oversampling operation is manipulated by removing all previously embedded zero samples. Ultimately, during the de-modulation process, the original data is retrieved utilising a previously decided modulation format.

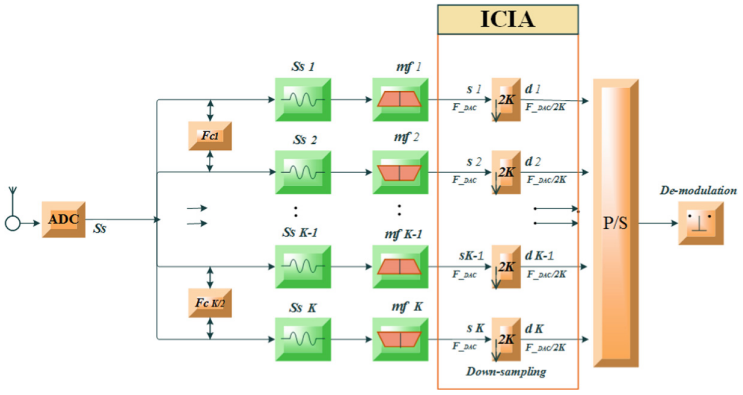


Fig. 2. Receiver side of the OGFDM with ICIA.

Mathematically speaking, the key required operations of the proposed transceiver system are demonstrated. Thus, on the transmitter side, the impulse responses of the h^{th} shaping pairs of Hilbert filters are written as follows [15]:

$$\begin{aligned} sf_h^A(t) &= g(t)\cos(2\pi f c_h t) \\ sf_h^B(t) &= g(t)\sin(2\pi f c_h t) \end{aligned} \tag{1}$$

fridaya@uef.fi

where fc_h represents the centre frequency of the h^{th} Hilbert filters pairs, and $g(t)$ is the baseband pulse which is expressed as follows [15]:

$$g(t) = \frac{\sin[\pi(1-\alpha)\gamma] + 4\alpha\gamma\cos[\pi(1+\alpha)\gamma]}{\pi\gamma[1-(4\alpha\gamma)^2]} \tag{2}$$

where $\gamma = t/\Delta t$, α parameter supervises the excess of filter BW and Δt is corresponding to the sampling interval before the oversampling. In addition, the in-phase and quadrature-phase of the Hilbert filter pair components are denoted by the superscripts A and B respectively. It's worth noting that each of those components is utilised independently to convey a sub-carrier. Moreover, on the receiver side, a version of impulse responses of matched filters for the corresponding shaping filters of Hilbert-pair are written as follows [15]:

$$\begin{aligned} mf_h^A(t) &= sf_h^A(-t) \\ mf_h^B(t) &= sf_h^B(-t) \end{aligned} \tag{3}$$

To recover the signal at the receiver side, convolution applied between shaping and matching filters achieves the following [15]:

$$sf_j^C(t) \otimes mf_j^D(t) = \left\{ \begin{array}{ll} \delta(t-t_0) & \text{if C = D and j = i} \\ 0 & \text{if C} \neq \text{D or j} \neq \text{i} \end{array} \right\} \tag{4}$$

where t_0 represents the probable delay time, the superscripts C and D indicate either in-phase or quadrature-phase level, and the subscripts i and j refer to the position of the fc.

To support the normal case of employing K filtered sub-carriers in the proposed system, the value of the oversampling factor (OV) is set to be equivalent to the total number of applied sub-carriers ($OV = K$). Furthermore, considering that the sampling speeds of the ADC and the DAC are identical, the decision of central frequency allocation for each pair of Hilbert filters is determined.

In this context, frequency responses of Hilbert filters are optimally distributed due to the perfect selection for each fc utilised in the spectral region of the obtainable DAC/ADC. Thus, filters accommodation process is performed orthogonally and sequentially considering both phases of each fc.

The key measurement value for representing the BW excess of each filtered sub-carrier is the α factor. Hence, whenever the α factor is increased, that means, the drop in the amplitude or power is raised and the level of interference between adjacent filters is growing.

To explain more about the introduced process, it is seen from Fig. 3 that each pair of orthogonal filters works on two copies of oversampled sub-carrier. For example, for each specific fc of filters' pair, the first copy of the first oversampled sub-carrier is taken by the in-phase filter (cosine filter) and the second one of the second oversampled sub-carrier is taken by the out of phase filter (sine filter) simultaneously. Consequently, odd and even sub-carriers are orthogonally filtered by the in-phase and out-phase filters respectively.

fridaya@uef.fi

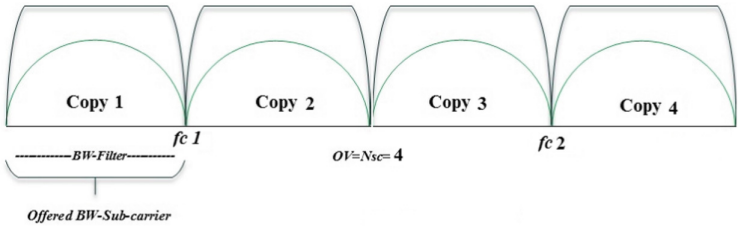


Fig. 3. Normal oversampling for 4 filtered sub-carriers.

The BW size of utilised filters F_{BW} depends on two major parameters. Firstly, the offered sampling rate of the employed sub-carrier Sub_S and secondly, the value of BW excess (α factor) of applied filters. The relation between the allocated filter BW and these two key factors is expressed as follows [16]:

$$F_{BW} = Sub_S * (1 + \alpha). \tag{5}$$

where $1 >= \alpha >= 0$.

Regarding the first significant factor, the BW of each filter varies according to the size of an individual copy in an oversampled sub-carrier. Such a variation is due to change in the configuration of oversampling operation, hence, size of the filter BW decreased whenever the value of the OV is increased and vice-versa.

To explain more about this, it is clear in Fig. 4 that if the specified BW for each convoluted sub-carrier is less than originally afford BW by the system, the aggregated channel capacity is affected badly, particularly, utilising a similar fc for different oversampling cases. Hence, compared to the optimally obtained sub-carrier BW, $(BW = F_{DAC}/OV)$, where the OV is equivalent to the number of applied sub-carriers K, the outcome BW for each filtered sub-carrier after applying the dual value of the OV will be reduced by a half, $(BW = F_{DAC}/OV, where, OV = 2K)$. As a result, the overall channel capacity is decreased according to Shannon theorem as follows [16]:

$$Cap_{ICIA} = (BW/2) * log_2(1 + SNR) \tag{6}$$

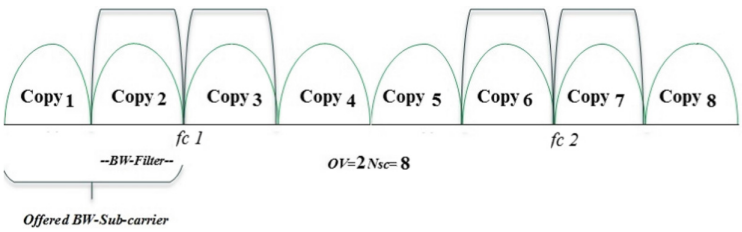


Fig. 4. Dual oversampling for 4 filtered sub-carriers.

fridaya@uef.fi

Regarding the second noteworthy factor, if α factor has a zero value, the allocated BW of each utilised filter is equal to the offered sampling rate of each employed sub-carrier. Thus, the produced brick wall filter for both in-phase and out-phase levels has no interference with its adjacent filters which is considered as an ideal case [17].

On the other hand, setting α value to its maximum (α =1) leads to double the occupied size of each applied filter. Hence, increase the interference between them, especially, with normal oversampling state [18].

From the time-domain perspective, oversampling operation does not affect transmission time specified for each OGFDM iteration absolutely. Hence, multiply the number of samples by X does not increase the overall delay time for each symbol since the smallest duration between any two adjacent samples will be divided by X [19].

For example, the calculated time T_1 for N transmitted samples of an individual OGFDM iteration before oversampling process is, $T_1 = N * dt$, where dt represents the closest duration between any two adjacent samples. After oversampling the sub-carrier by X, the counted time T_2 for transmitting NX samples of an individual OGFDM symbol is, $T_2 = NX * dt/X$, which is consequently equivalent to T_1.

To ensure that the convolution between an oversampled sub-carrier and the applied filter is achieved perfectly, the frequencies for both filter and sub-carrier must be identical [20]. Hence, the frequency for each adopted sample of an individual sub-carrier (f_i) is calculated accurately and then to be considered as a first input to the convolution process. The second input is the calculated frequency for the corresponding selected sample of the applied filter (f_j). If frequencies for samples of both oversampled sub-carrier and utilised filter are similar, ($f_i = f_j$), the convolution is achieved successfully.

Occasionally, the variance between any two adjacent taps in the applied filter is rather large, thus, the most common solution to come up with this issue is by increasing the number of utilised taps of the filter. The calculated frequency for a new tap achieved between two adjacent frequencies of the utilised filter (f_j, f_{j+1}) will be convoluted as well with its corresponding frequency of a sample of an employed sub-carrier.

Such type of treatment improves the BER efficiency of convolution process due to the raised resolution of the filter to recognize the intended in-between frequencies for samples of OGFDM sub-carrier. That, however, raises the computational complexity due to an increased number of multiplications of the filtration process.

On the other hand, the proposed approach which raises the oversampling ($OV = 2K$), can handle this problem utilising advantages of the enlarged α and benefit from the high flexibility in expansion between any two adjacent filters. Thus, the considerable relation between the doubled OV and the expanded value of α factor at accepted limits of errors results in mitigating the variance among the filter coefficients. Hence, whenever the assigned value of α is increased, a passband ripple of the configured filter is decreased and vice-versa.

fridaya@uef.fi

As a result, the increased degrees of the α factor ($\alpha > 0$) can address this issue without the need to promote the resolution of the utilised filter.

In contrast, with the minimum oversampling ($OV = K$), increase the value of the α comes up with a side effect. That's mean, further enlarged α factor can improve the BER performance of the filter's passband ripples but cause in an extra rise of interference between applied filters.

3 Experimental Work

To explore performance characteristics of the proposed OGFDM with ICIA system in terms of channel capacity and BER, a numerical simulation is extensively undertaken in the PHY of an electrical, digital, and wireless transmission system.

Through this paper, the utilised modulation formats for the frequency sub-carriers vary from the low formats like the BPSK to the higher schemes as the 256 QAM which is proposed recently to the future mobile technology [21,22].

Besides, depending on the employed approach for interference manipulation, the oversampling factor is mainly assigned for two values, ($OV = K$) for the conventional way, and ($OV = 2K$) for the proposed method (ICIA), where K refers to the number of applied sub-carriers.

Unless explicitly stated, four digital filters with their convoluted sub-carriers are considered. In addition, to simulate the generation and detection process for each filtered sub-carrier in the OGFDM system, the approach mentioned in [7] is adopted. Also, a sampling speed (DAC/ADC) equivalents to 4 GHz is applied.

According to the trade-offs relation between aggregated channel capacity of the OGFDM system and the flexibility of filter design, the key guidelines of the optimum filter scheme are recognised. As such, the filter BW access can be assigned to different values (dynamic range) through this simulation but it is initially fixed at its optimal value, ($\alpha = 0.1$).

Moreover, the co-existence between the α value of the overlapped frequency, and the weakened ripples of the frequency response rises the opportunity of achieving an optimum value of α, which, in turn, corresponds to a maximum channel transmission capacity.

To explain the key advantages and disadvantages of employing the ICIA, the overall channel capacity versus α filter parameter is examined for two adopted oversampling values, the normal one ($OV = 4$) that is used for the traditional processing of interference and the doubled method ($OV = 8$) which is employed for the developed scheme of interference manipulation (ICIA).

As is shown in Fig. 5, in comparative with the stable channel capacity of the doubled oversampling case (DOV), the fluctuations in channel capacity for normal oversampling case (NOV) are classified here into three main zones.

First is the 'Green Zone' (GZ), which represents all channel capacities above the stable line of channel capacity for the DOV. Thus, the achieved channel capacities for the NOV are better than the obtained gained DOV channel capacities. The GZ capacities are calculated for all α values of the period between 0 and 0.3, where ($0 <= \alpha <= 0.3$).

fridaya@uef.fi

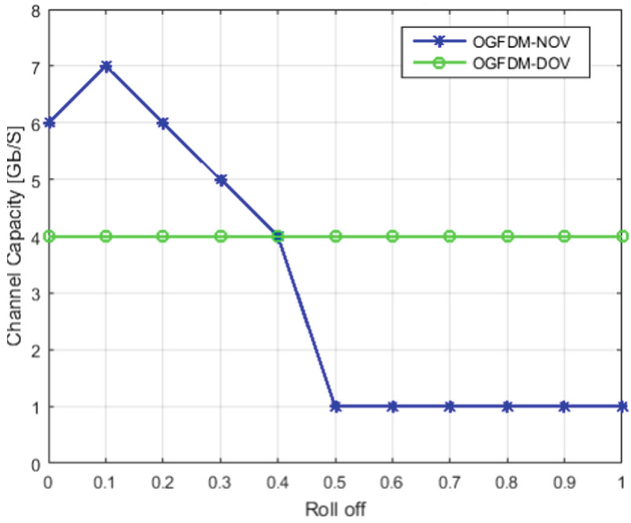


Fig. 5. Channel capacity of OGFDM with NOV/DOV for the entire range of α

Second is the 'Yellow Zone' (YZ), which is counted for the α value of 0.4. In this neutral area, the channel capacities for both the NOV and DOV are equivalent. From oversampling perspective, there are two logical reigns inside the YZ (safe and alarm). Hence, employing the NOV in the YZ puts the system under the alarm reign, while utilising the DOV promotes it to the safe reign. Thus, exceed the accepted limits of the GZ and working in YZ of the NOV results in ringing the alarm system to avoid any probable increase in the α factor (alarm status).

Third is the 'Red Zone' (RZ), which is introduced for all channel capacities under the stable line of the DOV channel capacity. Thus, it basically refers to all α values between 0.5 and 1, where ($0.5 <= \alpha <= 1$). It's worth noting that the transmission system for the RZ suffers from a sharp decline in channel capacity in comparison with the utilised DOV.

For this reason, the system is converted to the alternative solution (DOV), where the decrease in channel capacities is stopped despite continuous increases in the α factor where $\alpha <= 1$. Thus, it is necessary to replace the applied system (NOV) with the DOV to switch the context of transmission from the red to the yellow zone.

In this experiment, it is worth pointing out that using the NOV, the maximum channel capacity is achieved with the optimum value of the α ($\alpha = 0.1$). In addition, depending on the best channel capacity, the ratios of other calculated channel capacities for both the NOV and DOV are rated accordingly.

For example, in the GZ, the highest proportions of channel capacity, 85.7%, 85.7%, and 71.4%, are obtained for the α values of 0, 0.2 and 0.3 respectively. In addition, the middle ratio of channel capacity at the YZ is dropped to 57.1%

fridaya@uef.fi

for a α value equivalent to 0.4. Moreover, the lowest ratio of channel capacity (14.4%) is registered for the α values of the RZ between 0.5 and 1.

This, however, can be improved by utilising the DOV which is able to stop the bad effect of increased α for values larger than 0.4. Hence, the average ratio of channel capacity is settled at, 57.1% for all α values between 0 and 1.

Nevertheless, the gained ratio of channel capacity for the α values between 0 and 0.3 is less than the corresponding proportions of the NOV (85.7%, 71.4%). Thus, the offered DOV is more suitable for the RZ than the GZ, since the ratio of obtained channel capacity is raised, in the RZ, from 14.1% to 57.1%, which is equivalent to 4 times channel capacity of the NOV.

To mitigate such an issue, a hybrid solution is recommended utilising the advantages of both the NOV and DOV in one robust transmission system. In a such hybrid system, the NOV is applied for the α values between 0 and 0.4 while the values of α between 0.5 and 1 are allocated to the DOV. The hybrid treatment can produce a more robust transmission system with a minimum ratio of channel capacity equivalents to 57.1% at the worst condition and maximum ratio reaches 100% at the optimal value of the α.

As a result, the hybrid handling can offer both the higher channel capacity and more reliability due to its strong ability in supporting the accepted levels of the α factor for values between 0 and 0.4 and removing the serious influence of declined α for values between 0.5 and 1.

As is clear from Fig. 6, the calculated BW efficiency for the NOV has been taken firstly for the optimal value of α ($\alpha = 0.1$), where maximum BW efficiencies are achieved for different points of the SNR. These various peaks of BW efficiencies with gradually reduced levels of the SNR come from increased interference among neighbouring samples at the constellation table.

On the other hand, a sharp reduction of BW efficiency is recorded for a α equal to 0.5. Hence, using the NOV, the system shows a severe fluctuation in BW efficiency for different levels of the α.

This, nevertheless, can be solved by utilising the DOV which demonstrates a steady level of BW efficiency for different values of α, for instance, 0.1 and 0.5. In addition, the BW efficiencies with different stages of the SNR are varied for the same induced reason of the NOV.

Despite the halved spaces of the available BW due to the dual oversampling, the rates of BW efficiency for different α values (best and worst) are reduced by about 40% instead of 50% due to avoiding interference between adjacent filters for similar limits of the SNR. Thus, since the obtained BW efficiency is decided by the two important factors (BW and SNR), around 10% extra improvement is achieved by improved SNR, increasing the gained BW efficiency from 50% to about 60%.

In contrast, the NOV rates of the BW efficiencies are decreased by about 85% at worst cases of the α parameter. The main reason beyond this is the difficulty in dealing with the exaggerated interference between adjacent filters under the bad conditions ($\alpha >= 0.5$).

fridaya@uef.fi

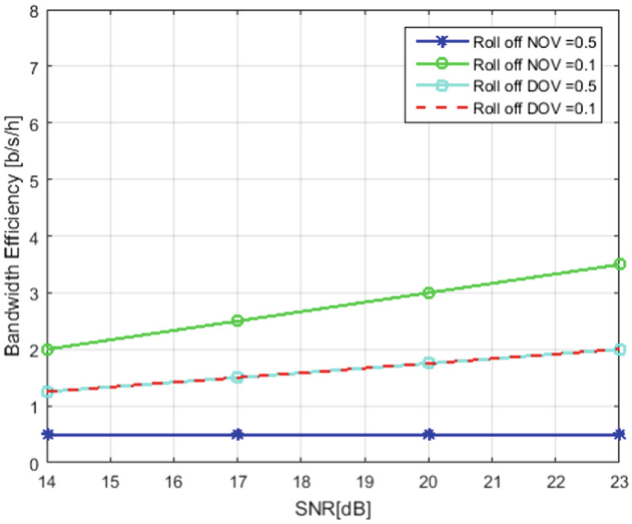


Fig. 6. BW efficiency of the OGFDM with NOV/DOV vs the SNR

Table 1. System parameters for the OGFDM

Parameter	Value
No. of $f_c s$	2
$F_{DAC/ADC}$	4 GHz
SNR	23
α	0–1
Filter length	32
Spectral efficiency	1 Hz
Oversampling	Normal - Dual
Modulation format	BPSK - 256 QAM
OGFDM symbols	2000
Filter type	Hilbert filter

It should be noted that the decreased rates of BW efficiency with both the NOV and DOV in this transmission system have been measured in comparative with the best BW efficiency of the NOV at SNR limit equivalents to 23 and α factor to 0.1.

The main parameters of this experiment are shown in Table 1, in addition, they have been applied for a wireless electrical back-to-back system.

fridaya@uef.fi

4 Conclusion

The OGFDM waveform with ICIA approach has been proposed, investigated and evaluated, for the first time. The proposed system extra sustains a higher channel capacity, more data reliability and better shareable resources in the PHY than the conventional way. Since the interference manipulation is an important process for the digital filters of the OGFDM system, a thorough investigation has been undertaken to explore the interconnected relationship between the filter's roll-off factor and the level of interference. It is worth noting that with the conventional process of interference which essentially works in the NOV mode, the aggregated channel capacity is decreased more sharply with relatively increased α parameter ($\alpha > 0.4$). Thus, the interference between the frequency responses of the applied filters is further increased due to extremely overlapped regions between similar filter types (in-phase/quadrature-phase). On the other hand, the findings show that utilising the proposed ICIA which converts to the DOV mode, the overall channel capacity is improved, particularly, for the enlarged values of the α ($0.5 <= \alpha <= 1$). Hence, by adopting a doubled value of the oversampling, the bad reaction of increased α value is accommodated optimally. Nevertheless, in comparison with the NOV, the α values between 0 and 0.3 at the DOV mode result in reducing the capacity of the channel due to the narrowed BW of the filtered sub-carriers. As a result, neither the DOV nor the NOV alone can achieve the optimum interference handling. Thus, the hybrid solution that employs both (NOV & DOV) is always recommended to boost the capacity of transmission for a diverse range of the α factor. Hence, such interference management presents a further optimization for the filtration process in the practical domain.

Acknowledgment. This research was funded by the Ministry of Higher Education and Scientific Research, Republic of Iraq, Scholarship 2633.

References

1. Dias, J.T., De Lamare, R.C.: Unique-Word GFDM transmission systems. **6**, 746–749 (2017)
2. Yazar, A., Onat, F.A., Arslan, H.: New generation waveform approaches for 5G and beyond (2016)
3. Mattera, D., Tanda, M., Bellanger, M.: Frequency domain CFO compensation for FBMC systems. Sign. Process. **114**, 183–197 (2015)
4. Geng, S., Xiong, X., Cheng, L.: UFMC system performance analysis for discrete narrow-band private networks. In: 2015 IEEE 6th International Symposium Microwave, Antenna, Propagation, EMC Technologies, pp. 303–307 (2015). https://doi.org/10.1109/MAPE.2015.7510319
5. Abdoli, J., Jia, M., Ma, J.: Filtered OFDM: a new waveform for future wireless systems. In: 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications, pp. 66–70 (2015). https://doi.org/10.1109/SPAWC.2015.7227001

fridaya@uef.fi

6. Farhang, A., Marchetti, N., Member, S., Doyle, L.E., Member, S.: Low-complexity modem design for GFDM. IEEE Trans. Signal Process. **64**, 1507–1518 (2016)
7. Kadhum, M.R., Kanakis, T., Al-sherbaz, A., Crockett, R.: Digital chunk processing with orthogonal GFDM doubles wireless channel capacity, pp. 1–6 (2018). https://doi.org/10.1007/978-3-030-01177-2-53
8. Gaspar, I. et al.: LTE-compatible 5G PHY based on generalized frequency division multiplexing. In: 2014 11th International Symposium on Wireless Communication Systems, pp. 209–213 (2014). https://doi.org/10.1109/ISWCS.2014.6933348
9. Prabu, R.T., Benisha, M., Bai, V.T., Yokesh, V.: Millimeter wave for 5G mobile communication application. In: 2016 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics, pp. 236–240 (2016). https://doi.org/10.1109/AEEICB.2016.7538280
10. Chiumento, A., Pollin, S., Desset, C., Van Der Perre, L., Lauwereins, R.: Scalable LTE interference mitigation solution for HetNet deployment. In: 2014 Wireless Communications and Networking Conference Workshops, WCNCW 2014, pp. 46–51 (2014). https://doi.org/10.1109/WCNCW.2014.6934859
11. Kadhum, M.R.: New multi-carrier candidate waveform for the 5G physical layer of wireless mobile networks. In: Proceedings of IEEE 2019 Wireless Days, Manchester (2019)
12. Gao, W.: Energy and bandwidth-efficient wireless transmission (2017). https://doi.org/10.1007/978-3-319-44222-8
13. Prommee, P., Thongdit, P., Angkeaw, K.: Log-domain high-order low-pass and band-pass filters. AEUE - Int. J. Electron. Commun. **79**, 234–242 (2017)
14. Widmann, A., Schröger, E., Maess, B.: Digital filter design for electrophysiological data - a practical approach. J. Neurosci. Methods **250**, 34–46 (2015)
15. Tao, L., et al.: Experimental demonstration of 10 Gb/s multi-level carrier-less amplitude and phase modulation for short range optical communication systems. Opt. Express **21**, 6459 (2013)
16. Im, G.H., et al.: 51.84 Mb/s 16-CAP ATM LAN standard. IEEE J. Sel. Areas Commun. **13**, 620–632 (1995)
17. Haboobi, H., Kadhum, M.R.: Impact study and evaluation of higher modulation schemes on physical layer of upcoming wireless mobile networks. Int. J. Adv. Comput. Sci. Appl. **10** (2019). https://doi.org/10.14569/IJACSA.2019.0100583
18. Mukumoto, K., Wada, T.: Realization of root raised cosine roll-off filters using a recursive FIR filter structure. IEEE Trans. Commun. **62**, 2456–2464 (2014)
19. Wang, Z., Zhou, S., Giannakis, G.B., Berger, C.R., Huang, J.: Frequency-domain oversampling for zero-padded OFDM in underwater acoustic communications. IEEE J. Ocean. Eng. **37**, 14–24 (2012)
20. Yli-Kaakinen, J., et al.: Efficient fast-convolution-based waveform processing for 5G physical layer. IEEE J. Sel. Areas Commun. **35**, 1309–1326 (2017)
21. Kadhum, M.R., Kanakis, T., Crockett, R.: Dynamic bit loading with the OGFDM waveform maximises bit-rate of future mobile communications. In: Proceedings of Computing Conference 2019, London (2019)
22. Haboobi, H., Kadhum, M.R., Al-Sherbaz, A.: Utilize higher modulation formats with heterogeneous mobile networks increases wireless channel transmission. In: Proceedings of Computing Conference 2019, London (2019)

fridaya@uef.fi

Decomposition of a Non-stationary Process Using Time-Dependent Components

Abdullah I. Al-Shoshan(✉)

Qassim University, Qassim, Saudi Arabia
drshoshan@gmail.com

Abstract. In this paper, an algorithm for estimating the time-varying components of a non-stationary signal, using orthogonal projection, is proposed. This analytical approach can also be used for estimating the time-frequency kernel, and hence the time-frequency distribution of the non-stationary process. The time-frequency kernel is then estimated with an order based on the dimensionality of the expanding orthogonal basis. Projecting each component of the signal onto an expanding orthogonal basis does the process. Some simulations and experimental results of modeling a non-stationary signal are presented when the signal has either single or multiple time-varying components.

Keywords: Multi-component signal analysis · Non-stationary process · Linear time-varying process · Orthogonal polynomial (OP) · Time-frequency distribution (TFD) · Evolutionary spectrum (ES)

1 Introduction

Due to the existence of non-stationary signals in many applications, the non-stationary signal processing techniques have become a trend [1–4]. An important issue in the analysis of a nonstationary signal is the knowledge of the degree of the signal non-stationary [1–5]. This is reflected by the expansion order of the time-varying components (TVCs) embedded in the signal [6], and hence, depending on the expansion order, the time-frequency (TF) kernel can be estimated [7]. Therefore, a method is proposed for estimating the order of expansion that will be used to give the minimum order needed for estimating the TVC, and accordingly the optimum time-frequency kernel, and thus the time-frequency distribution (TFD) of the energy of the non-stationary signal over time. In practice, experimental data and functions are usually represented or approximated by linear combinations of some predefined basis functions, which are usually of polynomial forms, are bounded and are closed under translation and scaling [8]. For these polynomials to exactly be able represent a given data sequence, they must be able to span the whole space of the data sequence. Since for curve fitting [9] (e.g. interpolation and regression), the least-square approach is used and, from a computational point of view, a set of orthogonal polynomials (Ops) is better used in the fitting problem. In the expansion approach, the set of OPs such as Hahn, Legendre, Laguerre, Jacobi, or Hermite polynomials can be used, however we need to know or estimate the number of OPs, which will be investigated in this paper, used in the fitting problem [10]. In this paper, the Legendre OPs [11, 12] are used.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 294–305, 2019.
https://doi.org/10.1007/978-3-030-22868-2_22

fridaya@uef.fi

In this paper, the evolutionary periodogram [13], an estimator of the evolutionary spectrum [7], is used in estimating the time-frequency kernel of the non-stationary signal. The formal derivation of the EP is explained as follows [13]. The Wold-Cramer representation of a non-stationary signal is defined as

$$x(n) = \int_{-\pi}^{\pi} H(n,\omega) e^{j\omega n} dZ(\omega) \tag{1}$$

where $Z(\omega)$ is a process with orthogonal increments such that

$$E\{dZ^*(\omega_1) dZ(\omega_2)\} = \begin{cases} \frac{S(\omega)d\omega}{2\pi} & \omega_1 = \omega_2 = \omega \\ 0 & \omega_1 \neq \omega_2 \end{cases} \tag{2}$$

where $S(\omega)$ is the spectrum of $x(n)$ and ω is the frequency parameter. Using the definition of the generalized transfer function of an LTV system, we have

$$H(n,\omega) = \sum_{m=-\infty}^{n} h(n,m) e^{-j\omega m} \tag{3}$$

So, the Wold-Cramer evolutionary spectrum of $x(n)$ is defined as [13]

$$S_x(n,\omega) = |\, H_x(n,\omega)\,|^2 \tag{4}$$

or simply

$$S(n,\omega) = |\, H(n,\omega)\,|^2 \tag{5}$$

Define

$$H(n,\omega) = A(n,\omega) dZ(\omega) \tag{6}$$

and consider that $x(n)$ can be modeled at a frequency of interest ω_0 as

$$\begin{aligned} x(n) &= x_0(n) + y_{\omega_0}(n) \\ &= A(n,\omega_0) e^{j\omega_0 n} dZ(\omega_0) + y_{\omega_0}(n) \\ &= H(n,\omega_0) e^{j\omega_0 n} + y_{\omega_0}(n) \end{aligned} \tag{7}$$

where $x_0(n)$ is the signal component at the desired frequency ω_0 and $y_{\omega_0}(n)$ is a zero-mean modeling error including other components of $x(n)$ at the frequency of interest. The EP is then obtained for finite length of signals, $\{x(n) \mid 0 \leq n \leq N\}$, by assuming

fridaya@uef.fi

that $H(n, \omega_0)$ can be expressed in time as expansion of OP functions, $\{\beta_m(n) \mid 1 \leq m \leq M, 0 \leq n \leq N-1\}$ so that

$$H(n, \omega_0) = \sum_{m=1}^{M} a_m \beta_m(n) = b^t(n) a \tag{8}$$

where

$$a = [a_1, a_2, \ldots, a_M]^t \tag{9}$$

is a vector of random expansion coefficients and

$$b(n) = [\beta_1(n), \beta_2(n), \ldots, \beta_M(n)]^t \tag{10}$$

is a vector of the OP functions at time n, where t stands for matrix transpose. After manipulating the above equations, we get the EP estimator of the ES expressed as [11]:

$$\hat{S}_{EP}(n, \omega) = \frac{N}{M} \mid \hat{H}(n, \omega) \mid^2, \; 0 \leq n \leq N-1 \tag{11}$$

where $\hat{H}(n, \omega)$ is an estimate of the TF kernel and can be estimated as follows:

$$\hat{H}(n, \omega) = \sum_{k=0}^{N-1} w_n(k) x(k) e^{-j\omega k} \tag{12}$$

which uses a window that varies with time and constructed from a combination of OPs such that

$$w_n(k) = \sum_{i=1}^{M} \beta_i^*(n) \beta_i(k) \tag{13}$$

The variation of $w_n(k)$ depends on the value of M, which is the order of the OP set used. Although the parameter M plays an essential role in estimating the EP, no way so far for has been proposed finding the most suitable value of M. Actually, M is a signal-dependent parameter which is connected directly to the degree of non-stationarity of the signal; therefore, this problem will be investigated throughout this paper. When $x(n)$ is stationary, then the most suitable value for M is one. However, when $x(n)$ deviates from stationarity, M should be increased accordingly, depending on how far $x(n)$ is from stationarity. Some experimental results of modeling a non-stationary signal are presented when the signal has either single or multiple time-varying components.

fridaya@uef.fi

2 Analysis of a Single Component Case

Consider a linear-phase process $x(n)$ with only one TVC $A(n)$ at $\omega = \omega_0$ such that

$$x(n) = A(n)e^{j\omega_0 n} \tag{14}$$

where $A(n)$ can be expanded using M OPs set $\{\beta_i(n)\}$ such that

$$A(n) = \sum_{i=1}^{M_0} a_i \beta_i(n) \tag{15}$$

where the set $\{a_i\}$ are the expansion coefficients of $A(n)$, and M_0 is the expansion order. Demodulating $x(n)$ by multiplying it by $e^{-j\hat{\omega}_0 n}$ will lead to

$$\hat{x}_0(n) = A(n)e^{j(\omega_0 - \hat{\omega}_0)n} \tag{16}$$

where $\hat{\omega}_0$, if unknown, is estimated from the conventional spectrum of the process and the difference between ω_0 and $\hat{\omega}_0$ is close to zero when the spectral estimator is good enough. This process is depicted in Fig. 1.

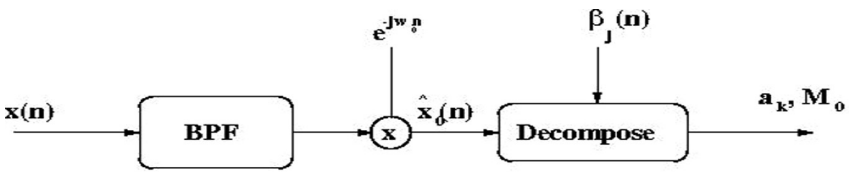


Fig. 1. A single component decomposition case.

Assuming that the set of OPs $\{\beta_i(n)\}$ are known and using their orthogonality property, then, using the inner product associated with them, we get

$$\begin{aligned} D(k) &= \sum_{n=0}^{N-1} \beta_k(n)x_0(n) = \sum_{n=0}^{N-1} \beta_k(n)A(n) \\ &= \sum_{i=1}^{M_0} a_i \sum_{n=0}^{N-1} \beta_k(n)\beta_i(n) = \begin{cases} a_k & k \le M_0 \\ 0 & k > M_0 \end{cases} \end{aligned} \tag{17}$$

where $a_k = 0$ for $k > M_0$ and we have used the property of orthogonality

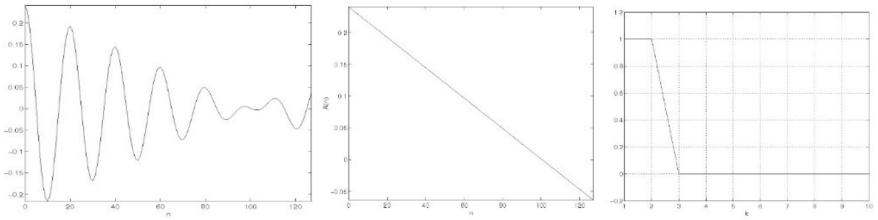
$$\sum_{n=0}^{N-1} \beta_k(n)\beta_i(n) = \delta(k-i) \tag{18}$$

fridaya@uef.fi

Therefore, if we increase k until we reach $D(k) = 0$ we will get the value of M_0 as the supremum of k. Hence if we define a set $K = \{k \mid D(k) \neq 0\}$, the minimum expansion order of the TVC $A(n)$ is defined as $\hat{M}_0 = sup_k\{K\}$. After estimating M_0, one then estimates the EP of $x(n)$ using an expansion order M equals to $\hat{M}_0$. Let us illustrate the above algorithm by means of some examples.

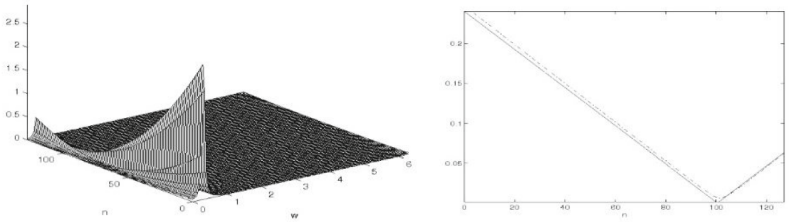
3 Experimental Results for a Single Component Case

Sequences such as the one in Eq. (14) may represent amplitude-modulated harmonics such as those arising in radar, sonar, and communication systems; an example occurring in active sonar systems is found in [14, 15]. Assuming we have the non-stationary process defined as in (14) with $\omega_0 = 0.1\pi$ radians and the TVC $A(n)$ is expanded using only two OPs $\{\beta_i(n)\}_{i=1..M_0}$ such that $A(n)$ is defined as in (5) with $\{a_i = 1\}_{i=1..M_0}$ and $M_0 = 2$. Figure 2(a) depicts the real part of $x(n)$ however Fig. 2(b) shows the time variation of $A(n)$. From Eq. (17) we get the estimated decomposition coefficients $\{a_i\}$ as depicted in Fig. 2(c). From this figure we observe that the coefficients of the decomposition go down to zero when the index k exceeds the minimum value of M, i.e. when it exceeds $k = 2$. This concludes that to track $A(n)$, we use an expansion order $M = 2$. The EP of $x(n)$ is shown in Fig. 3(a) when the expansion order $M = 2$ is used. Figure 3(b) shows the absolute value of the estimated amplitude compared with the absolute value of the actual one at the frequency of interest.



(a) (b) (c)

Fig. 2. (a) The real part of $x(n)$, (b) $A(n)$ for $M = 2$ and $\{a_i = 1\}_{i=1,2}$, (c) The estimated expansion coefficients of $A(n)$ when $M = 2$.

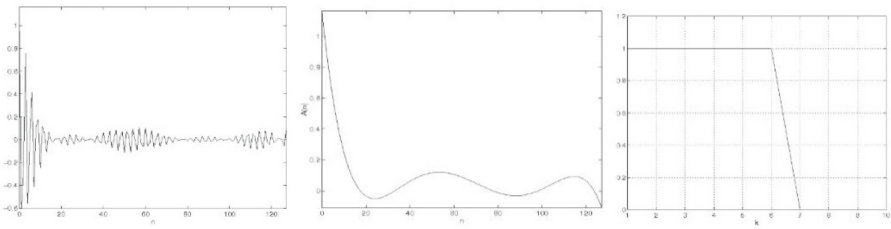


(a) (b)

Fig. 3. (a) The EP of $x(n)$ with $M = 2$, (b) $\hat{A}(n)$ (dotted-line) vs. $A(n)$ (solid-line) when $M = 2$.

fridaya@uef.fi

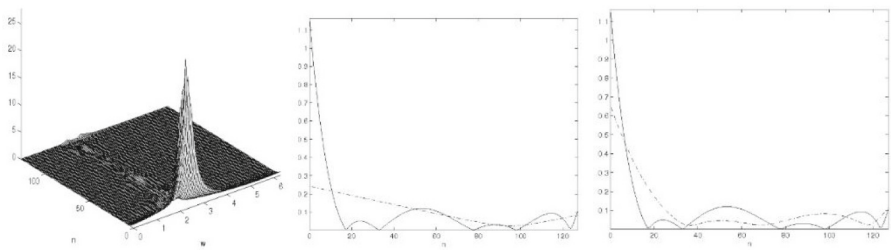
Consider now the above example but the expansion order of the TVC $A(n)$ is increased such that it can be expanded using six Ops $\{\beta_i(n)\}_{i=1\ldots6}$. In this case, assume that $\omega_0 = 0.7\pi$ and $\{a_i\}_{i=1\ldots6}$. Figure 4(a) depicts the real part of $x(n)$ however Fig. 4(b) shows the time variation of $A(n)$. Using a band-pass filter (BPF) centered at $\omega = 0.7\pi$, then after demodulating its output by $e^{-j0.7\pi n}$ we get the TVC $A(n)$ and then decompose it to get the expansion coefficients shown in Fig. 4(c). From Fig. 4(c) it is clear that the coefficients of the decomposition go down to zero when the minimum value of M is exceeded, i.e., when $k > 6$. This concludes that to be able to track the TVC $A(n)$ of the non-stationary process $x(n)$, the expansion order of the EP should be $M = 6$.



(a) (b) (c)

Fig. 4. (a) The real part of $x(n)$, (b) $A(n)$ for $M = 6$ and $\{a_i = 1\}_{i=1..6}$, (c) The estimated expansion coefficients $\{a_k\}$ for $M = 6$ and $\{a_i = 1\}_{i=1..6}$.

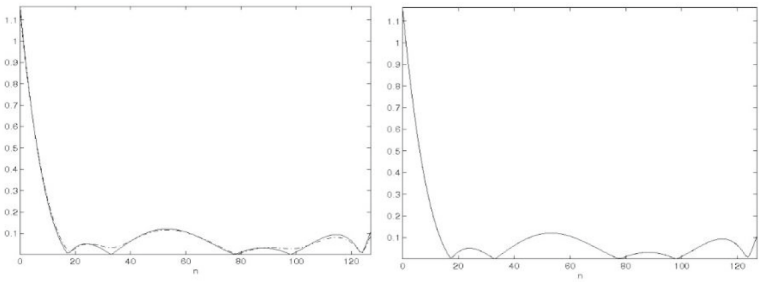
The EP of $x(n)$ is depicted in Fig. 5(a) with $M = 6$. Figures 5(b), (c) and 6(a), (b) show the absolute value of $A(n)$ compared to the absolute value of the estimated ones, with order of expansions $M = 2$, 4, 6 and 8, respectively. It is clear from the figures that when the expansion order M increases, $\hat{A}(n)$ becomes closer to $A(n)$, however, they coincide when it exceeds 6. This is obvious from Fig. 6(b) where $M = 8$.



(a) (b) (c)

Fig. 5. (a) The EP of $x(n)$ with $M = 6$, (b) $\hat{A}(n)$ (dotted-line) vs. $A(n)$ (solid-line) when $M = 2$, (c) $\hat{A}(n)$ (dotted-line) vs. $A(n)$ (solid-line) when $M = 4$.

fridaya@uef.fi



(a) (b)

Fig. 6. (a) $\hat{A}(n)$ (dotted-line) vs. $A(n)$ (solid-line) when $M = 6$, (b) $\hat{A}(n)$ (dotted-line) vs. $A(n)$ (solid-line) when $M > 6$.

4 Analysis of a Multi-component Case

Assuming that the process under consideration has more than one TVC, as depicted in Figs. 5 and 6 for two and multi-component cases, respectively.

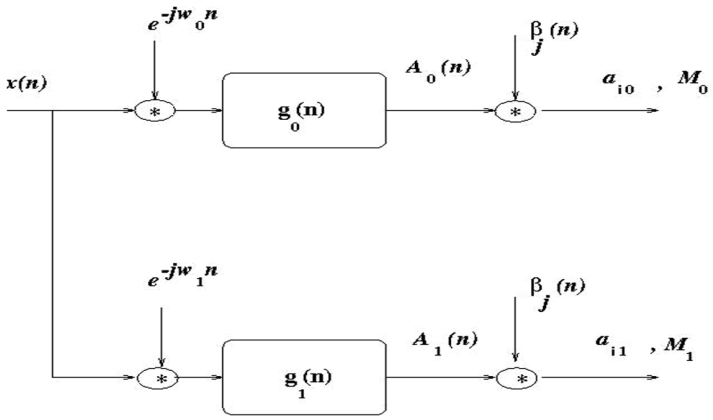


Fig. 7. An example of decomposing a signal with two components.

Using the sinusoidal representation of a TV spectra process, $x(n)$ can be represented as:

$$x(n) = \sum_{k=0}^{R-1} A_k(n) e^{j\omega_k n} \tag{19}$$

fridaya@uef.fi

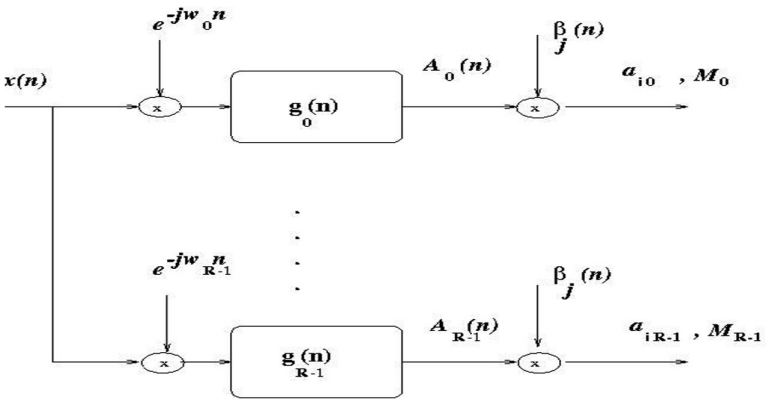


Fig. 8. A multiple component decomposition case.

where R is the number of sinusoidals in $x(n)$, and $A_k(n)$, the TVC of the k^{th} sinusoidal, is represented as

$$A_k(n) = \sum_{i=1}^{M_k} a_{ki}\beta_i(n) \quad 0 \le k \le R-1 \tag{20}$$

where $\{a_{ki}\}$ are the composition coefficients of $A_k(n)$, $\{\beta_i(n)\}$ is a set of OPs, and M_k is the number of coefficients used for expanding $A_k(n)$. To resolve these components, we demodulate $x(n)$ by $e^{-j\omega_r n}$ to get the sub-process

$$x_r(n) = \sum_{k=0}^{R-1} A_k(n)e^{j(\omega_k-\omega_r)n} = A_r(n) + \sum_{\substack{k=0 \\ k\neq r}}^{R-1} A_k(n)e^{j(\omega_k-\omega_r)n} \tag{21}$$

and then passing $x_r(n)$ through a low pass filter $\{g_r(n)\}$ whose bandwidth is larger than the bandwidth of $A_r(n)$ and assuming a non-overlapping case we get

$$\tilde{x}_r(n) = A_r(n) = \sum_{i=1}^{M_r} a_{ri}\beta_i(n) \tag{22}$$

and decomposing $\tilde{x}_r(n)$ using the OP set $\{\beta_i(n)\}$, as shown in Fig. 7, we get from the inner product

$$\begin{aligned} C_r(k) &= \sum_{n=0}^{N-1} \beta_k^*(n)\tilde{x}_r(n) = \sum_{n=0}^{N-1} \beta_k^*(n)A_r(n) \\ &= \sum_{n=0}^{N-1}\sum_{i=1}^{M_r} a_{ri}\beta_k^*(n)\beta_i(n) = a_{rk}[u(k-1) - u(k-M_r-1)] \end{aligned} \tag{23}$$

fridaya@uef.fi

where $u(k)$ is the discrete-time unit-step function. Since $a_{rk} = 0$ for $k > M_r$, $C_r(k) = 0$ for $k > M_r$ and $r > R - 1$. Therefore, the decomposition order k should be increased until $C_r(k) = C_r(k_{max}) = 0$, where $k_{max} = sup_k\{K\}$. Picking up the least upper bound of the estimated M's, i.e., $\hat{M} = max\{\hat{M}_0, \hat{M}_1, \ldots, \hat{M}_{R-1}\}$, to be the order of expansion of the EP estimator over the whole TF domain of the ES implies that $a_{rk} = 0$ for $k > M_r$. Another way of estimating the TF distribution (TFD) of $x(n)$, using the EP, is to use different values of M for different ranges of frequencies according to the different values of the estimated M's, i.e., dividing the TFD into a couple of bands each with different values of expansion order M. Two difficulties with this algorithm are: (1) the need to know the frequency of the sinusoidal components of $x(n)$; (2) since in practice we would not know the original polynomials, the value of M depends on the polynomials used to represent $A(n)$. The first problem can be remedied if the phase of each component is linear. When this is the case, the frequency of each sinusoidal can be estimated from the frequency spectrum of the process. For the second problem, one would expect different values for different sets of polynomials.

5 Experimental Results for a Multi-component Case

To illustrate the analysis approach proposed in the previous section, let us have a non-stationary process $x(n)$ such that

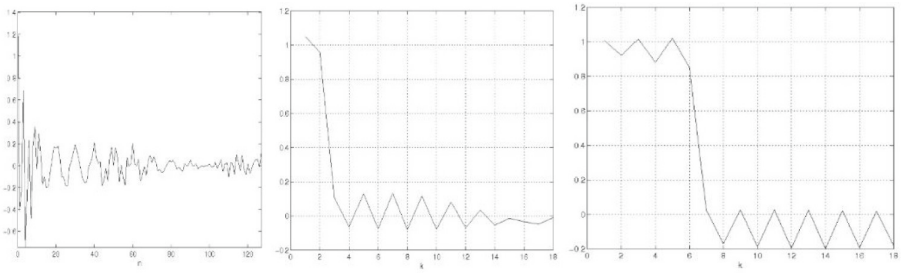
$$x(n) = \sum_{k=0}^{R-1} A_k(n) e^{j\omega_k n} \tag{24}$$

where R is the number of sinusoidals in the process and is equal to two, ω_0 is the angular frequency of the first sinusoid and is equal to 0.1π, ω_1 is the angular frequency of the second sinusoid and is equal to 0.7π, and the TVC of each sinusoid is represented as

$$A_k(n) = \sum_{i=1}^{M_k} a_{ki}\beta_i(n) \quad 0 \le k \le R-1 \tag{25}$$

with $M_0 = 2$, $M_1 = 6$ and $\{a_{ik} = 1\}_{\forall i,k}$. The real part of $x(n)$ is depicted in Fig. 9(a). Demodulating $x(n)$ by ω_0, and then by ω_1 we get $x_0(n)$ and $x_1(n)$, respectively, then, as shown in Fig. 8, passing these processes through the low pass filters $g_0(n)$ and $g_1(n)$, respectively. The filtering process is carried out under the condition that the bandwidths $BW_{A_0(n)} \le BW_{g_0(n)}$ and $BW_{A_1(n)} \le BW_{g_1(n)}$. Using an LPF with a cut-off frequency $\omega_c = 0.15\pi$, the outputs of the filters are $A_0(n)$ and $A_1(n)$, respectively. Decomposing each one of them separately, from Fig. 9(b) and (c), we get the estimated $\hat{M}_0$ and $\hat{M}_1$ to be approximately 2 and 6, respectively. Therefore, $\hat{M} = 6$ is the minimum expansion order of the EP estimator of $x(n)$. We observe from Fig. 9(b) and (c) the effects of the filtering process on the estimates of the M's as some ripples through the variation of k.

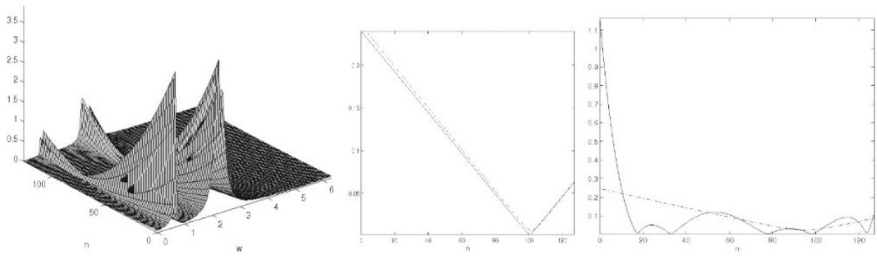
fridaya@uef.fi



(a) (b) (c)

Fig. 9. (a) The real part of $x(n)$ when $M_0 = 2$, $M_1 = 6$, and $\{a_{ik} = 1\}_{\forall i,k}$ (b) The estimated expansion coefficients, $\hat{a}_{ik}$, for $\{a_{ik} = 1\}$ when $M_0 = 2$, (c) The estimated expansion coefficients, $\hat{a}_{ik}$, for $\{a_{ik} = 1\}_{\forall i,k}$ when $M_1 = 6$.

Figure 10(a) shows the EP of $x(n)$ when the expansion order $M = 2$ is used. Figure 10(b) and (c) show the absolute values of $A_0(n)$ and $A_1(n)$ compared with $\hat{A}_0(n)$ and $\hat{A}_1(n)$, respectively, when $M = 2$ is used as the expansion order. For these two figures, we observe that $\hat{A}_0(n)$ is close to $A_0(n)$. However $\hat{A}_1(n)$ is not comparable to $A_1(n)$. This is due to the fact that the expansion order $M = 2$ is not enough for representing $A_1(n)$ since it needs an expansion order $M > 6$ to be able to track it.

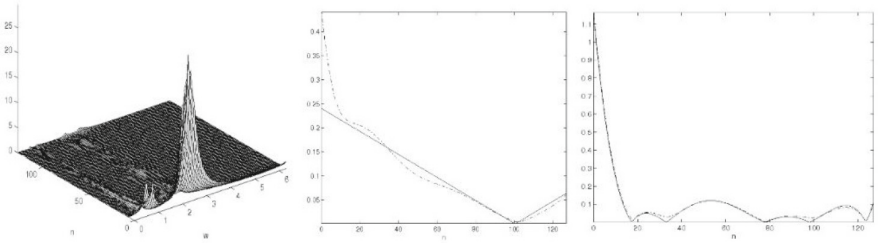


(a) (b) (c)

Fig. 10. (a) The EP of $x(n)$ using $M = 2$ as an expansion order, (b) $\hat{A}_0(n)$ (dotted-line) vs. $A_0(n)$ (solid-line) when $M = 2$, (c) $\hat{A}_1(n)$ (dotted-line) vs. $A_1(n)$ (solid-line) when $M = 2$.

Therefore, the expansion order must be increased to track the variation of $A_1(n)$. The EP of $x(n)$ is depicted in Fig. 11(a) using an expansion order $M = 6$. Figure 11(b) and (c) show the absolute values of $A_0(n)$ and $A_1(n)$ compared with $\hat{A}_0(n)$ and $\hat{A}_1(n)$, respectively, when $M = 6$ is used as the expansion order. It is clear in this case that an expansion order $M = 6$ is necessary to reasonably represent $A_1(n)$.

fridaya@uef.fi



(a) (b) (c)

Fig. 11. (a) The EP of $x(n)$ using $M = 6$ as an expansion order, (b) $\hat{A}_0(n)$ (dotted-line) vs. $A_0(n)$ (solid-line) when $M = 6$, (c) $\hat{A}_1(n)$ (dotted-line) vs. $A_1(n)$ (solid-line) when $M = 6$.

6 Conclusion

In this paper, an algorithm for estimating the time-varying components, and hence, the time-frequency kernel of a non-stationary process by using orthogonal projection theory was proposed. The algorithm was carried out by projecting the process onto an expanding orthogonal basis. The time-frequency kernel was then estimated using the EP with an order based on the dimensionality of the expanding orthogonal basis. For a noisy process, some noise removal algorithms for the non-stationary processes such as the ones in [2] and [14] may be used. For a non-stationary random process, the ensemble average of the process [8] can be used for estimating a suitable expansion order since the ensemble average sheds some light on the time-variation of the characteristics of the process. The order of the orthogonal polynomials plays an important role in estimating the time-varying parameters of the signal, which depends on the stationarity of the signal, so, a thorough investigation is suitable for future work. Simulations and experimental results for the proposed algorithms were also given.

References

1. Maheswari, R., Umamaheswari, R.: Trends in non-stationary signal processing techniques applied to vibration analysis of wind turbine drive train – a contemporary survey. Mech. Syst. Signal Process. **85**, 296–311 (2017)
2. Marple, S.L.: Digital Spectral Analysis with Applications. Prentice-Hall, Englewood Cliffs (1987)
3. Neuman, C.P., Schonbach, D.I.: Discrete (legendre) orthogonal polynomials - a survey. Int. J. Numer. Methods Eng. **8**(4), 743–770 (1974)
4. Gao, R., Yan, R.: Non-stationary signal processing for bearing health monitoring. Int. J. Manuf. Res. (IJMR) **1**(1), 18–40 (2006)
5. Deng, L., Rathinavelu, C.: A Markov model containing state-conditioned second-order non-stationarity: application to speech recognition. Comput. Speech Lang. Acad. Press Ltd. **9**, 63–86 (1995)

fridaya@uef.fi

6. Al-Shoshan, A.I., Chaparro, L.F.: Identification of non-minimum phase systems using the evolutionary spectral theory. Signal Process. J. **55**, 79–92 (1996)
7. Priestley, M.B.: Non-linear and Non-stationary Time Series Analysis. Academic Press, New York (1988)
8. El-Jaroudi, A., Redfern, M.S., Chaparro, L.F., Furman, J.M.: The application of time-frequency methods to the analysis of postural sway. Proc. IEEE **84**(9), 1312–1318 (1996)
9. Rao, S.T.: The fitting of nonstationary time series models with time-dependent parameters. J. Stat. Soc. **32**, 312–322 (1970)
10. Al-Shoshan, A.I.: Parameteric estimation of time-varying components using orthonormal bases. Telecommun. Syst. **13**(2–4), 315–330 (2000)
11. Cohen, L.: Time-frequency distributions - a review. Proc. IEEE **77**(7), 941–981 (1989)
12. Bhagavan, C.S.K.: On non-stationary time series. In: Hannan, E.J., Krishnaiah, P.R., Rao, M.M. (eds.) Handbook of Statics, vol. 5, pp. 311–320. Elsevier Science Publishers, Amsterdam (1985)
13. Kayhan, A.S., El-Jaroudi, A., Chaparro, L.F.: Evolutionary periodogram for non-stationary signals. IEEE Trans. Signal Proc. **42**, 1527–1536 (1994)
14. Boashash, B., Ristic, B.: Time-varying higher-order cumulant spectra: application to the analysis of composite FM signals in multiplicative and additive noise. In: ICASSP IV, pp. 325–328. Australia (1994)
15. Huang, N.C., Aggarwal, J.K.: On linear shift-variant digital filters. IEEE Trans. Circuits Syst. **27**(8), 672–679 (1980)

fridaya@uef.fi

Open Source Firmware Customization Problem: Experimentation and Challenges

Thomas Djotio Ndie[1(✉)] and Karl Jonas[2(✉)]

[1] National Advanced School of Engineering, CETIC, University of Yaounde 1, PoBox 8390, Yaounde, Cameroon
tdjotio@gmail.com
[2] Bonn-Rhein-Sieg University of Applied Sciences, Grantham-Allee 20, 53757 Sankt Augustin, Germany
karl.jonas@h-brs.de

Abstract. The great advantages of open source code-based projects like the case of open source firmware (OSF) are flexibility and freedom of customization. However, the difficulty inherent to their development process, which can be seen as a software composition issue, is the lack of structured approach and a teachable methodology for efficiently tackling such project. In this paper, we propose a 5-step pedagogical OSF's customization approach, coupled with an agile development process to guide the learner and ease his comprehension. We experience this approach to prototype WiAFirm, an OpenWRT-based firmware for operating IEEE 802.11x enabled WAP. Two groups of 04 students each were involved. After 2 months of experimentation, the group that applied the approach was able to integrate into the core OpenWRT a custom WifiDog captive portal feature as a built-in package; while the other group had barely understood the goal of customizing an OSF.

Keywords: WiAFirm · Open source firmware · Customization

1 Introduction

The industry of open hardware has given to scientists and researchers, the possibilities to design and develop open source firmware (OSF). It provides to end-users the freedom of customization, to engineers the freedom of improvement, and to developers the opportunity to contribute in the field by experimenting and optimizing existing (or creating new) features. An open source code (OSC) allows independent developers to create derived versions from the original one. Despite this flexibility, their development process is opaque, complex, lack structured documentation and a teachable methodology that can efficiently drive their conception. These shortcomings discourage newbies (e.g. students) to contribute in the field. We are particularly interested in the case of OSF customization issue, which can be assimilated to the software components integration problem. Which rules govern an OSC-based development project in a sustainable way? How can a learner be pedagogically introduced to gradually manage such a project?

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 306–317, 2019.
https://doi.org/10.1007/978-3-030-22868-2_23

fridaya@uef.fi

This paper aims at proposing a 5-step pedagogical approach inspired from [1], couple to an agile development process [2] to guide the learner in the customization process and to ease his comprehension. The main benefit of it is the better quality control of the OSF development process and the knowledge sharing. We applied it for prototyping WiAFirm, an OpenWRT-based firmware [3] for operating the WiABox appliance [4]. The experience consisted to integrate to an optimized OpenWRT base firmware [5] the WifiDog-based captive portal feature [6] for controlling users requesting wireless access to our campus network. Two groups of 04 students each were involved and, after 02 months, the group that applied the method was able to implement the project while the other group was almost to understand the aim of OSF customization.

The Sect. 2 of this paper presents the problem of OSF customization. The Sect. 3 describes the pedagogical approach used for prototyping WiAFirm (Sect. 4). The Sect. 5 discusses obtained results and briefly raises up some challenges.

2 The Problem of Open Source Code/Firmware Customization

If the problem of embedded system firmware development is formally addressed [2, 7, 8], it is not the case of OSC-based firmware also known as open source firmware. This issue can be formulated as a problem of software composition from third parties [9]. Composing the new firmware can then be simply like applying a puzzle's game strategy. The puzzle's pieces are assimilated to firmware packages available at their respective repositories. The new OSF results from a plan of the player, driven by the architectural concept he wants to build. The development and testing consist of arranging a game. Errors management consist of fixing or rebuilding a specific part of a worked form. The problem is that the puzzle game is intuition-oriented and has no formal rule. To the best of our knowledge, you build the figure you want given the set of pieces. If you want to build complex architectures, then you have to design new specific pieces of puzzle or remodel existing ones. It may require additional tools to glue together old, rebuilt and new parts. That is why learning from a customization experience is a complex issue and experience-oriented.

We illustrate this issue with the idea of creating an OpenWRT-based OSF [10–14]. This can result to play with more than 3000 packages from different repositories and can require more than a real strategy and experience. Some third parties open source packages can be outdated. In addition, some low-level modifications from code inspection can be required. E.g. the bridging and VLANs, intelligent and transparent layers 2 VPN configurations in the case of Freifunk [15, 16]. This implies writing some custom codes for achieving and automating desired functionalities. According to the project's objectives, new features may need to be developed and integrated in a "one-downloadable" and distributed image file. The need to have domain engineering knowledge, to know the system and development environment, the package integration, code inspection, compilation and debugging processes and, testing policies are required.

The customization development approach helps to get what a designer want, and the component composition approach helps to have it fast. This is only feasible if you

fridaya@uef.fi

are an experienced developer. The drawbacks of customization are dirty hacks, with difficulties for unexperienced developers to easily jump into the development process and hands-on the source code. There are difficulties to maintain and to add new features. Indeed, a package can be subject to bugs and security issues. If the OSF's development process was methodologically addressed, it could have reduced cost, time and, enhanced quality control. The documentation is also a great issue, even when it exists, it is unstructured, uncompleted and it is hard to see technical/scientific one.

Some activities related to the complex software customization process that we have identified are (but not limited to): (1) clone the base source code from its official repository (e.g. github); (2) use the image creation tool to select components (modules and packages) [9]; (3) create the system programming environment (sometimes multiple languages are involved [17, 18]); (4) perform raw compilation, debug, fix errors and integrate patches; develop third-party packages and glue code; (5) integrate and test; (6) decompress the source code, hack the system; (7) recover it from the break, just to name a few. For addressing an OSF project, what is the good starting point for better quality control? In the next section, we propose a pedagogical approach for designing and modeling an OSF-based project.

3 Presentation of the Pedagogical Approach

The process of forming a new OSF lacks an effective method. Most of the time, the project leaders are experts in their respective fields and almost know what they need and want and can directly start by building a prototype. We propose a pedagogical approach (illustrated in Fig. 1) in 5 phases: (1) the initialization phase (2) the conception phase, (3) the realization phase, (4) the termination phase and, (5) the maintenance phase.

Phase 1: The initialization. It comprises two sub-phases: the **system requirement** and the **domain analysis**. The objective of the **requirements analysis** is to draw the system specification and to capture the end user's needs and wants. Here is answered the question "What is the purpose of the system?" in an unambiguous and testable way [1]. The **domain analysis** helps to capture the client's knowledge framework: developing concepts, languages, terminology just to name a few. At the end of this phase, an analysis and planning document is produced.

Phase 2: The conception or design. This phase consists to determine how the OSF will provide all the needs and features. It comprises research and leads to model the system's architecture. The designer has to detect critical parts of the system e.g. those involved in real-time functionality. Here are identified eligible packages from in-depth study of the base OSF framework for customization. Their respective source code repositories are located. The need for domain engineering knowledge is necessary.

Phase 3. The realization. It consists of the effective customization of the eligible existing source codes and to write new codes, modules/packages if necessary, according to the system design requirements and specifications. Customized and new packages are glued together in the new framework.

fridaya@uef.fi

Phase 4: The termination. It consists of the deployment and the integration of the developed and compiled packages in form of binary files (e.g., *.ipk* and *.bin* in our case). Then various tests (unit and global tests), performance evaluation and validation are made. The system testing, system deployment and, the system evaluation are important steps to validate and insure the quality control of the final build.

Phase 5: The Maintenance. It consists to collect users experience and change management for future improvements and, some can take us back to reviewing design, analysis phases. The documentation is an important artefact during the project lifecycle.

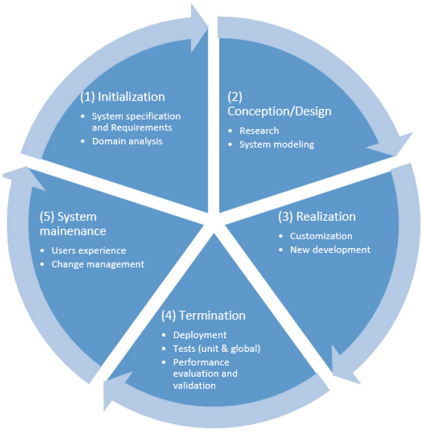


Fig. 1. Proposed model for an OSF-based project development process

In this section, we realize that it is definitely not an easy task to start the development of a new firmware project from existing OSFs daily maintained by free, enthusiastic, passionate and, motivated communities like the case of OpenWRT or Freifunk-Gluon. This is particularly a difficult problem if the project aims to have professional, technological and scientific impact like the case of WiAFirm.

4 On Modeling WiAFirm an OpenWRT-Based Firmware for WiABox Appliance Framework

WiAFirm is a modular OpenWRT-based framework mostly inspired by the Freifunk-Gluon firmware [15] and OpenWISP [18] for operating multi-band IEEE 802.11x wireless access routers. It is part of the WiABox project initiative [4].

4.1 The Initiation Phase: System Requirements and Specification of WiAFirm

WiAFirm will be the base firmware for building different wireless community networks (WCN) with varied expectations, with special focus to hard-to-access areas. It MUST

fridaya@uef.fi

[19] be easy to setup, customizable, pre-configurable and, compatible with other OpenWRT-like firmware and hardware. It MUST provide a web GUI to non-skill users for configuring and monitoring network settings. For supporting the wireless network access and control (WNAC), WiAFirm MAY implement a customized hotspot and captive portal solution for end-user's authentication through an Authentication, Authorization, Accounting (AAA) backend server. WiAFirm MAY be used to deploy private WLANs. Therefore, the use of different credential types including digital certificates, user-names and multifactor passwords will be possible. It inherits Freifunk-Gluon essential requirements: security, configuration and much more. The anonymity of a **user** MUST be ensured in public places and in-line with local regulations. It MUST guaranty an end-to-end protection of end-user data. It SHALL implement auto-configuration and self-healing features to deliver secure, scalable and reliable service to end-users. It MUST provide (1) geolocation to ease remote and graphical monitoring, (2) mesh VPN on LAN and WAN to OPTIONALY allow someone to share part of his Internet bandwidth connectivity with the community.

In this paper, we choose to implement the WNAC requirement only by building a customized captive portal feature based on WifiDog [6].

4.2 Phase 2: Conception and Design of WiAFirm

We start with an in-depth study of the base OpenWRT OSF and the structure of its package. Table 1 below gives a snapshot of the list of eligible packages needed in regard to the experimentation.

Table 1. List of features for package selection.

Features	List of packages	Source repositories	Notes
Kernel modules	OpenWRT base modules	OpenWRT	IEEE 802.11x Fundamentals
Target system or OHW architecture	Ar71xx-generic	OpenWRT	
EAP-AAA backend server	FreeRadius v3*	FreeRADIUS	802.1x EAP
WNAC: Hotspot & Captive portal	WifiDog	WifiDog	UAM
VPN Mesh protocols	VPN on LAN, on WAN (fastd, tinc, alfred)	Freifunk-Gluon	O-SP*
Enhanced L2&3 routing and mesh protocols	Bird, Batman, Batman-adv (meshclient), babel, bmx6 (meshbackbone)	Freifunk-Gluon	O-SP*
IPV4/IPV6 addressing	radvd	Freifunk-Gluon	O-SP*
Mapping, GIS monitoring services	OWGIS,	Freifunk-Gluon	

fridaya@uef.fi

In this table, the content commented "O-SP*" mentions the features that are out of the scope of this paper. As above stated, we will limit our experimentation on only the implementation of one feature: the captive portal.

The System Architecture. The design considerations and component's identification leads to the conception of the framework architecture illustrated in Fig. 2. The users of the WiAFirm framework, in addition to benefiting to WiABox wireless community services [4], will always have options to use wireless services natively provided by OpenWRT. The concept of plugins inherited from Freifunk-Gluon is used to extend the framework by customizing some existing features and developing new community-based services.

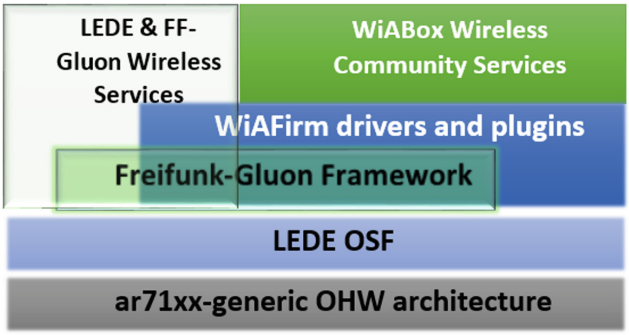


Fig. 2. The WiABox Appliance framework's architecture.

Conceptual architecture of the Captive Portal: WiAGate. Its architecture is illustrated in Fig. 3 and consists of two logical parts respectively running on a Web server and on the wireless router: The WiAGate_Server and the WiAGate_CP. Its complete presentation is out of the scope of this paper. It is the full hotspot and captive portal solution for WiABox project.

WiAGate_CP is a WiFiDog-based package to provide the captive portal service. It communicates with WiAGate server which integrates some specific system processes among which are:

- **WiAGate_Auth.** It intercepts all requests from WiAGate_CP and performs authentication check, sends the response back to it. Based on the response value, the WAP can be authorized or not to connect the end user to the network.
- **WiAGate_Acc.** It provides the data access history and accounting for each hotspot, WISP and end-user for pricing, billing and payment purposes [20].
- **WiAGate_Mon.** It gives the possibility to specify some restrictions such as duration/time limit of the connection, the volume or the maximum usable bandwidth.

fridaya@uef.fi

4.3 Phase 3: Prototyping of WiAFirm, Design and Integration of WiAGate_CP

Customizing the Candidate Package. The eligible package for customization is WifiDog V1.3.0 [6]. The resulted customized package to be integrated as an OpenWRT build-in feature is WiAGate_CP. The required development resources are: the C programming language (gcc on Linux Kernel), the source code of the WifiDog and the source code of the stock LEDE release 17.02.3.

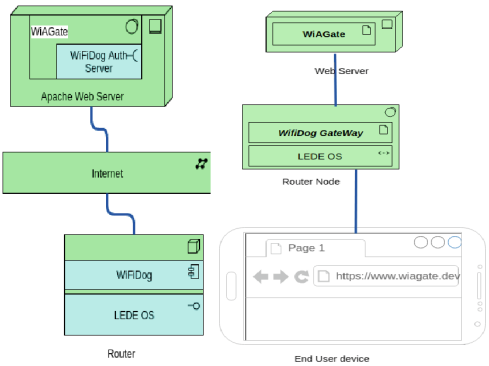


Fig. 3. Interaction between: (a) the WAP (router) and (b) End user device and the WiAGate Server through Internet

The steps to follow for creating and integrating a package as a OpenWRT build-in package are inspired from the "helloword" example tutorial described in the developer manual available here [21]. They are summarized in 12 steps below:

(1) Preparing, configuring and building the necessary tools.
(2) Adjusting the path variable for project to be created.
(3) Creating the source directory and files for hosting the base source code.
(4) Creating a package as a standalone "c" program. Here we cloned the source code of WifiDog into a local directory e.g. the 'wiagatecp' folder in this case. We inspected the source code and brought all the necessary changes and patches according to our expectations.
(5) Compiling and testing the application.
(6) Creating the .ipk package from the application.
(7) Creating the package Manifest file "Makefile" for the resulted package.
(8) Including the package into the OpenWRT system by updating the feeds.conf. default or feeds.conf file.
(9) Update and install feeds from the/home/buildbot/source directory. A web feed is a query-based system designed to bring the user relevant, immediate, useful and regularly updated stream of data or information. It aims to make a collection of web resources accessible in one spot.
(10) Testing WiAGate_CP as a standalone application.

fridaya@uef.fi

(11) Publishing the WiAGate_CP in a new repository.
(12) Customizing the OpenWRT stock firmware with the generated WiAGate_CP package. We used an extended and optimized stock base LEDE build for TP-Link Archer C7V2 [22]. It includes some preconfigured packages and kernel modules that feature some of our performance needs and wants.

Compiling the Stock LEDE with WiAGate_CP as a Build-in Package. All the 7 steps described hereafter are done on Linux Ubuntu 16.04 LTS. We followed the steps advised on the OpenWRT's official website [13, 21] to process this task. It comprises:

Step 1: Installing the dependencies.
Step 2: Updating the sources and getting the stock LEDE source code with Git.
Step 3: Update and install the source from feeds.
Step 4: Getting the WiAGate_CP package from our repository and unzipped it into a local folder.
Step 5: Configuring the build image by using the builtroot tool build [21].
Step 6: Compiling the build (building WiAFirm image). This can take time, depending to the performance of the host computer. It can lead to fixing errors if any until having an error-free process.
Step 7: Resetting the device with WiAFirm, the resulted customized build. The target wireless access router is TP-Link Archer C7V2. After a successful reset process, the WiAGate_CP package can be verified in the running build image through the CLI or luci interface. WiAFirm is ready for deployment and testing.

4.4 Termination Phase: Prototype Deployment and Testing

The test environment consisted of a 64 bits Linux OS Ubuntu 16.04 LTS. We used three (03) TP-Link AC1750 C7V2 wireless dual band Gigabit router. The additional tools are programing languages and frameworks: PhP, NodeJs, VueJs and Laravel, FreeRADIUS v3, and some Debian software systems utilities. The WiAGate_Server is up and running on the Linux host. It is the backend server also performing AAA capabilities. Its presentation is out of scoop of this paper.

The test is possible from any browser on any Wi-Fi enabled end-device. Any user will be warmed that an authentication is needed prior getting access to network. A login or a registration is required.

4.5 Maintenance Phase

The system and package testing, system deployment, system maintenance and system evaluation are important steps to assure the quality of the final build. Here we collect users' experience, results and measurements, analyze them and review all the process if necessary. This can take us back to review the complete customization process. The documentation is an important artefact during the project lifecycle, e.g. the design document of the project.

fridaya@uef.fi

5 Related Works and Business Idea

Almost all the related OSFs OpenWRT, Freifunk-Gluon [15, 16], OpenWISP [17, 18, 23], Chillifire [24], DDWRT [25] just to name a few, lack structured and scientific documentation and consequently raise shortcomings of learning from their experiences. Some are partially closed source like the case of Chillifire and DDWRT. The above-related projects are almost widely in use in their local countries and they each have active developer communities, so it will be a bad idea to think starting from scratch. OpenWRT/LEDE currently manages more than 3000 software packages. When considering each fork, more additional specific packages need to be taken into account. Exploiting such huge list of components is a great challenge. By applying the "pick-on-desire" policy, one just has to select desired packages, configure them and in doing so, build his own architecture. When missing, he can write his own package and integrate it based on specific rule. This approach suits to general-purpose projects with a low degree of customization on low-end devices for building each own solution by means of few configurations. Nevertheless, with no experience, no methodology and required skills, this really leads to great confusion.

The business idea is to build free Wi-Fi and share it with family and friends, in café, in the community and expand it town-wide, then rural-wide then, collect users' experience, surveying and provide valued-added services on top of the built WCNs. Each WCN is autonomous in terms of structuration and organize trainings and professional events for promoting the community-oriented services [26]. End users are (but not limited to), SMBs, researchers, academics, network administrators and developers, commercial, education and health centers, city councils, remote communities living in hard-to-access areas, hotels, camping, cafés, bars, residences, remote schools and colleges. People living in the neighborhood of the above-mentioned places can automatically benefit from the hotspot services providing contributions from other users with their internet connection, bandwidth sharing and their packet forwarding ability. Competitors are private local WISP and cybercafés able to provide Wi-Fi community-wide.

6 Discussion and Challenges

It may happen that a package has several feeds. It is recommended to focus on the most active one. The criteria used to select a package are stability, functionality, security, strong community, documentation even if incomplete, development language, openness, maintenance and communication. The anatomy of the base OSF code shows that some source packages are natively included in its official repository. Example is the case of OpenWRT and its forks OpenWISP [17, 18, 23] and Freifunk-Gluon [15, 16]. It is not recommended to select them as such to expect good results because they have a separate trunk and are not updated with the same frequency. That is why it is recommended to directly clone the official repository of the candidate source packages, customize them according to our needs and wants, create bin or ipk files and piece them all together with the OSF base kernel. For new developed packages emerged from the

fridaya@uef.fi

system analysis, we exactly do the same after successful compilation and tests. For integrating all the identified packages, the process is summarized as follows:

```
For each package do:
  search for the right version,
  update the /etc/apt/sources.list file,
  import right keys for validating the downloaded sources,
  update and upgrade the system (in Linux Ubuntu apt-get update
& apt-get upgrade),
  install the dependencies,
  clone the source code,
  configure it in regards of expected results,
  create the build accordingly with the required command (make,
make menuconfig, make install, make image),
  do post-configuration actions: fix errors if any
until process is error-free.
```

The aim of customization is to create working code as quickly and as accurately as possible. However, this process involves code inspection, testing and debugging which are activities most of the time reserved to certain elite, that is predominantly professional.

Most of the features to be implemented in WiAFirm are inspired from Freifunk-Gluon, OpenWISP and Chillifire. These solutions are optimized for urban areas and for developed countries. Some are partially closed source and efforts to use them as such is difficult. They were designed with the assumption that the Internet connectivity will be always available, but in our context it is not the case and, building a business model with low-income people is not an easy task.

As an important element in the definition and design of WCNs, an OSF in its current architecture, despite its flexibility of multi-context adaptation becomes inappropriate in an ubiquitous service environment [5, 27]. This is why we are considering using the SOA paradigm to propose a complete service-oriented architecture as a response to this problem. A package would therefore be seen as a service (**PckaaS**) in the SOA perspective and would instead offer business functionality to other applications or packages hosted in a service repository. Considering WiAFirm as a FaaS (firmware as a service) will allow, instead of integrating the functionality of the application itself at the WAP level, the possibility to implement instead a communication interface with the packet registry in order to use the package it needs on demand.

7 Conclusion and Future Works

In this paper, we have addressed the issue of modeling a new firmware from existing, proven most used and, tuned OSFs like OpenWRT/LEDE and Freifunk-Gluon. Because it lacks a formal methodology to tackle such project, we proposed a 5-phase pedagogical approach as a pragmatic development process model that can be applicable to any OSC-

fridaya@uef.fi

based project. The approach consists of the: (1) initialization, (2) conception, (3) realization, (4) termination and, (5) maintenance phases. We illustrated it with in-depth steps on how to customize a WifiDog-based captive portal (WiAGate_CP) and, integrate it into an optimized stock LEDE, as a build-in feature for controlling access to end-users in the WCN. The resulted build is the prototype of WiAFirm, inspired from Freifunk-Gluon for operating WCNs in rural DCs.

The main benefit of the proposed approach is the better quality control of the OSF development process and the knowledge sharing. The learner is methodologically introduced and guided step-by-step in the OSF engineering paradigm.

Regarding the difficulties encountered and to benefit from the great potential of WCNs in a ubiquitous service environment, we consider in the future using the SOA paradigm to model WiAFirm as FaaS (firmware as a service). Thus, instead of integrating the application functionality itself at the WAP level, it will instead implement a communication interface with the package registry to use the package it needs on demand.

Acknowledgments. Special gratitude respectively goes to Alexander von Humboldt Foundation; GIZ-CIM Organisation; the University of Applied Sciences Bonn-Rhein-Sieg and the CETIC Project of the University of Yaounde 1 for their various supports.

References

1. Barr, M.: Firmware architecture in five steps (2009). http://www.embedded.com/design/prototyping-and-development/4008800/1/Firmware-architecture-in-five-easy-steps. Accessed Nov 2019
2. Greene, B.: Agile methods applied to embedded firmware development. In: AGILE Development Conference 2004, pp. 71–77. IEEE, Salt Lake City, Utah, USA (2004)
3. LEDE homepage. https://lede-project.org/. Accessed June 2017
4. Djotio, N.T., Jonas, K.: WiABox 2507: Project Initiative for an Open Wireless Access Technology alternative for Broadband Internet access in rural areas. In: WACREN Workshop, TANDEM Project Dakar Senegal (2016)
5. Brown, E.: OpenWRT adds IPv6, preps for IoT future (2014). http://linuxgizmos.com/openwrt-adds-ipv6-and-prepares-for-iot-future/. Accessed June 2018
6. WifiDog Captive Portal Suite project homepage. http://dev.wifidog.org/wiki/doc. Accessed may 2018
7. IPA/SEC.: Embedded System Development Process Reference guide. Version.2. Software Engineering Center, Technology Headquarters, Information-Technology Promotion Agency, Japan (2012)
8. Geiger, L., Siedhof, J., Z¨undorf, A.: µFUP: a software development process for embedded systems. In: Proceedings Dagstuhl-Workshop MBEES: Modellbasierte Entwicklung eingebetteter Systeme I, Schloss Dagstuhl (2005)
9. Capretz, L.F.: Y: a new component-based software life cycle model. J. Comput. Sci. **1**, 76–82 (2005)
10. NAGY, P.: Configuration of OpenWRT System Using NETCONF Protocol. Bachelor's thesis. Brno University of Technology, Faculty of Information Technology (2016)
11. OpenWrt history. https://wiki.openwrt.org/about/history. Accessed Aug 2018

fridaya@uef.fi

12. Fainelli, F.: OpenWrt/LEDE: when two become one. LEDE (2016). https://events.linuxfoundation.org/sites/events/files/slides/ELC_OpenWrt_LEDE.pdf. Accessed Oct 2017
13. OpenWRT homepage. https://www.openwrt.org/. Accessed Jan 2019
14. Capoano, F.: Do you really need to fork OpenWRT? Sharing our experience with OpenWISP Firmware (2015). https://prplworks.files.wordpress.com/2015/10/do-you-really-need-to-fork-openwrt.pdf. Accessed Oct 2017
15. Freifunk Gluon Documentation. https://media.readthedocs.org/pdf/gluon/latest/gluon.pdf. Accessed Aug 2018
16. Freifunk gateway. https://wiki.freifunk-franken.de/w/Freifunk-Gateway_aufsetzen. Accessed June 2018
17. Ferraresi, A., Goretti, M., Guerri, D., Latini, M.: CASPUR Wi-Fi Open Source. In: GARR Conference 2011 (2011)
18. Openwisp home page. http://openwisp.org/. Accessed Aug 2018
19. RFC 2119: Key words for use in RFCs to Indicate Requirement Levels (1997). https://tools.ietf.org/pdf/rfc2119.pdf
20. Djotio, N.T.: An entity-based black-box specification approach for modeling wireless community network services. In: Proceedings of SAI Computing Conference 2019 (2019)
21. OpenWRT Developer guide. https://openwrt.org/docs/guide-developer/start. Accessed Dec 2018
22. Optimized LEDE. https://forum.lede-project.org/t/gcc-6-3-build-optimized-tp-link-archer-c7-v2-ac1750-lede-firmware/1382. Accessed Dec 2018
23. Gijeong, K., Sungwon L.: Openwincon: open source wireless-wired network controller software defined infrastructure (SDI) approach for fixed-mobile-converged enterprise networks. In: ICN 2016: The Fifteenth International Conference on Networks (includes SOFTNETWORKING 2016), pp. 58–59 (2016)
24. Support.chillifire.net: Chillifire Hotspot Router Installation Guides. http://www.chillifire.net/. Accessed July 2018
25. DDWRT homepage. https://dd-wrt.com/. Accessed Jan 2019
26. Freifunk homepage. https://freifunk.net/. Accessed Jan 2019
27. Open vSwitch. http://openvswitch.org/. Accessed June 2018

fridaya@uef.fi

An Entity-Based Black-Box Specification Approach for Modeling Wireless Community Network Services

Thomas Djotio Ndié(✉)

National Advanced School of Engineering, CETIC,
University of Yaounde 1, PoBox 8390, Yaounde, Cameroon
tdjotio@cetic.cm

Abstract. Wireless community networks are emerging as a better alternative to bridge the digital divide in underserved areas. As such, they can stimulate a proximity economy and allow the emergence of local digital service operators like wireless service providers (WISPs). To achieve this goal, we propose in this paper, an entity-based black-box specification approach to formally describe: (1) the concept of wireless community network service (WCNS), (2) the mechanisms of interaction with the latter through the authentication-authorization-accounting (AAA) model, and (3) the profitability mechanisms that complement the AAA model which are: pricing, billing and payment (PBP). Our case study is a campus-type WCN. We have defined an abstract representation of the network service concept through characteristics common to all services. The access control and profitability functions are described through well-defined and justified formalisms. These functions are also well illustrated by practical cases.

Keywords: Network service · Service access · AAA · PBP · WCN

1 Introduction

The goal of building wireless community networks (WCNs) [1, 2] is usually for social and not-for-profit purpose. Built based on grassroots approach [3] and depending on the project's initiators, they offer free and unrestricted access to network services to users in the need. Economic profitability is therefore not an initial concern for the designers of such projects [4]. Yet, they can be an excellent alternative for the development of the digital proximity economy in developing countries (DC). For this to be possible and especially in a context of ubiquitous network service, it is important to challenge some well-established models such as access control models for interaction with network services, in order to effectively integrate the profitability dimension into community wireless projects adapted to underserved areas.

Many works present the AAA model (authentication, authorization and accounting) [5] as a solution to this concern, except that it is more suitable for business networks that require more security through well-known standard protocols like RADIUS [6] or Diameter [7]. The AAA model was designed without taking into account financial

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 318–335, 2019.
https://doi.org/10.1007/978-3-030-22868-2_24

fridaya@uef.fi

appreciation functions, functions that could be very useful in a service-oriented WCN context. To this shortcoming we propose the PBP model for pricing, billing and payment. We obtained a more complete model called 3A-PBP, composed of 06 sub-models or components, each of which fully describes input and output spaces for the full model.

Our problem is thus reduced to that of the description of the input and output spaces of a system, as well as the functions linking the outputs to inputs. The realization of the functions allows to operate the system. Our solution is based on the entity-based black box specification method, which is more appropriate, for intuitively modeling WCN's services, by using mathematical formalisms to describe the properties and the behavior of our system. This would solve the problem of the variety of interpretations and offer flexibility and neutrality in terms of the implementation environment of the 3A-PBP model. Entities are the elements of the A-A-A-P-B-P protocol; the outputs produced by a component are the inputs for the next component. This method requires 04 steps: (1) The definition of the entry and the output spaces; (2) The definition of the system entities; (3) The definition of constraints and; (4) The definition of the relationship between inputs and outputs. Figure 1 illustrates the architecture of the whole system and highlights the actors and its different components.

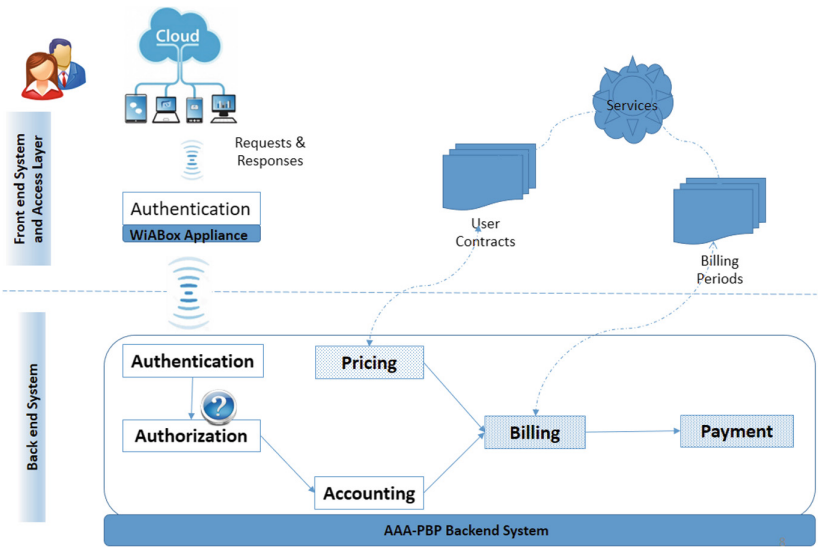


Fig. 1. The 3A-PBP architecture model.

The Sect. 2 defines and characterizes the wireless community network service concept. The Sect. 3 describes WiABox.Net, our virtual WISP for the simulated practical cases and set the assumptions as well. The Sect. 4 models the 3A-PBP ecosystem through its 06 entities using well-defined functions and properties. Before concluding, the Sect. 5 raises up some discussions and challenges.

fridaya@uef.fi

2 The Concept of Service: Definition and Characterization

2.1 Definition of the Service Concept

Just like a product, a service is something that can be sold, but is intangible. Talking about sales involves 03 actors: a producer, a supplier and a consumer. In the context of computer networks, a producer creates and administers digital services through an administration system. The provider subscribes to the services and makes them available to consumers through a platform. The consumer requests the services provided by the platform according to certain predefined constraints. In the concern of this paper, we will consider only two actors: the supplier and the consumer (also called user or customer).

2.2 Dimension of a Service

The concept of service has 04 intrinsic characteristics that distinguish it from the notion of product: intangibility, inseparability, perishability and heterogeneity.

Intangibility means that the service cannot be materialized, e.g., in the case of a digital library service, this is illustrated by the fact that access to the library is not materializable. Inseparability characterizes the fact that the production of the service and its consumption are simultaneous. For example, the bandwidth provided by a Wireless Internet Service Provider (WISP) is instantaneously consumed (totally or partially depending on the network throughput). Perishability means the service cannot be returned or stored. The example of the Internet bandwidth illustrates this property as well. Heterogeneity denotes the fact that the quality is not constant from one service to another; and the performance of a service, even with the same provider, varies with the user. For example, the actual Internet throughput will not be the same for 02 users who have subscribed to the same package.

2.3 Actors: Service Provider (SP) vs Service Consumer (SC)

We only consider two main actors: the SP and the SC, who both interact and trade at the business level. The SP indefinitely produces services and earns money. The SC indefinitely consumes services and spends money. In the context of a WCN for example, the profit made by a SP can be financially nil, but socially important. Several factors come into play. We are interested in evaluable mechanisms or methods for controlling and accessing provided network services. The SPs must provide authenticatable, traceable, accountable, priceable, billable and payable services. The system must enable relevant analysis, must be error-free so that SCs can usefully spend their money for the requested services. The access to services is constrained by the availability of resources whose allocation must comply with certain formal rules. SPs must allocate resources efficiently and to do so, they will need to study and analyze the trend in the use of available resources, to know for example, how many services they have or can be instantly provided.

The SC is the beneficiary of the service. He requests access to the service, this access is conditioned by the positive verification of a certain number of information, in

fridaya@uef.fi

particular (1) the capacity of the latter to compensate the service, (2) the capacity of the system to provide the requested service and, (3) the availability of the service itself.

Three service payment alternatives are possible: (a) prepaid, (b) postpaid and, (c) on demand (on the go). SCs are constrained by the throughput or the "quantity" of the provided and available service. In order to be able to equate the service used with the cost to be paid by the SC, a SP must be able to quantify the service used (e.g. by the time of use) in order to charge the SC. Services are served to SCs in general according to several packages or flat rate formats. A package is simply a parameterized service. E.g. for the Internet access, these parameters can be the connection throughput and the exchangeable data volume. Therefore, we can have settings like package1 (2Mbps; 1Go), package2 (1.5Mbps; 2Go as example of packages. This not only allows the SC to have several choices to satisfy his needs, but also for the SPs to offer several pricing options.

2.4 The Concept of Time

The time is the element that helps to differentiate moments or instants [8]. In the collection of information, it is necessary. It allows the study of the trend of the use of the service. The time is much more a statistical tool; it offers the possibility to study the population of SCs and SPs and allows the counting of the "quantity" of the service used.

3 Base Assumptions and Presentation of WiABox.Net, Our Virtual WISP

3.1 Network Service Example and Presentation of WiABox.Net

The network service example that will support our practical illustrations is *Internet access*. This service will be provided according to well-defined constraints by WiABox.Net, our virtual WISP for simulation purpose. To simplify the model, we will assume that WiABox.Net does not operate in a competitive environment. All users of the Internet service (SP) will have the same settings: ID; Password; Name; FirstName; Email; Package; RegistrationDate and, Address.

3.2 Package Concept

WiABox.Net supports SCs based on multiple packages. A package is simply a combination of the bandwidth and the amount of data that are allocated to a user. WiABox.Net parameterizes 05 packages named: packageA, packageB, packageC, packageD and packageE. We will assume to simplify the model that WiABox.Net has only 01 long range wireless access point (WAP), and it can only grant access to up to 10 users simultaneously. The users are all distributed (a priori not uniformly) in the perimeter of the WAP.

3.3 Concept of Physical User and Logical User

A physical user refers to a SC or client. He can be identified by his email address and password. In contrast, a logical user is a user account registered in the database. It can

fridaya@uef.fi

be identified by the combination of his ID, email address, password and/or package. A SC can have one or more user accounts. A package plan is associated with an account that uniquely identifies it. A SC can then subscribe to one or more different packages.

4 On Modeling the 3A-PBP Ecosystem

In this section we model the ecosystem of the 3A-PBP (heard triple A-PBP) through its six principle components (or entities) using well-defined functions and properties.

4.1 Authentication

The authentication is the mapping and comparison of a received identity to that of a stored one (equipment or user) in a system. We apply the OAD process, which consists first to *obtain* the credentials from a host, and then *analyze* them, and finally *determine* if they are associated with the requesting entity. Authenticating a user would then consist to compare and match his identity (e.g. his email address and password) to all the stored information that refers to him (his name, first name, email address, phone number, password, etc.). To authenticate a user of a service, the system therefore needs information that uniquely identifies him and a test function.

The set I of the inputs is the set of information of all [possible] types that can be persistently inserted into the system. A fortiori it contains, e.g. all the strings of characters that can be used to identify a user [9]. If the system needs an email address and a password as input data, then the system requires two strings. Hence, I can be simply seen as the set of strings and the set of input information is then an element of I^2.

Authentication Function. The *Auth* authentication function is the function that tests whether the information entered by a user and processed by a host, correspond to a user included in a certain group for identification purposes. We define the *Auth* function as following:

$$\begin{gathered} Auth : I^n \times \mathcal{P}(U) \rightarrow \{0,1\} \\ (i,E) \mapsto Auth(i,E) \end{gathered}$$

- i is an element of I^n, where n is chosen by the SP, it is the set of identifiers;
- E is a part of U, the set of all users of the system;
- When this search is successful, the function returns 1, or 0 otherwise.

The authentication function in most of the cases will be the search and comparison for $i \in I^n$ in a projection, according a subset of the set of users includes in the set E.

Practical Case. Considering WiABox.Net our WISP defined above, only the ID that is automatically generated by the system or the couple (email_address, password) will allow to uniquely identifying a SC. We will then use it for these purposes.

fridaya@uef.fi

What happens if a SC has multiple packages and uses the same email address and password for all his user accounts?

As stated in the specifications of WiABox.Net, a user is fully defined by his email address, password and package. Thus, a SC with three different packages: f1, f2 and f3, has three user accounts. However, since the system just asks for the email address and password for identification process, this would mean that if a SC has more than one user account, the system will have to make a choice among its accounts. The system will choose his oldest user account.

Since these are strings that are processed, then I is the set of all possible strings. So the global information provided by the client is an element of I^2. In addition, only its users U of the service are concerned. As a WiABox.Net user is an 8-tuple of the type (ID, password, last_name, first_name, email_address, package, registration_date, address), then $U \subset I^8$. Finally, to authenticate a SC who sends i = (email_address, password) as a set of identifiers, the *Auth* function is explicitly defined by:

$$Auth(i, E) = 1_{E_{|val_5, val_2}}(i)$$

- $E_{|val_5, val_2}$ is the projection of E according to the items email_address and password.
- Remember that for a given set E, 1_E denotes the indicator function of E.

4.2 Authorization

In the AAA model, the principle of authorization consists to decide which resources will be allocated to an authenticated user on the basis of verifiable criteria. The answer to the request for using a resource by a user x is positive, if and only if the proposition '$p(x)$*: x is eligible for using the service*' is true.

Now Let us assume that WiABox.Net has only 05 available hosts to offer Internet access and that 07 of its subscribers wish at the same time to consume the service according to their packages. For each customer x, the proposition $p(x)$ is true, but the constraint of available resources dictates that only 05 of them will be eligible. How to objectively choose these 05 customers among the 07, knowing that each deserves as much the service as the others?

We realize that the Boolean logic is limited and inappropriate to provide an effective and relevant solution to this problem. Indeed, the proposition $p(x)$ can only have 02 possible values of truth: True (V) or False (F), which is insufficient to efficiently classify or order SCs while limiting frustrations. We must find other additional criteria such as seniority or type of package, to be taken into account in the decision-making. This leads us to introduce an additional merit variable as the degree of verification of the proposition $p(x)$. The appropriate mathematical tool in this case is the Fuzzy logic [10].

The Fuzzy logic is an extension of the Boolean logic based on the mathematical theory of fuzzy sets [11, 12]. It introduces the notion of the degree of truth-value of a proposition into the classical logic. It defines new operators called fuzzy set operators

fridaya@uef.fi

e.g. the t-norm operator which corresponds to the fuzzy AND (intersection). Each t-norm has an associated t-co-norm which corresponds to the fuzzy OR (union).

For illustration, let us suppose that we have 02 fuzzy sets "big" and "old". Let an individual x such that $\alpha_{big}(x) = 0.7$ and $\alpha_{old}(x) = 0.5$. How to determine $\alpha_{big \cap old}(x)$? The corresponding fuzzy model is as follows:

Let T be a t-norm and S a t-co-norm, A and B 02 fuzzy sets. They are respectively defined by:

$T : [0, 1] \times [0, 1] \rightarrow [0, 1]$ and, $S : [0, 1] \times [0, 1] \rightarrow [0, 1]$.

The t-norm is then defined by: $\alpha_{A \cap B}(x) = T(\alpha_A(x), \alpha_B(x))$. The t-co-norm is then defined by: $\alpha_{A \cup B}(x) = S(\alpha_A(x), \alpha_B(x))$.

Modeling of the Authorization Process: Authorization's Functions. The authorization is the process of selecting the most deserving users of the service. It does 02 things: *ordering* users according to the merit and *selecting* the most deserving users according to the available resources for service's access.

Ordering the Users. At a time t, there is a set E of users ($E \subset U$) who wish to access the service. However, the number n of services that can be provided is limited. The aim is to order the users for the system to choose the first n users. For that, the SP defines a number of linguistic variables. Users will be element of sets of Fuzzy values. To simplify the problem, we will only study one fuzzy value: e.g. the value *favorable*.

Let us now consider the universe that contains all the linguistic variables that the SP can consider, and let us define among them the linguistic variable merit that characterizes the merit of each user. In the context of fuzzy logic, it will have a fuzzy value "is deserving" which designates the merit rate of the considered user. The fuzzy set described by the fuzzy value "is deserving" of the linguistic variable merit will be the fuzzy intersection of all fuzzy sets associated with the fuzzy values favorable to each linguistic variable considered by the SP.

Example: Let us consider the linguistic variables *distance* and *age* chosen by an SP, who favors *far* and *aged* users. Then the fuzzy values *favorable* for these 02 variables will be *far* and *aged*, and so the linguistic value "is deseserving" will then be defined by: "*is deseserving*" = $far \cap aged$. In other words, the fuzzy set defined by "is deseserving" will be the fuzzy intersection of fuzzy sets defined by *far* and *aged*.

The linguistic value "is deseserving" involves the logic AND ($\bigwedge$), so we must define the t-norm to be used. From now on, we will instead use the probabilistic t-norm: $T(x, y) = xy$, because it approaches at best the (geometric) average since we want a value that gathers all the truth-values equitably. The SP will then have a m-tuple of characteristic functions $\alpha = (\alpha_i)_{i=1}^{m}$, where α_i is the characteristic function of the fuzzy set whose fuzzy value is v_i. Thus, we introduce the *Ord* function for ordering users defined by:

$$\begin{aligned} Ord \ &: \mathcal{P}(U) \times \left([0, 1]^U\right)^m \rightarrow U^n \\ &(E, \alpha) \mapsto Ord(E, \alpha) = \bar{E} \end{aligned}$$

fridaya@uef.fi

Where:

- E is the set of users to be ordered, those who wish to access to the service.
- α is the m-tuple of characteristic functions; m being chosen by the SP.
- $\bar{E}$ is a n-tuple of users. The users are then ranked in ascending order of their merits in $\bar{E}$; n being fixed by the SP.

Selection of Users. After ordering the users, it is now necessary to classify them according to the capacity of the SP in terms of the number of available services. An SP f has a number p (which can be infinite) of available services. We define the selection function *Select* by:

$$\begin{aligned} Select : U^n \times \bar{\mathbb{N}} &\rightarrow \mathcal{P}(U) \\ (\bar{E}, p) &\mapsto Select(\bar{E}, p) = E_0 \end{aligned}$$

Where:

- $\bar{E}$ is a n-tuple of users. This is the result of the *Ord* function.
- p is the number of attributable services at a given time.
- $\boldsymbol{E_0}$ is the set of users to whom the service is granted.

$$Property : card(E_0) \leq p$$

Practical Case. Let us consider WiABox.Net our WISP. It is assumed that there are 06 successful authenticated users $(u_1 \ldots u_6)$, who wish to benefit the access to the Internet. For exceptionally simplifying this case study, we will assume that the maximum number of attributable services is 04. Our goal is to determine among 06 users the 04 that deserve or merit the most the access to the service. The specificities of the SP are the following:

- On a scale of 0 to 10, the priority of the packageA is 10, the priority of the packageB is 8, the priority of the packageC is 7, the priority of the packageD is 3 and finally the priority of the packageE is 0.5.
- If x denotes the *distance* (in 10^4 m) between the SC and the WAP, then his priority is given by the function $x \mapsto e^{-x}$. The closer we are, the more priority we have.
- The *seniority* gives the priority to access the service by the fact that if you are a Blue (≤ 3 months of loyalty), we have a priority of 10/10. If you are a Veteran (>3 months and ≤ 5 years), then we have a priority of 3/10. Finally if you are an X-Veteran (>5 years), then we have a priority of 7/10. We assume the following information about the users were collected:
  - u_1: packageC, 4.7 km, 1 month;
  - u_2: packageA, 20 km, 3 years;
  - u_3: packageE, 380 m, 2 months;
  - u_4: packageD, 10 km, 5 years;
  - u_5: packageA, 5.5 km, 10 years;
  - u_6: packageB, 8.5 km, 4 years.

fridaya@uef.fi

Linguistic Variables. The specificities of WiABox.Net show 03 linguistic variables that are *package, distance, seniority*. Each of these linguistic variables respectively proposes as a favorable fuzzy value: *best_package, close, satisfactory_duration*. We can then define the characteristic functions α_1 for best package, α_2 for close and α_3 for satisfactory duration.

Fuzzy Value 'Best_package': The characteristic function of the fuzzy set defined by this fuzzy value is given by:

$$\alpha_1(u) = \begin{cases} 1 & if\ package(u) = packageA \\ 0.8 & if\ package(u) = packageB \\ 0.7 & if\ package(u) = packageC \\ 0.3 & if\ package(u) = packageD \\ 0.05 & if\ package(u) = packageE \end{cases}$$

Where $package(u)$ is the package of the user u.

Fuzzy Value 'Near'. The characteristic function of the fuzzy set defined by this fuzzy value is given by:

$$\alpha_2(u) = e^{-distance(u)}$$

Where $distance(u)$ is the distance (in 10^4 m) between the user u and the WAP of the WISP.

Fuzzy Value 'Satisfactory_duration'. The characteristic function of the fuzzy set defined by this fuzzy value is given by:

$$\alpha_3(u) = \begin{cases} 1\ if\ fidelity\,(u) \leq 3\ months \\ 0.3\ if\ 3\ months < fidelity\,(u) \leq 5\ years \\ 0.7\ if\ fidelity\,(u) > 5\ years \end{cases}$$

Where $fidelity(u)$ is the fidelity of the user u.

Calculation of the Merit of the Users.

- *User u1*: The data of the user u_1 are the following: packageC, 4.7 km, and 01 month. Then, the merit of u_1 is $0.7 \times 0.67 \times 1 = 0.469$.
- *User u2*: The data of the user u_2 are the following: packageA, 20 km, and 03 years. Then, the merit of u_2 is $1 \times 0.135 \times 0.3 = 0.040$.
- *User u3*: The data of the user u_3 are the following: packageE, 380 m, and 02 months. Then, the merit of u_3 is $0.05 \times 0.962 \times 1 = 0.048$.
- *User u4*: The data of the user u_4 are the following: packageD, 10 km, and 05 years. Then, the merit of u_4 is $0.3 \times 0.367 \times 0.3 = 0.033$.
- *User u5*: The data of the user u_5 are the following: packageA, 5.5 km, and 10 years. Then, the merit of u_5 is $1 \times 0.576 \times 1 = 0.576$.
- *User u6*: The data of the user u_6 are the following: packageB, 8.5 km, and 4 years. Then, the merit of u_6 is $0.8 \times 0.427 \times 0.3 = 0.102$.

fridaya@uef.fi

Selection of the Users. The results of the previous calculations are $u_1 : 0.469$; $u_2 : 0.040$; $u_3 : 0.048$; $u_4 : 0.033$; $u_5 : 0.576$; $u_6 : 0.102$. This implies the following:

$$Ord(\{u_1, u_2, u_3, u_4, u_5, u_6\}, (\alpha_1, \alpha_2, \alpha_3)) = (u_5, u_1, u_6, u_3, u_2, u_4)$$

Knowing that the limit number of users of our WISP was 04, we then have:

$$Select((u_5, u_1, u_6, u_3, u_2, u_4), 4) = \{u_5, u_1, u_6, u_3\}$$

In conclusion, the access to the service will be granted to users u_1, u_3, u_5 and u_6.

4.3 Accounting

Definition of the Measured Spaces [13].
Measured space of users. Let us consider U the set of users of our system. This set is influenced by the time. For the reasons of simplicity, we will choose as tribe on U, the discrete tribe $\mathcal{P}(U)$ designating all the parts of U. Since the entities are countable, we choose to define on $\mathcal{P}(U)$ the counting measure that is noted μ_U as a measure on this space. Finally, we obtain the measured space $(\mathrm{U}, \mathcal{P}(\mathrm{U}), \mu_{\mathrm{U}})$: space of users.

Measured space of providers. Let F denoting all the SPs of the system. In a similar way to U, one defines the measured space $(\mathrm{F}, \mathcal{P}(\mathrm{F}), \mu_{\mathrm{F}})$: space of providers.

Measured space of the actors. An actor will therefore designate either a provider or a user. We note A all the actors of the system. Then we have the definition $\mathrm{A} = \mathrm{F} \cup \mathrm{U}$. We thus define in a similar way for U, the measured space $(\mathrm{A}, \mathcal{P}(\mathrm{A}), \mu_{\mathrm{A}})$: space of the actors.

Space of time. In our model, we took into account the time dimension because the state of the entities of our system varies over time. The set of instants is $\mathbb{R}_+$. The tribe on $\mathbb{R}_+$ will be $\mathcal{B}_{\mathbb{R}_+}$, the Borelian tribe [13–15] of $\mathbb{R}_+$. We have taken the measure of Lebesgue λ[16] as the most appropriate one. Thus, the measured space $\left(\mathbb{R}_+, \mathcal{B}_{\mathbb{R}_+}, \lambda\right)$ results to the space of time.

Universe space. What is called the universe here is the resultant set of the Cartesian product of the actors' space and that of the time. If we denote Ω the universe, then by definition we will have $\Omega = \mathrm{A} \times \mathbb{R}_+$. We choose as tribe on Ω, the tribe product of $\mathcal{P}(\mathrm{A})$ and $\mathcal{B}_{\mathbb{R}_+}$, $\mathcal{T}_\Omega$ defined by $\mathcal{T}_\Omega = \mathcal{P}(\mathrm{A}) \otimes \mathcal{B}_{\mathbb{R}_+}$. Before defining the measure here, we must consider the following two facts: the space depends on the time and the time depends on the space.

- ***1st case: The space depends on the time***. In the universe, places are governed by the rules of time. To demonstrate this, we can take the Big Bang theory [8]. In our context, we know that A is countable and its cardinal depends on time. This leads to a chaotic model where we cannot describe in detail the variations of A depending on the time. In the remainder of the work, the notation M_t will be used to show that the set M is a function of time.
- ***2nd case***: ***The time depends on the space***. In fact, Albert Einstein has shown that time passes at different speeds depending on where you are. This is the principle of

fridaya@uef.fi

the dilation of the time of the theory of general relativity [8]. In our context, we will suppose that the time defined on the set A does not depend on A. Indeed, we are in a situation where the elements of our system will be people or machines, which are in the same reference frame: Repository. In addition, that the nature of the environment is homogeneous for all actors. To suppose space as a variable of time would require us to take considerations irrelevant to our study. In conclusion: A depends on $\mathbb{R}_+$ but $\mathbb{R}_+$ does not depend on A.

Based on the previous assumptions, we can now define a measure on $\mathcal{T}_\Omega$. $\mathcal{T}_\Omega$ being the tribe product from tribes on A and on $\mathbb{R}_+$, it would be wise to define on $\mathcal{T}_\Omega$, the measurement product from the measurements on $\mathcal{P}(A)$ and on $\mathcal{B}_{\mathbb{R}_+}$. We note it μ_Ω and we will finally have: for $M = M_1 \times M_2 \in \mathcal{T}_\Omega = \mathcal{P}(A) \otimes \mathcal{B}_{\mathbb{R}_+}$, $\mu_\Omega(M) = \mu_\Omega(M_1 \times M_2) = \mu_A(M_1) \times \lambda(M_2)$.

μ_Ω defined on $\mathcal{P}(A) \otimes \mathcal{B}_{\mathbb{R}_+}$ self-extends uniquely to a measure on the tribute product $\mathcal{P}(A) \otimes \mathcal{B}_{\mathbb{R}_+}$.

We then obtain the measured space $(\Omega, \mathcal{T}_\Omega, \mu_\Omega)$: universe space.

For $E' \in \mathcal{T}_\Omega$, i.e. any measurable set E'; we define its measurement $\mu_\Omega(E')$ as follows:

$$\mu_\Omega(E') = \int_{E'} d\mu_\Omega$$

If $E' = E_t \times I$ with $E_t \in \mathcal{P}(A)$ *and* $I \in \mathcal{B}_{\mathbb{R}_+}$.

Thus,

$$\mu_\Omega(E') = \int_{E'} d\mu_\Omega = \int_I \left(\sum_{x \in E_t} 1_A(x) \right) dt$$

At this level, since E_t depends on time, then we first sum on the space's dimension, then on the time dimension. The opposite is not possible according to the reasons given above. Where:

$$\mu_\Omega(E') = \int_I \left(\sum_{x \in E_t} 1_A(x) \right) dt = \int_I \left(\sum_{x \in A} 1_{E_t} \cdot 1_A(x) \right) dt = \int_I \left(\sum_{x \in A} 1_{E_t \cap A}(x) \right) dt$$

$$\mu_\Omega(E') = \int_I \left(\sum_{x \in A} 1_{E_t}(x) \right) dt = \int_I \mu_A(E_t) dt = \int_I card(E_t) dt$$

Notes: The calculation here cannot be exhaustive because, until now we do not have any scheme, which "a priori", can help to describe the variations of E_t according to t (the time).

Definition of Functions

The function unit of period. The modeling of the space of time in a continuous way is not only not accessible but also especially unusable on the practical level since the service flows are most often quantified per unit of time defined by the SP and this unit

fridaya@uef.fi

can itself vary over time. We introduce for this purpose, the function θ that, at any time t, returns the time unit corresponding to the service flow provided at this time. The function unit of period is defined hereafter.

$$\begin{aligned}\theta : \mathbb{R}_+ &\rightarrow \mathbb{R}_+ \\ t &\mapsto \theta(t)\end{aligned}$$

Example: For Internet access, θ is usually constant and is worth one second.

Service utilization rate function. It is defined by the function v as follows:

$$\begin{aligned}v : U \times \mathbb{R}_+ &\rightarrow \mathbb{R}_+ \\ (u,t) &\mapsto v(u,t)\end{aligned}$$

Where $v(u,t)$ is the utilization rate of the service by the user $\boldsymbol{u}$ at the time t.

Example: Let us consider the Internet access service. Here, the function v will be the bitrate (a measure of the quantity of digital data transmitted per time unit), e.g. $v(u,t) = 42Mbits/s$, $\forall (u,t) \in K \times \mathbb{R}_+$ where $K \subset U$ (Here the unit of period corresponds to 1 s (one second)).

Service Consumption's Measure. As we saw earlier, in practice, the continuous model has limitations. We must then consider that for the user u, the value of v at t is conserved on the unit of period provided by θ, i.e. $v(u,t') = v(u,t)$, $\forall t' \in [t, t+\theta(t)]$. Thus, the unit consumption from t will be worth:

$$unit(u,t) = g(u,(t,t+\theta(t))) = v(u,t) \times \theta(t)$$

It is the consumption at t. This defines the unit function. Finally, the function $\hat{\boldsymbol{g}}$ (discrete modeling) by:

$$\begin{aligned}\hat{g} : U \times D &\rightarrow \mathbb{R}_+ \\ \big(u,(t_i,t_f)\big) &\mapsto \hat{g}\big(u,(t_i,t_f)\big) = \sum_{k \in J_{(t_i,t_f)}} unit(u,t_k)\end{aligned}$$

Where the sequence (t_k) is defined by:

$$\begin{cases} t_0 = t_i \\ t_{k+1} = t_k + \theta(t_k)\ \forall k \geq 0 \end{cases} \text{and} J_{(t_i,t_f)} = \big\{k \in \mathbb{N} / t_i \leq t_k \leq t_f\big\}$$

The functions *unit* and $\hat{g}$ being defined, we can now define their respective set versions that are *Unit* and $\hat{G}$.

$$\begin{aligned}Unit : \mathcal{P}(U) \times \mathbb{R}_+ &\rightarrow \mathbb{R}_+ \\ (E,t) &\mapsto Unit(E,t) = \sum_{u \in E} unit(u,t)\end{aligned}$$

fridaya@uef.fi

$$\hat{G} : \mathcal{P}(U) \times D \to \mathbb{R}_+$$
$$\big(E, (t_i, t_f)\big) \mapsto \hat{G}\big(E, (t_i, t_f)\big) = \sum_{k \in J_{(t_i,t_f)}} Unit(E, t_k) = \sum_{k \in J_{(t_i,t_f)}} \sum_{u \in E} unit(u, t_k)$$
$$= \sum_{u \in E} \sum_{k \in J_{(t_i,t_f)}} unit(u, t_k) = \sum_{u \in E} \hat{g}\big(u, (t_i, t_f)\big)$$

Example: In the case of Internet access service, if we consider $\theta(t)$ constant and equal to 1 s, then, for a user u benefiting from a constant bit rate during the time of v $v(u,t) = 2Mb/s$, its consumption at time $t' = 30\,s$ is $unit(u,t') = v(u,30) \times \theta(30) = 2 \times 1 = 2Mb$.

In addition, that from 30 s to 45 s will be:

$$\hat{g}(u,(30,45)) = \sum_{k \in J_{(30,45)}} unit(u, t_k) = \sum_{k \in J_{(30,45)}} 2 = 2 \times (45 - 30 + 1) = 32Mb$$

The function Resource. Considering the WISPs, one defines the resource function noted r that, for a given WISP f, gives the remaining quantity of the principal resource at the time t. We then have:

$$r : F \times \mathbb{R}_+ \to \mathbb{R}_+$$
$$(f,t) \mapsto r(f,t)$$

The principal or main resource is at the origin and of the provided service. Without it, there could be no real, genuine and adequate provision of service. E.g. Let us consider the case of a printing service. If one supposes that the WISP f only has 400 paper sheets at the beginning of the day $(t = 0)$, then $r(f, t = 0) = 400$. The set version of the resource function r is the function R defined by:

$$R : \mathcal{P}(F) \times \mathbb{R}_+ \to \mathbb{R}_+$$
$$(E,t) \mapsto R(E,t) = \sum_{f \in E} r(f,t)$$

The function Service Access Capability. It is necessary to couple to the resource function r, one another called *service access capability* noted c. For a WISP f, the function c specifies the maximum number of critical tools he has for the simultaneous provision of the service. In other words, it characterizes his capacity to simultaneously serve users requesting access to the service. We then define the function c as follows:

$$c : F \to \mathbb{R}_+$$
$$f \mapsto c(f)$$

fridaya@uef.fi

Example: Let us consider a cybercafé that has 05 computers available for a full-time use. The WISP f, here can only serve simultaneously up to 05 users at the maximum. Thus, we have $c(f) = 5$. The set version of the function capacity c is the function C defined by:

$$C : \mathcal{P}(F) \rightarrow \mathbb{R}_+$$
$$E \mapsto C(E) = \sum_{f \in E} c(f)$$

The function Accessibility. This is the function that, at any time t, returns the remaining number of instances of the service that a SP f can simultaneously offer at this time. It is noted a, and somehow, the capacity $c(f)$ is its maximum value. In other words, it is the difference between the overall capacity and the number of service instances already offered. For that, let us introduce the following function o which returns, at any time t the number of instances of service already offered by f:

$$o : F \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$$
$$(f,t) \mapsto o(f,t)$$

Moreover, we will have:

$$a : F \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$$
$$(f,t) \mapsto a(f,t) = c(f) - o(f,t)$$

Example: Let us take a cybercafé, knowing that the WISP f has a maximum of 05 computers and that 03 of them are currently used by the SCs at the time t: we will have $a(f,t) = 5 - 3 = 2$.

Properties: Let Δ denoting a duration defining a period.

1. $\forall (E,t) \in \mathcal{P}(A) \times \mathbb{R}_+,\ R(E \cap F, t) \geq \hat{G}(E \cap U, (t, t+\Delta))$;
2. $\forall (f,t) \in F \times \mathbb{R}_+,\ r(f,t) = 0 \Rightarrow a(f,t) = 0$

The first property translates the fact that the overall consumption in a locality E of all users over a period can never exceed the services production of their WISPs. The second reflects the fact that one cannot access the service when the resource is exhausted and a fortiori consume it; the accessibility of the service is dependent on the resource.

Vocabulary: Let $(f,t) \in F \times \mathbb{R}_+$ be fixed:

- When $r(f,t) = 0$ then "the service is unavailable".
- When $a(f,t) = 0$ then "the service is inaccessible".
- "An unavailable service is, a fortiori, inaccessible".

fridaya@uef.fi

4.4 Pricing

Next to the accounting of the services, it is question of fructifying the consumptions of SCs. For that, the WISPs will have to define a pricing policy or a tariff plan, i.e., to associate tariffs with the various predefined packages [17, 18]. A WISP f defines the set FO_f of all his packages for the considered service. It is then a question of defining a function Pr that will link a package to a tariff for a given WISP $f \in F$.

$$\begin{aligned} Pr : FO_f &\rightarrow \mathbb{R} \\ x &\mapsto Pr(x) \end{aligned}$$

Each WISP will define this function Pr, which will thus make a correspondence between its packages and the unit prices of the service.

Note: For the definition of service packages, the WISP will rely on the qualities of the service.

4.5 Billing

The service billing [19] is a part of the PBP model. Let us note that charging in our context consists to apply a price to a consumption. We then define the following function:

$$\begin{aligned} Bil : U \times \mathbb{R}_+ &\rightarrow \mathbb{R}_+ \\ (u, q) &\mapsto Bil(u, q) = p \end{aligned}$$

- u is the logical user whose features include the package;
- q is the consumption of the service by the user u;
- p is the price calculated for the quantity q consumed by the user u.

4.6 Payment: Prepaid, Postpaid and on-the-Go

Prepaid. The access to the service is conditioned by the billing. In other words, the user pays first before actually consuming the service. This billing mode is similar to that of a fuel pump at the service station. When the user pays an amount of money, his tank is topped up with the fuel (the resource in our case). His tank will progressively empty when he will drive during a certain period. As illustrated in Fig. 2, after the fuel finishes or reaches a certain level, he will top up again the fuel in the tank. Here, consuming the service consists to driving the car.

Postpaid: The payment method illustrated in Fig. 3 below is made after the consumption of the service.

On the Go: The payment is made as the user consumes the service.

fridaya@uef.fi

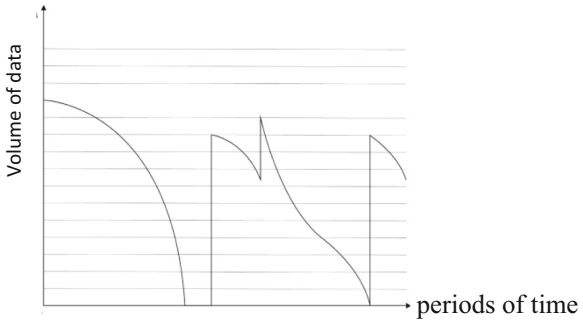


Fig. 2. Illustrative representation of the prepaid mode

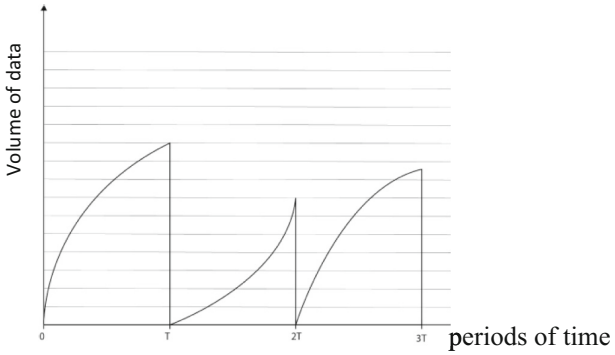


Fig. 3. Illustrative representation of the postpaid mode

5 Discussions and Open Issues

This model can be used to build generic profiles of users in a multi-operator (telecom or internet) context. Consumer scoring can thus be optimized by coupling a machine-learning engine at the Authorization component. To be more efficient, we would obtain classes of users to address offers of targeted services, ranging from free to paid ones. In a rural community, health centers, schools and colleges can benefit from free services, while small businesses can financially contribute to help the former. This is considered by the same platform that implements a unique model: 3APBP.

We can also use this model for scoring merchants and customers of an e-marketplace or e-commerce platform; it would thus serve as a basis of a services' recommendation system to clients and service providers.

Finding a family of mathematical functions to model the Payment component is not easy, that is why we have instead illustrated it with graphs while continuing our investigation.

fridaya@uef.fi

6 Conclusion and Future Works

In this paper, we have proposed a formal approach based on Black Box Specification Method by Entities for modeling first network services, then the access and control mechanisms and functions of these services. This resulted into the 3A-PBP (authentication, authorization, accounting - pricing, billing and payment) model which includes each element of the AAA standard protocol augmented by the PBP services. Each entity, modeled as a mathematical function, represents one component of the whole system. The output parameters of each component are the input parameters of another one.

We first gave a definition of a network service, then we built the mathematical ecosystem based on set and fuzzy logic theories and finally we brought out all the relevant functions and properties that contribute to the solution of our initial problem.

This work forms the basis of a long-term project for the operationalization of wireless community networks in developing regions. This is why we have designed it more generic to facilitate its scaling for better adaptation. Our wireless community network model must support the operation of proximity and long-range digital services with minimal customization effort. The implementation of this 3APBP model, which extends the best-known AAA model, is in progress, in order to compare its performance with that of some AAA frameworks such as FreeRADIUS.

Acknowledgments. Special gratitude respectively goes to Alexander von Humboldt Foundation; GIZ-CIM Organisation; the University of Applied Sciences Bonn-Rhein-Sieg and the CETIC Project of the University of Yaounde 1 for their various supports.

References

1. Luiz, A.D.: Wireless community networks: evolution and technical challenges. In: Workshop on Nationwide Internet Access and Online, Applications, Dhaka, Bangladesh (2004)
2. Center for Neighborhood Technology: Community Wireless Networks: Cutting Edge Technology for Internet Access (2006)
3. Frappier, M., St-Denis, R.: EB3: an entity-based black-box specification method for information systems. Softw. Syst. Model. **2**, 134–149 (2003)
4. Freifunk. http://freifunk.net/. Accessed December 2018
5. de Laat, C., Gross, G., Gommans, L., Vollbrecht, J., Spence, D.: Generic AAA Architecture, Internet Society, RFC2903 (2000)
6. The FreeRADIUS Project. http://freeradius.org/. Accessed June 2018
7. freeDiameter. http://www.freediameter.net. Accessed June 2018
8. Hawking, S.: A Brief History of Time. Bantam Books, New York (1988). Physics Bulletin
9. Matt, B.: Computer Security: Art and Science. Addison-Wesley, Reading (2004)
10. Dernoncourt, F.: Introduction à La Logique Floue (Introduction to fuzzy logic) (2011). developpez.com
11. Gacogne, L.: Éléments de logique floue (Element of Fuzzy logic). Hermès (1997)
12. Leekwijck, W.V., Kerre, E.E.: Defuzzification: criteria and classification. Fuzzy Sets Syst. **108**(2), 159–178 (1999)

fridaya@uef.fi

13. Berberian, S.K.: Borel spaces. The University of Texas at Austin (1988)
14. Navara, M.: An algebraic generalization of the notion of tribe. Fuzzy Sets Syst. **192**, 123–133 (2012)
15. Anatolij, D.: MV-observables and MV-algebras. J. Math. Anal. Appl. **25**(9), 413–428 (2001)
16. Cenap, D., Birsen, S.: Non-newtonian comment of lebesgue measure in real numbers. J. Math (2017). https://doi.org/10.1155/2017/6507013
17. Chen, X.: Resource allocation and pricing in virtual wireless networks. Masters Theses 1896, University of Massachusetts (2014)
18. Belghith, A., Trabelsi, S., Cousin, B.: Realistic per-category pricing schemes for LTE users. In: Proceedings of IEEE 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WIOPT) (2014)
19. Lakhtaria, K.I., Jani, N.N.: Design and modeling billing solution to next generation networks. Int. J. Adv. Netw. Appl. **1**(5), 290–294 (2010)

fridaya@uef.fi

Downlink High Quality 72 GHz Millimeter

Fawziya Al-Wahaibi(✉) and Hamed Al-Rwashidi(✉)

Brunel University London, West of London, Uxbridge UB8 3PH, UK
{fawziya.al-wahaibi,Hamed.Al-Raweshidy}@brunel.ac.uk

Abstract. This paper investigates the performance of downlink transmission for millimetre wave in radio over a fiber system. Firstly, a high clarity optical 72 GHz millimeter wave signal is generated by using a frequency quadrupling technique based on carrier suppression. The optical carrier suppression ratio of the generated 72 GHz signal is 71 dB. Thus, highly suppression of optical carrier and harmonics lead to obtain high purity millimeter-wave which will decrease suffering from impairment due to fiber dispersion. Then, the performance of a modulated millimeter in downlink data transmission is investigated by using an additional optical modulator. Two different optical modulators are used to compare the performance of achieved mm-wave; Mach Zehnder modulator and Electro Absorption modulator. It is found that the signal when modulated using EAM has better performance than when this done with MZM. The Max Q factor for 80 km are nearly up to 20 when utilizing the former, but just 2.3 when the latter is deployed.

Keywords: Millimeter wave over fiber · Optical carrier suppression ratio · Mach Zehnder modulator · Electro Absorption Modulator

1 Introduction

With the huge growth in both the number of wireless communication users and the bandwidth required per subscriber, migrating to higher frequencies from lower radio frequency to millimetre-wave carriers, is an essential solution.

Radio-over-Fiber (RoF) technology provides a way for the convergence of wireless and optical systems by merging the technical advantages of a wireless and optical system. RoF has introduced system flexibility and enlarged the coverage area without increasing the cost and complexity. However, the demand for greater bandwidth that is used in backhaul of the network and the subscribers' various applications has meant that the research industry has needed to utilize higher frequencies. Hence, millimeter wave frequencies have become a solution to increase the power of ROF technology. Millimeter-wave (mm-wave) for ROF technologies have been a powerful area of research and investigation over the past decade, which combines the benefits of mm-wave frequencies and ROF technologies [2]. Mm-wave frequencies are allocated between 30 GHz and 300 GHz in the spectrum frequencies which offer wide bandwidth that is available for carrying ultra-high data rates in wireless transmission at this range. In sum, mm-wave over fiber has ability to overcome the difficulties of spectral congestion at lower frequency ranges that provide broadband wireless access services [1].

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 336–344, 2019.
https://doi.org/10.1007/978-3-030-22868-2_25

fridaya@uef.fi

An example of an mm-wave for ROF is shown in Fig. 1, where there is a CS that contains data resource, optical generation, transmitter (TX) and a receiver (RX) with a laser and photodiodes (PDs).

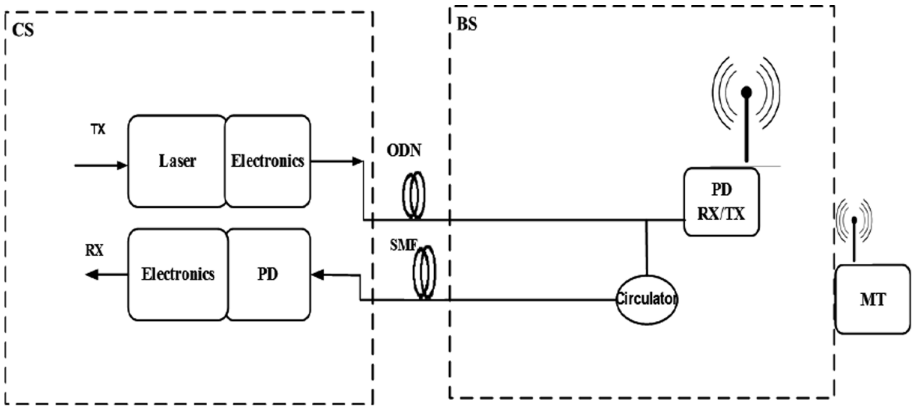


Fig. 1. General structure of millimetre wave over fiber

1.1 Basic Structure of Mm-Wave Over Fiber

The CS is a part of the trunk network, which has connection to the data network provider and the Internet, as well as having switching capabilities. The CS converts the electrical signal from the electrical domain to optical domain then transmits it over fiber to BS, the direction of which is called downlink. The BS includes a photo detector in order to obtain the desired mm-wave signal in the electrical domain and radiates it over small areas using microwave antennas into a mobile terminal (MT) end user in the mm-wave bands. In the other direction, termed the uplink direction, the BS receives the signal from an MT in mm-wave bands, which can be modulated before sending it through fiber to CS [1].

1.2 Downlink Data Transmission

In downlink data transmission there are two important processes: generating millimeter (mm)-waves and downlink data modulation [2]. Optical generation of a millimeter mm-wave signal is one of the main stages in mm-wave over fiber technology in order to get a high quality signal. Many approaches have been reported recently to generate such waves, including: direct modulation, optical heterodyning, and external modulation. Optical generation of mm-waves using the external modulation technique is likely to be more accurate, with such schemes having shown the most potentiality for high-frequency signal generation owing to the large tunable range, excellent stability, high spectral purity and lower system cost [3]. The external modulation method implementation comprises a continuous wave (CW) laser and an external modulator (EM), which modulates the laser light with a radio frequency (RF). A Mach Zhender Modulator (MZM) an Electro Absorption Modulator (EAM) and a Phase Modulator are

fridaya@uef.fi

examples of devices that can be used as external modulators in RoF [1]. Two techniques are demonstrated for generating mm-waves using external modulation, with optical frequency multiplication, and without it [2]. External modulation leads to the generation of higher order harmonics since it is a source of non-linearities in the transmission characteristic. The generation of harmonics associated with the desired signal depends on setting of bias voltage and modulation index [4]. Generating mm-wave signals using an external modulator was proposed by O'Reilly in 1992, where the signal was generated by biasing a Mach Zehnder Mmodulator (MZM) to suppress even-order optical sidebands. He implemented the frequency doubling based on carrier suppression to get a microwave signal of 36 GHz generated from a drive signal of 18 GHz [5]. Different modulation schemes using MZM have been deployed to generate mm-waves, such as single sideband (SSB), double sideband (DSB) and optical carrier suppression (OCS), with each having different resistance for the fibre dispersion [6, 7]. A among them, OCS is preferred because of its advantage in enhancing the receiver sensitivity [8]. In addition to that, the OCS modulation scheme has a simple configuration and low-frequency bandwidth requirement for both the electrical and optical components. Moreover, desired frequency multiplication factor can be accomplished by the proper setting of the RF LO phases, modulation index and the biasing voltages of the MZMs. Different methods of combinations of MZMs were utilised for the generation of mm-waves with desired frequency multiplication of input frequency by different settings of the modulation index and RF LO phases. In [9], the researcher used two cascaded MZMs based on a multiplication factor of four to generate a photonic microwave frequency. A multiplication factor of six was demonstrated to generate mm-waves for ROF in [10]. In [11], a frequency sextupling scheme for high-quality optical millimetre-wave signal generation based on carrier suppression was investigated. In [12], two cascaded dual-electrode MZMs, MI-insensitive were utilised and a filterless frequency octupling scheme based on two parallel dual-parallel Mach Zehnder modulators. In [13], the frequency decoupling techniques using two parallel DD-MZMs and an interleaver were deployed to get a 60 GHz OCS mm-wave signal. A 72 GHz mm-wave was generated using quadrupling techniques by using three different MZMs with OSSR of 36 dB. However, these proposed systems need either more than one optical filter or external modulators to remove undesired optical sidebands which significantly increase system complexity and cost. Therefore, the frequency quadrupling technique without optical filters is of great interest. A different approach that utilized to modulate the data with mm-wave in downlink transmission as it is summarized in reference [14]. In this work, we investigate a tunable photonic frequency quadrupling technique for the generation of mm-waves at 72 GHz by two parallel DD-MZMs which is based on carrier suppression. Then, the performance of the downlink data modulation is investigated using two types of optical modulation, EAM and MZM, in order to obtain best performance of a modulated mm-wave.

fridaya@uef.fi

2 Principle

The technique of frequency quadrupling (see Fig. 1). A continuous wave (CW) laser with finite spectral width is injected into the dual parallel configuration of upper and lower MZM, having extinction ratio (ER) of 60. Both the MZMs are biased at the Minimum Transmission Point of the MZM. Also, upper and lower MZMs are driven by an RF local oscillator with an equal phase shift of 90° introduced between the RF-driving signals applied to each MZM drive electrodes. The output of upper MZM contains second and fourth sidebands. Then lower MZM is driven by the same RF signal with 180° phase difference, this will causes the polarities of fourth order sidebands at the output of lower MZM to be in opposition to those at the output of upper MZM. Then, the outputs of the MZMs are combined using an optical coupler, which has only one band output which is second order sidebands, 72 GHz. Therefore, after the output coupler, the odd-order sidebands generated by lower MZM cancel out the odd-order sidebands generated by MZ-up and only the optical sidebands of the order of second, sixth, and tenth remain (see Fig. 2c). Finally the second order sidebands are detected using photo detector and the detected signal is observed using an electrical spectrum analyzer.

3 Simulation Setup and Analysis

The Opt system software is used to generate and investigate the performance of 72 GHz in downlink transmission. The simulation is divided into two parts: firstly, where the mm-wave is generated at 72 GHz. Secondly, the performance of the obtained mm-wave is investigated when it is modulated with data using MZM. Then, it is repeated to investigate the performance of this wave when it is modulated with data using EAM.

3.1 Generating Mm-Wave at 72 GHz

The simulation set up for generating 72 GHz mm-waves (see Fig. 2), with the parameter simulation set up being as displayed in Table 1. The light signal of spectral width of 10 MHz is emitted and injected into the MZMs using the CW laser source. The upper and lower MZMs are driven by 18 GHz with a phase shift of 90°. Both MZMs are biased with a switching bias voltage of 4 V. The biasing voltage for both MZMs is set to zero in order to overcome the chirp issues is caused by MZM. Because both MZMs are biased at the peak operation point, only carrier and even order sidebands are generated (see Fig. 2b).

Table 1. Parameters setup for generating 72 GHz.

Device	Parameter	Value
CW-laser	Optical signal frequency	193.1 THz
	Optical power signal	10 dBm
	Spectral width	10 MHz

(*continued*)

fridaya@uef.fi

Table 1. (*continued*)

Device	Parameter	Value
LO	Radio frequency	18 GHz
MZM	Bias voltage	0
	Switching bias voltage	4 V
	RF voltage	4 V
PIN-PD	Responsivity	7 A/w

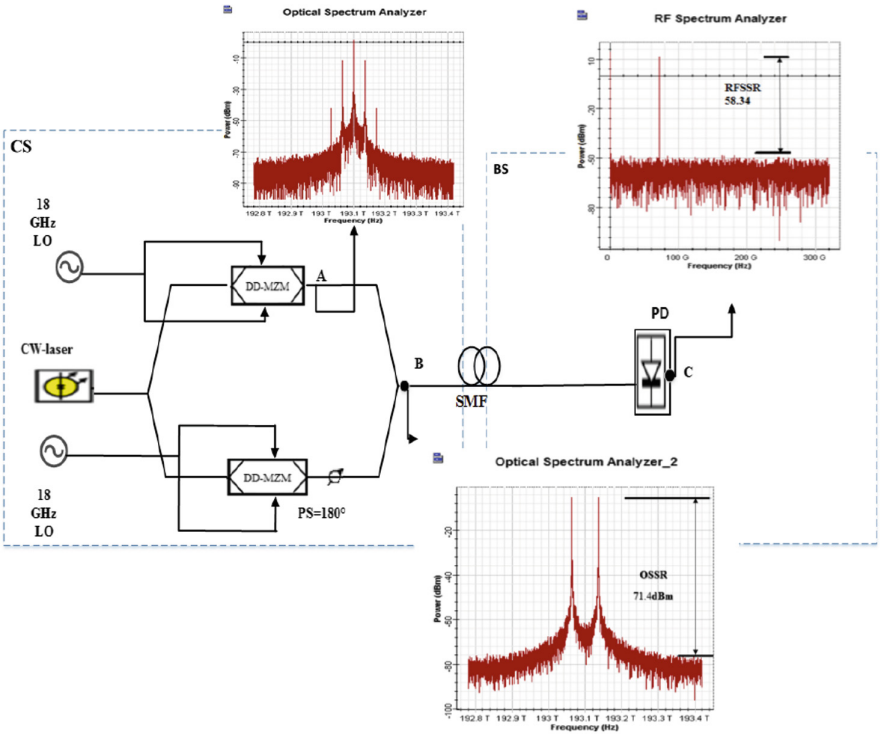


Fig. 2. Simulation design to generate 72 GHZ, A. mm-wave without carrier suppression in the optical domain, B. Carrier suppressed mm-wave in optical domain, C. mm-wave at 72 GHz in the electrical domain.

Figure 2a illustrates the generated mm-wave related to 193 GHz using quadrupling MZM techniques with carrier, second, and fourth sidebands. It shows that, the OSSR of obtained mm-wave in the optical domain is 71.4 dB, which is a very clear signal without a carrier and sidebands. While Fig. 2b shows the desired mm-wave at 72 GHz for the electrical domain with an RFSSR of 60 dB. From the figures, the highly quality of the obtained signals can be seen which lead to avoid strongly use the electrical filter at the receiver. Thus, the obtained mm-wave has greater immunity to chromatic dispersion that caused by fiber.

fridaya@uef.fi

3.2 Downlink Data Transmission

To investigate the quality of the generated optical mm-wave over a single mode fiber (SMF), the system is simulated using Opt system software with additional optical devices. The obtained mm-wave is modulated with a baseband signal in two different scenarios. In the first scenario, the obtained mm-wave signal is modulated with data using a single drive (SD-MZM), as shown in Fig. 3. For the second scenario, the acquired mm-wave signal is modulated using EAM, as shown in Fig. 4 in order to compare the performance of these waves between two types of optical modulators. The same simulation parameters set up are used as shown in Table 1, with additional parameters set up for additional components that are added. Moreover, two values of PD responsivity are considered to check the validity of the optical device used to examine the performance of the obtained mm-wave in downlink transmission of 7 A/W and 1 A/W.

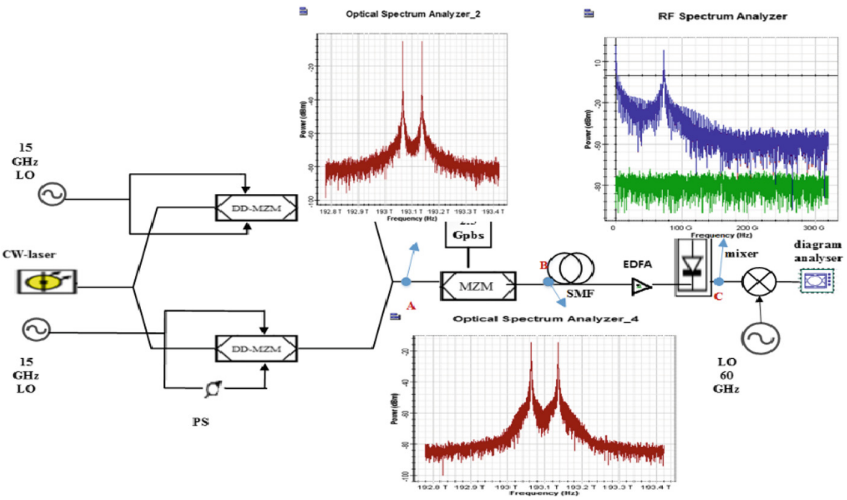


Fig. 3. Simulation design using MZM for data modulation A. mm-wave in optical domain, B. modulated mm-wave, C. mm-wave in electrical domain.

A. PD Responsivity at 7 A/w

After the mm-wave is generated, it will be modulated with 2.5 Gbps data by using SD-MZM as it is shown in Fig. 3B. Then it will be transmitted through SMF to PD in order to extract the mm-wave in electrical domain as it is shown in Fig. 3C. Same steps is repeated when the mm-wave is modulated with 2.5 Gbps by using EAM as it is presented in Fig. 4. The modulated mm-wave is shown in Fig. 4B. Then it will be received by PD to obtain the mm-wave in electrical domain as it is shown in Fig. 4c. The both systems are simulated to observe the performance of mm-wave when the responsivity of PD is 7 A/w. Figure 5 shows the max Q factor of modulated mm-wave signal has been achieved when it is modulating by EAM and MZM. As it is seen in Fig. 5, at 20 km the max Q factor when EAM is used to modulate optical mm-wave with data is 38.1, and max Q factor when the mm-wave is modulated using MZM is 28.

fridaya@uef.fi

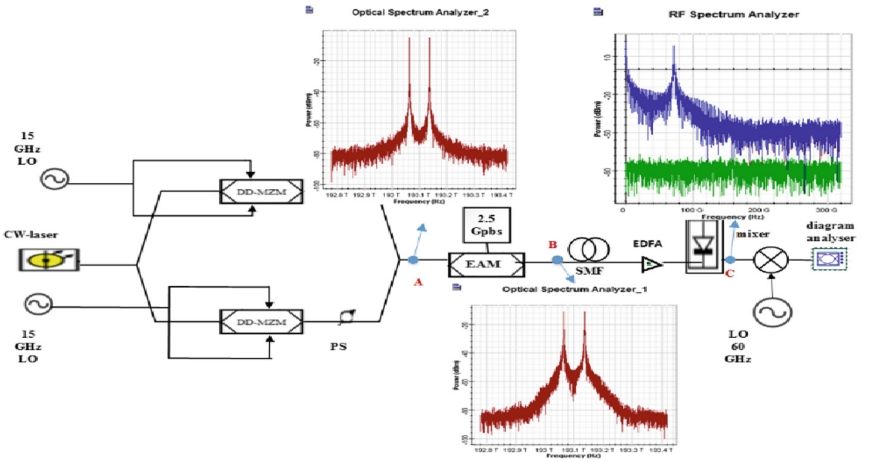
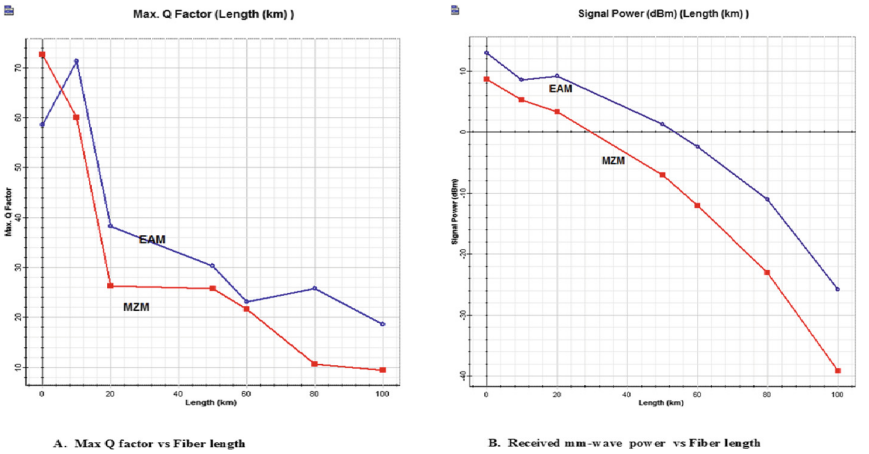


Fig. 4. Simulation design using EAM for data modulation A. simulation design using MZM for data modulation A. mm-wave in optical DOMAIN, B. modulated mm-wave, C. mm-wave in electrical domain.



A. Max Q factor vs Fiber length

B. Received mm-wave power vs Fiber length

Fig. 5. Max Q factor vs fiber length when PD responsivity is 7 A/W

However the Max Q factor starts to decrease when the modulator is MZM to reach 9.02 at 100 km and 20 at 100 km when the system used EAM. The signal power as it is seen in Fig. 5B for both scenarios decease slightly at 20 km the modulated mm-wave signal power using EAM is 9.3 dBm while the received mm-wave power using MZM is 4.8 dBm. However the received signal power for modulated mm-wave have shared same power decreasing as long as fiber length increases. At 80 km, the modulated mm-wave power by using EAM is −11 dBm but for MZM the received signal power reach −25 dBm. Thus the mm-wave power using EAM is slightly higher than received modulated mm-wave power by using MZM modulator.

fridaya@uef.fi

B. PD Responsivity at 1 A/w

In addition for that; the performance of modulated mm-wave in two cases of two types of optical modulation is investigated when responsivity of PD is 1 A/W. It is found that, the system using EAM almost have same high quality performance of obtained mm-wave when PD responsivity is 7 A/W as it is shown in Fig. 6a. For example; at 80 km, the max Q factor for design using EAM is 18, while using MZM the max Q factor is 2.3. However the effect of reducing the responsivity of PD will appears clearly in received electrical powers as it is presented in Fig. 6b. The design using EAM, the electrical received power of generated mm-wave is −18 dBm for 20 km fiber length, and for 80 km almost the mm-wave power reached −36 dBm. On other hand, the received signal power reached −80 dBm for 100 km for second scenarios.

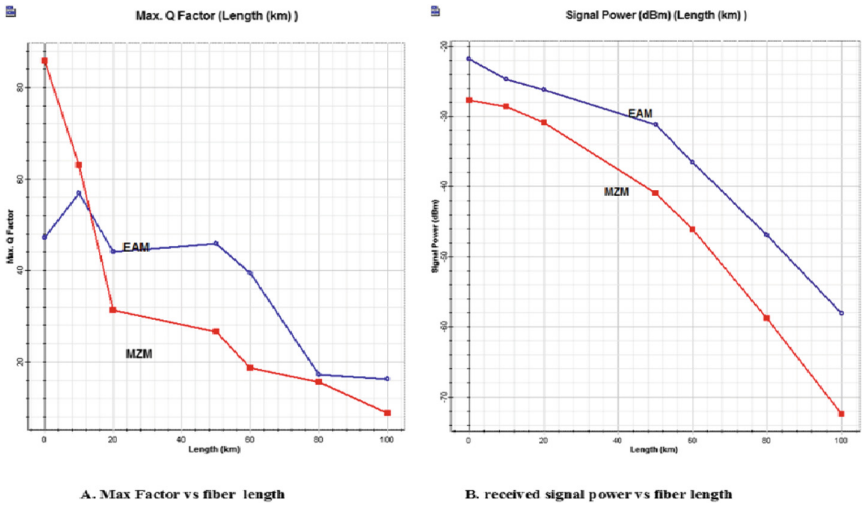


Fig. 6. Max Q factor vs fiber length when PD responsivity is 1 A/W

Thus, the received electrical power of modulated mm-wave in second scenario is very week compared to what is achieved in the first scenario as it is seen in Fig. 6b performance of mm-wave using EAM has better performance than the mm-wave than is obtained by using MZM according to the result obtain. On other direction, the effect of reducing the responsivity of PD effect the received power of the signal.

4 Conclusion

This work approved that 72 GHz can be generated with high quality using quadrupling techniques. The optical carrier suppression ratio (OSSR) of obtained 72 GHz mm-wave signals is 71 dB. High suppression of the optical carrier and harmonics leads to decrease impairment due to fiber chromatic dispersion. In addition for that; with high quality of obtained signal, the filter can be avoided at the receiver which will reduce the cost of implementing studied design. Among Modulating mm-wave in downlink

fridaya@uef.fi

transmission, this work demonstrated that modulation mm-wave with data using EAM optical modulator has achieved high quality modulated mm-wave which can be transmitted over 80 km with max Q factor around 20 with acceptable received electrical power. Modulation mm-wave with data using EAM don't affected by reducing the responsivity of PD. On other hand modulating the generated mm-wave by MZM requires increasing the value of responsivity of PD.

References

1. Varghese, A., Mohammed, H., Lajos, H.: Millimeter-wave radio over fiber optical up conversion techniques relying on link nonlinearity. First Quart. 2016 IEEE Commun. Surv. Tutor. **18**(1), 29–53 (2016)
2. Joaquın, B., Gerardo, C., Ivan, A., Alejandro, A., Gabriel, C.: Millimeter-wave frequency radio over fiber systems: a survey. IEEE Commun. Surv. Tutor. **15**(4), 1593–1619 (2013)
3. Guohua, Q., Jianping, Y., Joe, S., Stéphane, P., Claude, B.: Generation and distribution of a wide-band continuously tunable millimeter-wave signal with an optical external modulation technique. IEEE Trans. Microw. Theory Tech. **53**(10), 3090–3097 (2005)
4. Wangzhe, L., Jianping, Y.: Investigation of photonically assisted microwave frequency multiplication based on external modulation. IEEE Trans. Microw. Theory Tech. **58**(11), 3259–3268 (2010)
5. O'Reilly, J.J., Lane, P.M., Heidemann, R., Hofstetter, R.: Optical generation of very narrow linewidth millimetre wave signals. Electron. Lett. **28**(25), 2309–2311 (1992)
6. Jing, L., Aijun, W., Huixing, Z., Xiaome, H.: A novel MI-insensitive and filterless frequency octupling scheme based on two parallel dual-parallel Mach-Zehnder modulators. Optik **125**, 6887–6890 (2014)
7. Xie, J., Huang, X.G., Tao, J.: A full-duplex radio-over-fiber system based on a novel double-sideband modulation and frequency quadrupling. Opt. Commun. **283**(6), 874–878 (2010)
8. Jianjun, Y., Zhensheng, J., Lilin, Y., Yikai, S., Gee-Kung, C., Ting, W.: Optical millimeter-wave generation or up-conversion using external modulators. IEEE Photonics Technol. Lett. **18**(1), 265–267 (2005)
9. Jian, Z., Hongwei, C., Minghua, C., Tianliang, W., Shizhong, X.: A photonic microwave frequency quadrupler using two cascaded intensity modulators with repetitious optical carrier suppression. IEEE Photonics Technol. Lett. **19**, 1057–1059 (2007)
10. Mohmoud, M., Xiupu, Z.,Bouchaib, H., and Ke, W.: Frequency sixupler for millimeter-wave over fiber systems. OSA Opt. Express **15**, 10141–10151 (2008)
11. Qin, Y., Junqiang, S.: Frequency sextupling technique using two cascaded dual-electrode Mach-Zehnder modulators interleaved with Gaussian optical band-pass filter. Opt. Commun. **285**, 2911–2916 (2012)
12. Jian, Z., Hongwei, C. Minghua, C., Tianliang, W., Shizhong, X.: A photonic microwave frequency quadrupler using two cascaded intensity modulators with repetitious optical carrier suppression. IEEE Photonics Technol. Lett. **19**, 1057–1059 (2007)
13. Muthu, K., Raja, A., Sevendran, S.: Optical generation of millimeter waves through frequency decupling using DP-MZM with RoF transmission. Opt. Quant. Electron. **51**, 1572–817X (2017)
14. Chun-Ting, L., Po-Tsung, S., Jason, C., Wen-Qiang, X., Peng-Chun, P., Sien, C.: Optical millimeter-wave signal generation using frequency quadrupling technique and no optical filtering. IEEE Photonics Technol. Lett. **20**(12), 1027–1029 (2008)

fridaya@uef.fi

Study the Performance of Optical Millimetre Wave Based on Carrier Suppressed by Using an Inverted Optical Filter

Fawziya Al-Wahaibi(✉), Rajagopal Nilavalan, and Hamed Al-Rwashidi(✉)

Brunel University London, West of London, Uxbridge, UK
{fawziya.al-wahahibi,Hamed.Al-Raweshidy}@brunel.ac.uk

Abstract. In this paper, the method of supressing the carrier for optical millimetre wave signal using inverted optical filter is investigated. Therefore, the quality of obtained 60 GHz millimetre wave (mm-wave) signal based on carrier supersession using inverted optical filter is demonstrated. Specifically, a simple and cost effect design is used to generate it by just using a single dual drive Mache-Zehnder modulator (DD-MZM) and an inverted optical filter. This work has two simulation schemes; the first pertains is generating mm-wave signal based on carrier suppression, where in the first, the Optical spurious suppression ratio of 30.56 dB is achieved. In the second simulation scheme the performance of downlink data transmission of 2.5 Gpbs with obtained mm-wav over a single mode fibre transmission line is analysed. Whereby, two scenarios are considered to modulate the data with obtained optical mm-wave in the obtainable design. First scenario is modulating the data directly with DD-MZM while generating the optical mm-wave, second scenario when the data is modulated with obtained optical mm-wave signal. In this study, measuring the max Q factor for different bit rates are considered to observe the performance of the obtained mm-wave signal over the single mode fibre. With the proposed design for 2.5 Gbps, the max Q factors of 21 over 20 km fibre length are achieved. In addition for that Max Q factor of obtained signal is measured with different bit rates at 2 km, where for 10 Gbps, max Q Factor of 8 can be achieved. This design approved that the mm-wave signal can be generated effectively using inverted optical filter.

Keywords: Millimeter wave · Mach–Zehnder modulator · Inverted optical filter · Optical carrier suppressed · Mm-wave over fiber

1 Introduction

Mm-wave over fibre is a promising technique, which merges the benefits of mm-wave frequencies and fibre technology. That is, optical fibre technology provides high bandwidth and supports long transmission links, while the mm-wave frequencies offer large bandwidths in the wireless domain as well as overcoming the problem of spectral congestion at lower frequency ranges that provide broadband wireless access services [1]. There are different methods that has been used to generate mm-wave; direct modulation, optical heterodyning, Four Wave Mixing (FWM) and external modulation

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 345–357, 2019.
https://doi.org/10.1007/978-3-030-22868-2_26

fridaya@uef.fi

(EM) [2]. The optical generation of mm-wave using the external modulation technique is likely to be more accurate, because of its stability and low phase noise characteristics [3]. The structure of implementing EM to generate mm-wave involves of a Continuous Wave (CW) laser followed by an EM that modulates the laser light with radio frequency. The EM used in this technique can be an intensity modulator a (Mach-Zhender Modulator (MZM) or an Electro Absorption Modulator (EAM)), or a Phase Modulator (PM) [4]. The EM generates higher order harmonics which is required to implement Optical Frequency Multiplication (OFM). OFM technique has been proposed to produce mm-wave signals at frequencies above the EM bandwidth using electrical signals (fi). There are two main approaches of using OFM technique. The first is to take advantage of the generation of higher order harmonics with proper bias and modulation index. The second approach is utilizing the OFM technique for suppressing the carrier [4]. Different modulation schemes by using EM have been deployed to generate mm-waves, such as single sideband (SSB), double sideband (DSB) and optical carrier suppression (OCS), with each having different resistance for the fibre dispersion [5, 6]. It has been found that, OCS has an advantage in improving the receiver sensitivity. Furthermore, the OCS modulation scheme has a simple configuration and low-frequency bandwidth requirement for both the electrical and optical components [7]. Different methods have been demonstrated using external modulation to supress the optical carrier for example; as mentioned in study [7]. Frequency sixutpler techniques is used for millimetre-wave over fiber systems as in reference [8] and [9]. In study [10] frequency sextupling scheme was used to generate high-quality optical millimetre-wave signal based on carrier suppression without optical filter, in [11] two cascaded MZMs with a Gaussian optical BPF were used to supress the optical carrier, MI-insensitive and filterless frequency octupling scheme based on two parallel dual-parallel Mach–Zehnder modulators were used for carrier suppression [12]. In study [13] millimetre waves are obtained by frequency decupling using DP-MZM in RoF transmission. In study [14], the mm-waves are generated by using external modulation and Fiber Brag Grating (FBG) FBG to remove the carrier in order to obtain tunable millimetre wave signal of (32 GHz, 44.6 GHz and 49 GHz). However, the issue of using FBG is sensitive to temperature, which is required a complex control system to stabilize its operation. Furthermore, it consumed due to the FBG reflection [6]. In this paper, mm-wave of 60 GHz is generated by using DD-MZM with quadrupling techniques. Then the carrier is supressed by using inverted optical filter, then the performance of generated mm-wave is investigated in order to validate the quality of obtained signal by this method. At the output of DD-MZM, two strong optical harmonic related to optical carrier are generated at the output of DD-MZM. Thus a frequency spacing of four times of driving RF to the single DD-MZM is obtained. OPTISYSTEM software is used in this study since it is a software that enables researchers to plan, test and simulate advanced optical networks effectively.

fridaya@uef.fi

2 Principle

The proposed design of generating a 60 GHz mm-wave with inverted optical filter is shown in Fig. 1. An electrical drive signal is applied to a Dual Drive –MZM (DD-MZM), which is biased to supress the odd-order optical sidebands. An inverted optical filter is connected at the output of DD-MZM to remove the optical carrier and two second order optical sidebands are obtained at the output of inverted optical filter. A beat signal with four times the frequency of electrical drive signal is obtained at the PD.

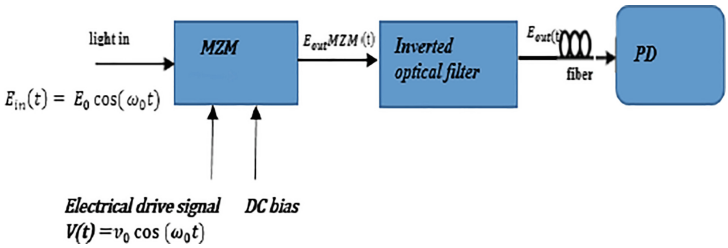


Fig. 1. The principle of generating millimeter wave based on carrier suppression

As it is known the input electrical field at the input of MZM can be expressed by

$$\boldsymbol{E_{in}(t) = E_0 \cos(\omega_0 t)} \tag{1}$$

The output of dual port MZM $E_{out}\,MZM(t)$ can be expressed by

$$\boldsymbol{E_{out}\,MZM(t) = E_0 \cos\left[\frac{\phi V(t)}{2}\right].\cos(\omega_0 t)} \tag{2}$$

Where electrical field amplitude is donated by $\boldsymbol{E_0}$ ***and*** $\boldsymbol{\omega_0}$ indicates angular frequency of input optical carrier, applied electrical drive voltage is indicated by V (t). $\phi \boldsymbol{V(t)}$ is optical phase difference caused by V(t) between the two arms of the MZM. *If* the MZM is driven by sinusoidal electrical signal and biased with a constant dc voltage $\phi \boldsymbol{V(t)}$ is expressed by

$$\phi \boldsymbol{V(t)} = \emptyset_0 + \frac{\pi}{\boldsymbol{V_\pi}}.\boldsymbol{V_e \cos(\omega_0 t)} \tag{3}$$

$\emptyset_0$ is a constant phase shift determined by the dc-bias voltage, $\boldsymbol{V_\pi}$ is the half wave voltage. Where $\boldsymbol{V_e}$ and $\boldsymbol{\omega_0}$ are amplitude and angular frequency of the electrical drive signal, respectively. Considering phase modulation index as β which $= \frac{V_0}{V_\pi} \cdot \frac{\pi}{2}$, while the Bessel function of first kind of order n is noted as $\boldsymbol{J_n}$. The electrical field of the output optical signal can be written when substituting (3) in (2) which will be equal to:

fridaya@uef.fi

$$
\begin{aligned}
&\boldsymbol{E_{out}\,MZM(t)} \\
&\quad = \boldsymbol{E_0\,cos}\left(\tfrac{\emptyset_0}{2}\right)\boldsymbol{J_0\,(\beta)\,cos(\omega_0 t)} + \boldsymbol{E_0\,cos}\left(\tfrac{\emptyset_0}{2}\right) \\
&\quad \times\left\{\sum_{n=1}^{\infty}\boldsymbol{J_{2n}\,(\beta)\,cos(\omega_0 t - 2n\omega_e t + n\pi)} + \boldsymbol{cos(\omega_0 t + 2n\omega_e} \right. \\
&\quad \left. \boldsymbol{- n\pi)}\right\} - \boldsymbol{E_0\,sin}\left(\tfrac{\emptyset_0}{2}\right)\times \\
&\left\{\sum_{n=1}^{\infty}\boldsymbol{J_{2n-1}(\beta)\,c}\begin{bmatrix} sin(\omega_0 t - (2n-1)\omega_0 t + n\pi - \frac{\pi}{2}) \\ -sin(\omega_0 t + (2n-1)\omega_0 t - n\pi + \frac{\pi}{2}) \end{bmatrix}\right\}
\end{aligned}
\tag{4}
$$

Equation (4) indicates that input power of the carrier will be spread out among first order, second order and third order higher order optical sidebands. The amplitude distribution of these sidebands is directed by the variation of Bessel functions parameterized by $\boldsymbol{\beta}$ and values of $\emptyset$ will affect their amplitude. If all these optical sidebands are fed to PD, harmonics of the electrical drive signal will be generated and can be obtained higher order electrical harmonic like second order and third order sideband. Frequency doubled electrical signal can be considered as first order sideband is then achieved at PD which has more efficiency than other electrical signal orders like second order sideband which can be considered as second order. The proposed method in this paper, the dc bias of the MZM is tuned to have 0 which lead to remove all of other order sidebands in order to improve the efficiency of obtained signal. The optical signal now can be approximately expressed as

$$
\begin{aligned}
&\mathrm{E_{out}\,MZM\,(t)} \cong \\
&\mathrm{E_0J_0\,(\beta)\,\cos(\omega_0 t) - E_0J_2\,(\beta)\,\cos(\omega_0 t - 2\omega_0 t) - E_0J_2\,(\beta)\,\cos(\omega_0 t \mp 2\omega_0 t)}
\end{aligned}
\tag{5}
$$

However when this generated signal is applied to a PD, a strong frequency which is double electrical signal (first order sideband) is generated and weaker frequency quadrupled electrical signal (second order sideband) will be achieved. Then these generated optical signal that is composed of two sidebands and optical carrier is transmitted over single fibre which will suffer from chromatic dispersion induced the high power penalty. In addition for that the presences of first order will cause interference to the process of the second sideband electrical signal in system application. In this proposed method inverted optical filter (Notch filter) is used to eliminate the first order sideband and remove the carrier (see Fig. 2). The electrical field of optical signal at the output of inverted optical filter (notch filter) will be expressed as

$$
\mathrm{E_{out}\,(t) \cong -E_0J_0\,(\beta)\,[\cos(\omega_0 t - 2\omega_0 t) + \cos(\omega_0 t + 2\omega_0 t)]} \tag{6}
$$

When this obtained optical sidebands fed to PD, a signal that has four timed the frequency of the electrical drive signal will be generated (see Fig. 3), which can be written as

fridaya@uef.fi

$$V_{out}(t) = CJ^2(\beta)\cos(4\omega_e t) \tag{7}$$

Where C is a constant that is related to the responsivity of the PD. Since the two optical sidebands originate from the same optical source.

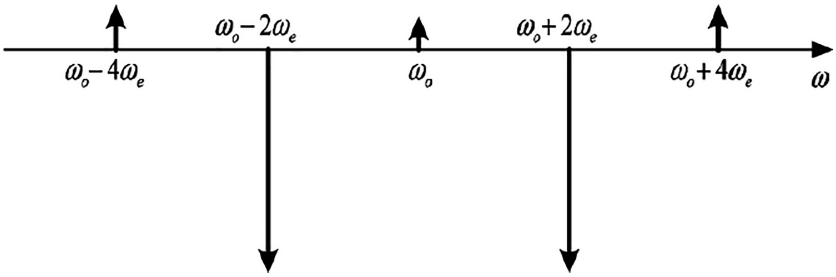


Fig. 2. Optical spectrum at the output of inverted optical filter

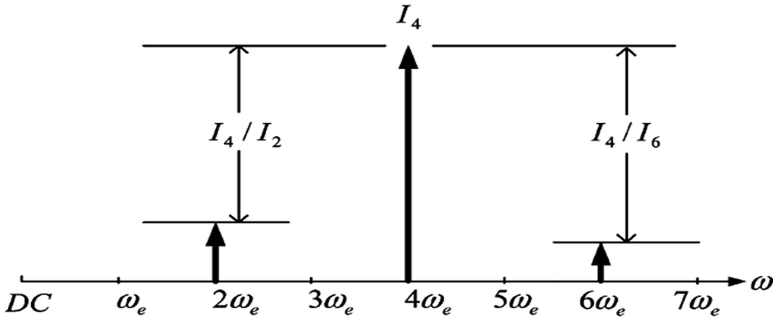


Fig. 3. Electrical spectrum at the output of a PD

3 Simulation and Results

The proposed design is simulated using OPTISYSTEM software to generate an optical carrier suppressed mm-wave signal. Further, the performance of data transmission of generated millimeter wave (mm-wave) over fiber using data modulation schemes are also analyzed.

3.1 Generating a Mm-Wave

For the proposed scheme regarding the generation of a 60 GHz mm-wave signal with carrier suppression (see Fig. 4). The entire system is simulated using OPTISYSTEM software (see Table 1).

fridaya@uef.fi

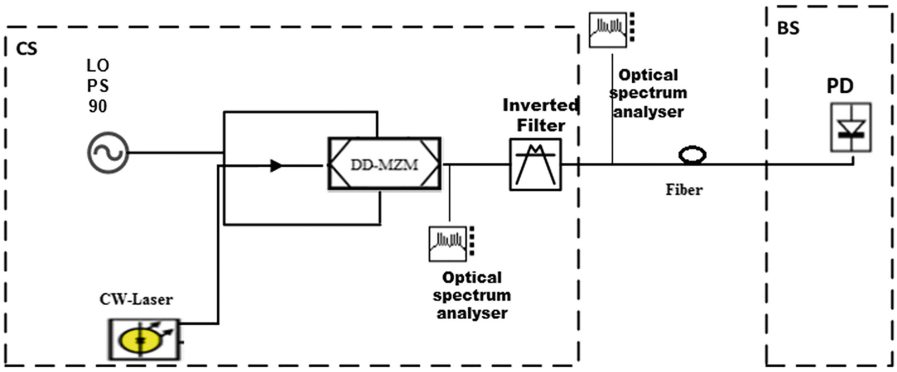


Fig. 4. The proposed design to generate mm-wave of OCS 60 GHz

Table 1. Simulation parameter setup

Device	Parameter	Value
(CW) Laser	Optical frequency	193.1 THz
	Optical Power	10 dBm
	Linewidth	10 MHz
LocalOscillator (LO)	*Radio frequency (RF)*	*15 GHz*
Fiber	*Fiber lengths*	(0–110) km
DD-MZM	*Bias voltage*	0 V
	Extinction Ration	60 dBm
	Switching voltage	4 V
PD	*Responsivity*	1 A/W

The obtained optical spectrum at the output of the DD-MZM is seen in Fig. 5(a) which is measured by optical spectrum analyzer. The obtained optical spectrum composed of the carrier, first order sideband and second order sideband whilst that of the optical filters (see Fig. 5b). The optical spectrum that obtained at the output of inverted optical filter (notch filter) composed of first and second order sidebands. Thus input power of the carrier will be spread out among first order, second order and third order higher order optical sidebands. It clearly shows that the optical spurious suppression ratio (OSSR) is 30.56 dB has been achieved which lead to avoid using electrical filter at the receiver. At the PD, when the system is used to generate mm-wave without using inverted optical filter; mm-waves are obtained, namely, 30 GHz which is double the input frequency, 60 GHz which quadruple the input frequency (see Fig. 6a) that is measured by Radio frequency analyser which is connected at the output of PD. The obtained mm-waves at the PD when the system is using inverted optical filter (see Fig. 6b). It shows that double the input frequency (30 GHz) which is first order sideband is removed and the power level of quadruple the input frequency (60 GHz) is increased.in addition for that, the 90 GHz which sextuple the input frequency is generated but with RFSSR of 25 dB.

fridaya@uef.fi

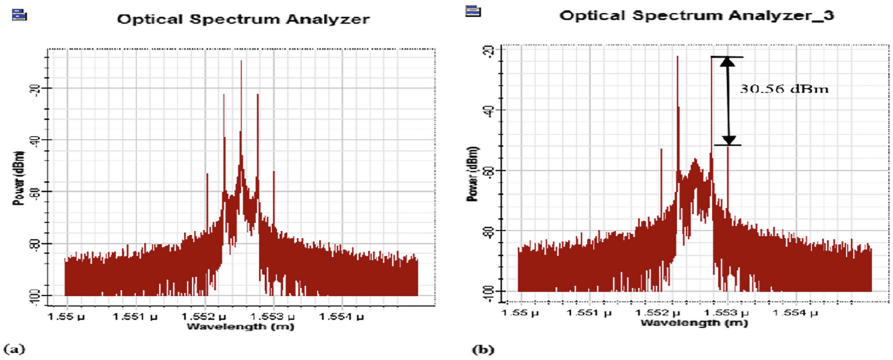


Fig. 5. The obtained optical spectrum (a) At the output of DD-MZM (b) At the output of inverted optical filter.

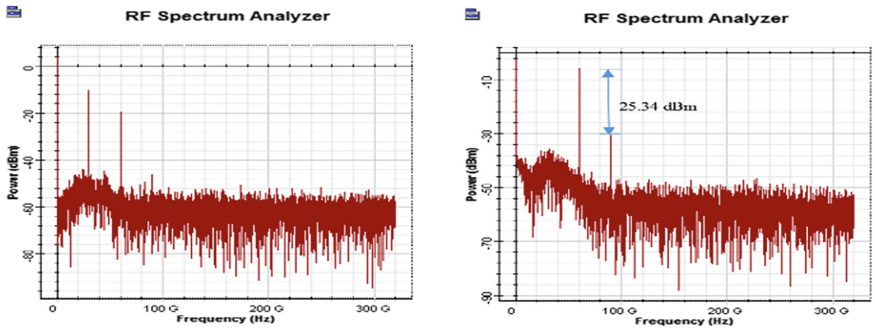


Fig. 6. The obtained mm-wave signals at the PD (a) when the carrier is not suppressed; (b) when the carrier is suppressed.

While Fig. 4b indicates that the RFSSR is increased to 25.34 dB. Thus, removing the carrier by using inverted optical filter adds the benefits of enhancing the RFSSR of the quadruple the input signal. With proposed design, the power of desired mm-wave signal at 60 GHz is −1.600 dBm, which is clearly greater than that of other generated frequencies who they are nearly to −33 dBm.

3.2 The Data Transmission

The transmission performance of the obtained mm-wave signal over single mode fiber (SMF) is evaluated by modulating the obtained optical signal with baseband signal of 2.5 Gbps. Firstly, the performance of data transmission is evaluated when the system is not using the proposed inverted optical filter which is refer to design A (see Fig. 7). Then Two configurations are considered to modulate the base band signal the data in the proposed design and comparing the performance of the modulated mm-wave in

fridaya@uef.fi

each. First configuration when the data is modulated directly with DD-MZM while generating optical mm-wave signal which refer to Design B (see Fig. 8), second configuration when the data is modulated with obtained optical mm-wave signal which, namely, design C (see Fig. 9).

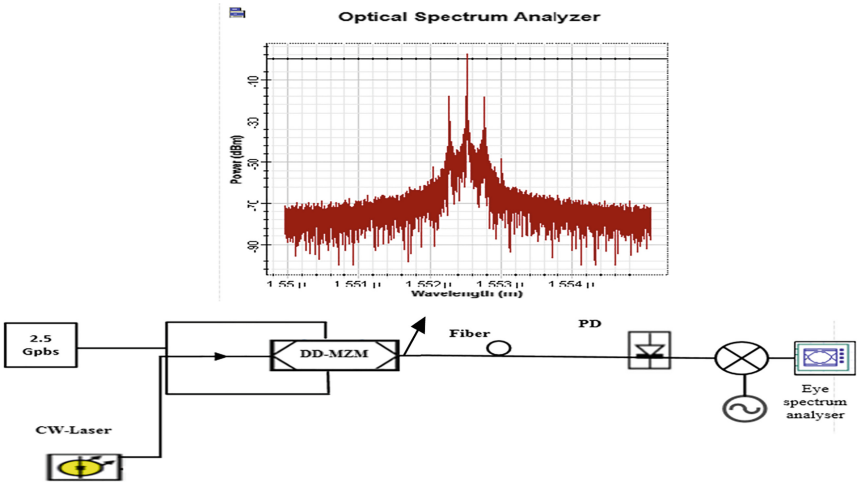


Fig. 7. Simulation setup of Design A

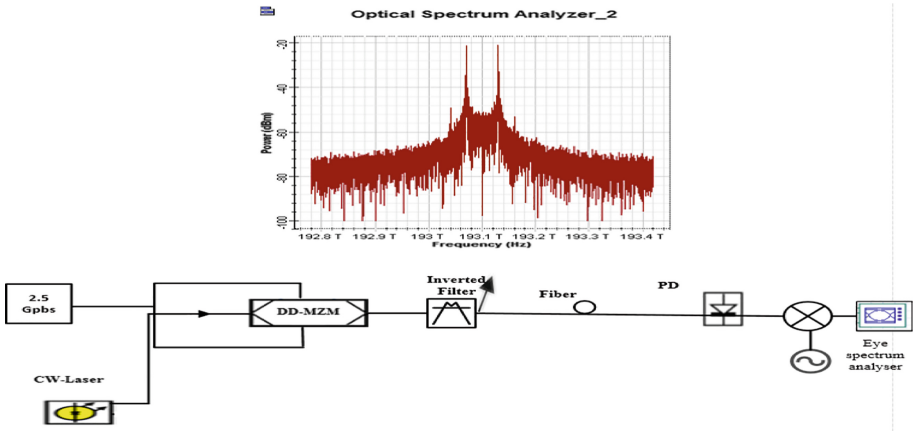


Fig. 8. Simulation setup of Design B, where the inverted optical filter is used

fridaya@uef.fi

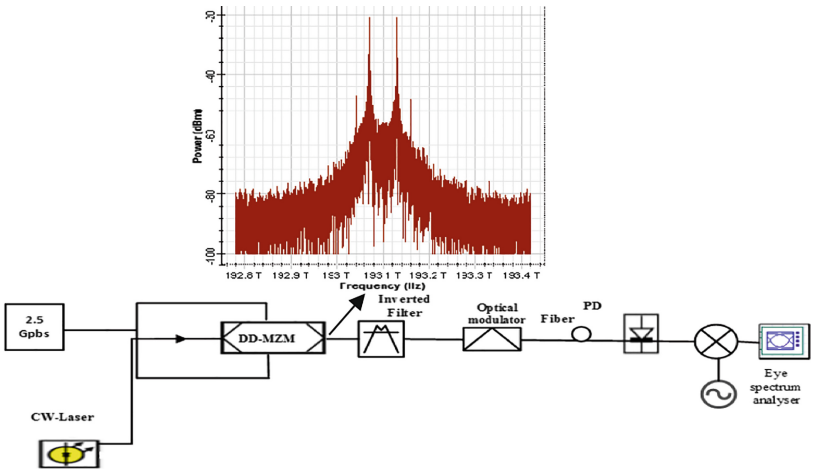


Fig. 9. Simulation setup of Design C

All of designs are simulated using OPTISYSTEM over Single mode Fiber transmission line in different transmission distances stating Back to Back system (B to B), 2 km, 20 km and 40 km. The SMF attenuation is set to 0.2 dB/km and 16.75 ps/nm/km dispersion. At the receiving part, PD is used to detect the modulated mm-wave signal at 60 GHz. Then, the data are recovered to evaluate the performance of the mm-wave by observing the eye diagram analyser for each design. An Eye diagram analyser is used in OPTISYSTEM to detect the obtained modulated mm-wave signal by monitoring the opening eye diagrams, which indicate the quality of the obtained signals (see Figs. 10, 11, and 12) and the Bit Error Rates.

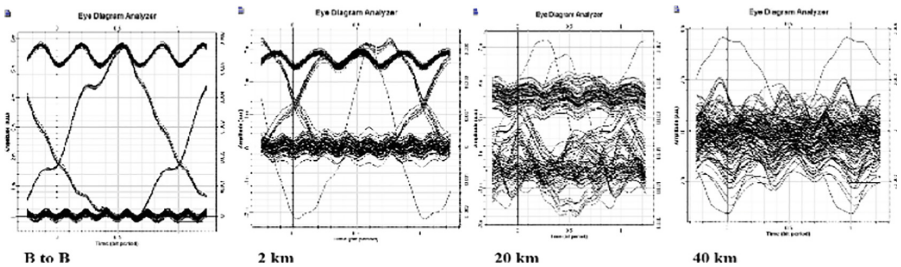


Fig. 10. Eye diagrams for Design A

fridaya@uef.fi

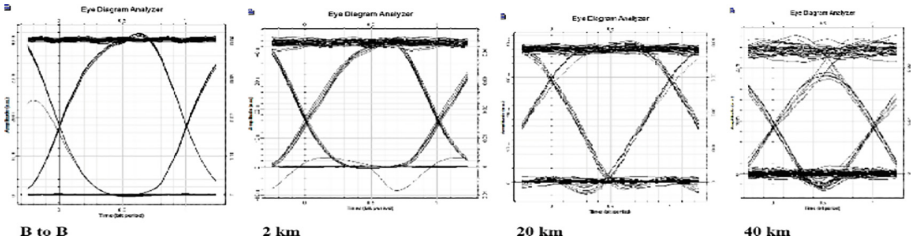


Fig. 11. Eye diagrams for Design B

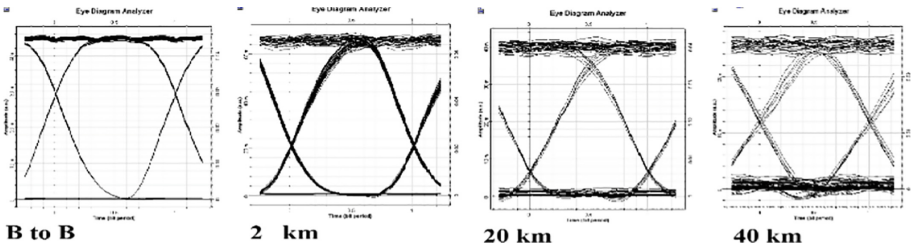


Fig. 12. Eye diagrams for Design C

Figure 10, 11, and 12 illustrates that the eye diagrams that are obtained for designs A, B, and C, respectively. Figure 10 shows that the obtained signal from design A has a large performance degradation where the carrier is not supressed. Figure 11 shows that the eye diagram of the received signal at 2 km is perfectly clear and widely opened. When transmission distance is increased eye diagrams started to be distorted due to the optical fiber impairments such as; attenuation and chromatic dispersion. On other hand the eye diagrams that are obtained from design C is perfectly clear and widely open with less distortion than that in design B which means has more capable to overcome the effect of attenuation and chromatic dispersion. Eye diagrams obtained demonstrated that removing the carrier add benefits to transmit the data with obtained mm-wave more effectively. Figure 13 displays the Max Q factor at different fibre distances, demonstrating that max Q factor decreases as the distance increases for all the designs. However, Design B and Design C have a better Max Q factor than Design A as the fibre length increases. For Back to Back fibre distance the Max Q factors for Design A, Design B and Design C are 22.42, 85 and 44.5, respectively. Moreover for 20 km fibre length the Max Q factors for designs A, B and C are 7.39, 13.085 and 17.4, respectively. Further, at 40 km the max Q factor is very low for Design A at 4.42, whereas Design B and Design C have Max Q factors of 10.45 and 21.42 at this distance. Hence, the results obtained in Fig. 11 demonstrate that the designs when the carrier is suppressed have better performance than when it is not. In addition for that, using the inverted optical filter to remove the carrier can generate an effective mm-wave.

fridaya@uef.fi

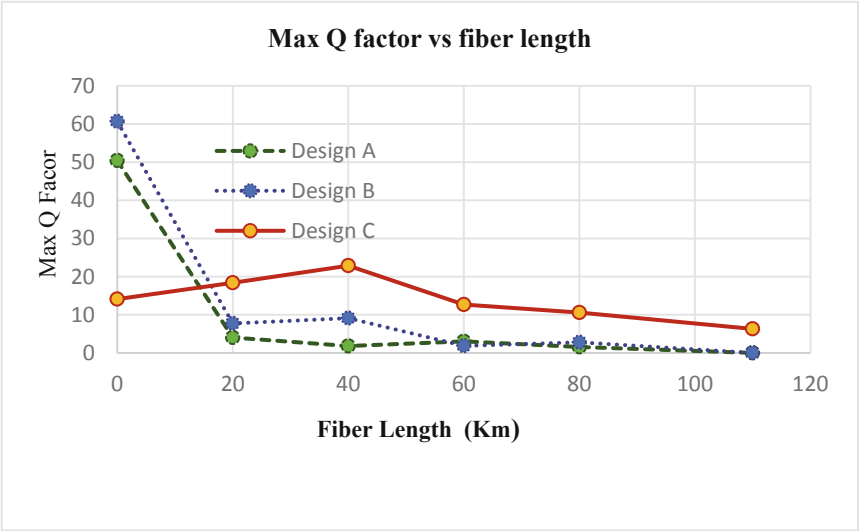


Fig. 13. Max Q factor for different designs at different fiber length distances

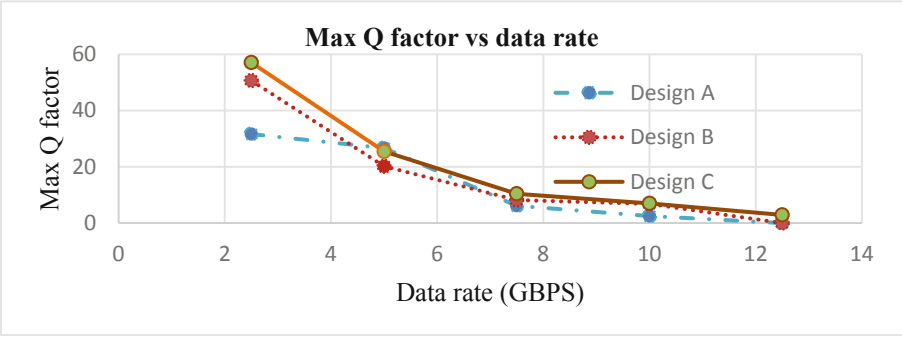


Fig. 14. Max Q factor with different bit rates.

Moreover, Fig. 14 shows that the proposed designs B and C can work effectively until 10 Gbps over 2 km with same parameter setup. As shown in Fig. 14, the Max Q factor decreases when there are increasing in data rates. As illustrated in the figure, at 10 Gbps the Max Q factor for Design A is 2 while those Design B and Design C are 7.8 and 10, respectively.

4 Conclusions

In this paper, it has been approved that mm-wave signal can be generated effectively with carrier suppression using inverted optical filter (notch filter) in a simple low cost design. Using an inverted optical filter adds the advantage of suppressing the carrier,

fridaya@uef.fi

which leads to balancing of the received signal power of the desired obtained signal. Moreover, what is proposed reduces the number of optical components, thereby making it easy to implement. The OPTSYSTEM simulation has demonstrated that the proposed design has good transmission performance in both scenarios of implementing the data; when the data is modulating while generating optical mm-wave and when the data is modulation after the optical mm-wave is generated. However, the results show that modulating the data with the obtained mm-wave has better performance. For 2.5 Gbps data transmission, the max Q factor for 40 km are 10.45 and 21.42 for at this distance. Moreover, the Max Q factor of the obtained signal for 20 km SMF length is 17.4 and that for 40 km is 8.42. With the scenario of modulating the data directly to MZM, the Max Q factors are 13.083 for 20 km and 7.45 for 40 km. In addition to this, the performance of the obtained mm-wave signal is effectively until 10 Gpbs for 2 km in both scenarios of modulating the data. It is concluded that carrier suppression using inverted optical filter can be considered to generate effective mm-wave.

References

1. Xiaoqiong, Q., Jiaming, L., Xiaoping, Z., Liang, X.: Fiber dispersion and nonlinearity influences on transmissions of AM and FM data modulation signals in radio-over-fiber system. IEEE J. Quant. Electron. **46**(8), 1170–1177 (2010)
2. Davinder, S., Preeti, S.: Ttechniques of millimetre-wave signal generation in ROF systems: A review. Int. J. Comput. App. Inf. Technol. **I**(II), 2278–7720 (2012)
3. Guohua, Q., Jianping, Y., Joe, S., Stéphane, P., Claude, B.: Generation and distribution of a wideband continuously tunable millimeter-wave signal with an optical external modulation technique. IEEE Trans. Microw. Theory Tech. **53**(10), 3090–3097 (2005)
4. Beas, J., Castanon, G., Aldaya, I., Aragón-Zavala, A., Campuzano, G.: Millimeter-wave frequency radio over fiber systems: a survey. IEEE Commun. Surv. Tutor. **15**(4), 1593–1619 (2013)
5. Fang, Z., Ye, Q., Liu, F.: Progress of millimetre wave subcarrier optical fibre communication technologies. Chinese J. Lasers **33**, 481–488 (2006)
6. Jianjun, Y., Zhensheng, J., Lilin, Y., Yikai, S., Gee-Kung, C., Ting, W.: Optical millimeter-wave generation or up-conversion using external modulators. IEEE Photonics Technol. Lett. **18**(1), 265–267 (2006)
7. Jian, Z., Hongwei, C., Minghua, C., Tianliang, W., Shizhong, X.: A photonic microwave frequency quadrupler using two cascaded intensity modulators with repetitious optical carrier suppression. IEEE Photonics Technol. Lett. **19**(14), 1057–1059 (2007)
8. Mohmoud, M., Xiupu, Z., Bouchaib, H., Ke, W.: Frequency sextupling for millimeter-wave over fiber systems. OSA Opt. Express **15**, 10141–10151 (2008)
9. Peiming, S., Song, Y., Zekunl, I., Jien, S., Yaojun, Q., Wanyi, G.: A frequency sextupling scheme for high-quality optical millimetre-wave signal generation without optical filter. Opt. Fiber Technol. **17**(3), 236–241 (2011)
10. Qin, Y., Junqiang, S.: Frequency sextupling technique using two cascaded dual-electrode Mach-Zehnder modulators interleaved with Gaussian optical band-pass filter. Opt. Commun. **285**, 2911–2916 (2012)
11. Lei, S., Aijun, W., Bo, L., Tonggang, W.: A filterless optical millimeter-wave generation based on frequency octupling. Optik **123**(13), 1183–1186 (2012)

fridaya@uef.fi

12. Jing, L., Aijun, W., Huixing, Z., Xiaome, H.: A novel MI-insensitive and filterless frequency octupling scheme based on two parallel dual-parallel Mach-Zehnder modulators. Optik **125**, 6887–6890 (2014)
13. Muthu, K., Raja, A., Sevendran, S.: Optical generation of millimeter waves through frequency decupling using DP-MZM with RoF transmission. Opt. Quant. Electron. 51, 1572–817X (2017)
14. Qi, G., Yao, J., Seregelyi, J., Bélisle, C.: Generation and distribution of a wide-band continuously tunable millimeter-wave signal with an optical external modulation technique. IEEE Trans. Microw. Theory Tech. **53**(10), 3090–3097 (2005)

fridaya@uef.fi

Check for updates

Visualization: A Tool to Study Nodes in Multi-dimensional Networks

Fozia Noor[1,2(✉)], Muhammad Usman Akram[3], Asadullah Shah[2], and Shoab Ahmad Khan[3]

[1] Yanbu University College, Yanbu, Saudi Arabia
noorf@rcyci.edu.sa, fozia_noor_khan@yahoo.com
[2] International Islamic University Malaysia (IIUM), Gombak, Malaysia
[3] National University of Sciences and Technology (NUST), Islamabad, Pakistan

Abstract. The world of data has become progressively complex with the explosive growth of the web and social networks. The use of complicated multi-dimensional models for such data has made visualization and analysis of data challenging. The multiple dimensions of data contain valuable information, but it is hard to analyze as the information is scattered among the dimensions. Sophisticated appropriate visualization of such networks can help in extracting the network and node characteristics to improve decision making. This papers study different multi-dimensional dataset using visualization tool to explore the different ways that can be used to study node characteristics.

Keywords: Multi-dimensional networks · Multilayer visualization · Mono-plex visualization

1 Introduction

Data visualization has become the de facto standard in SNA and is mainly used to quickly interpret the data and adjust different variables to see their effects or to monitor and ensure the performance of models under study. Effective visualization of networks offers a unique way to explore and analyze complex systems [1]. Different ways can be used to model the network. The common representations are adjacency matrix, adjacency lists, and node-edge diagrams. The adjacency matrix is used to represent a single graph indicating whether pairs of nodes are adjacent or not in the graph representing the relationship among them [2]. Visualization using node-edge diagram has a rich history and has been used as an analytical tool [4]. Many stand-alone tools have been developed to model node-edge diagrams of numerous different domains i.e. Gephi [5], Pajek, LaNet-vi, GraphViz, etc. These tools cover only mono-plex network paradigms.

The multi-dimensional network includes multiple dimensions representing different kinds of relations [3]. Multi-dimensional representation of real-world systems yields valuable insight into data during analysis. Simple adjacency matrix or node-edge diagrams cannot represent the Multi-dimensional network. Multi-dimensional network includes not only intra-edges but also inter-edges. Henceforth, instead of the simple adjacency matrix, supra-adjacency matrices are used to represent them. Similarly, basic node-edge visualization tools generate different unrelated graphs for each dimensions

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 358–367, 2019.
https://doi.org/10.1007/978-3-030-22868-2_27

fridaya@uef.fi

or represent the multi-dimensions of multi-layer networks in single graph connecting nodes with all possible relations using different colors. In both types of graphs, it is quite challenging to find the similarities in structure or compare the network characteristics. Figure 1 shows an adjacency matrix, node edge diagrams and supra adjacency matrix for Krackhardt high tech network (multi-dimensional network with 3 dimensions) [6] as an example.

In this paper, we have explored muxViz [7] and have generated multi-layer graphs to study the nodes from different perspectives. Multiple different datasets i.e. Noordin terrorist network [8] with 10 dimensions, Lazega law firm [9] with 3 dimensions and Krackhardt networks with 3 dimensions are used to model and analyze the nodes in multi-layer paradigm.

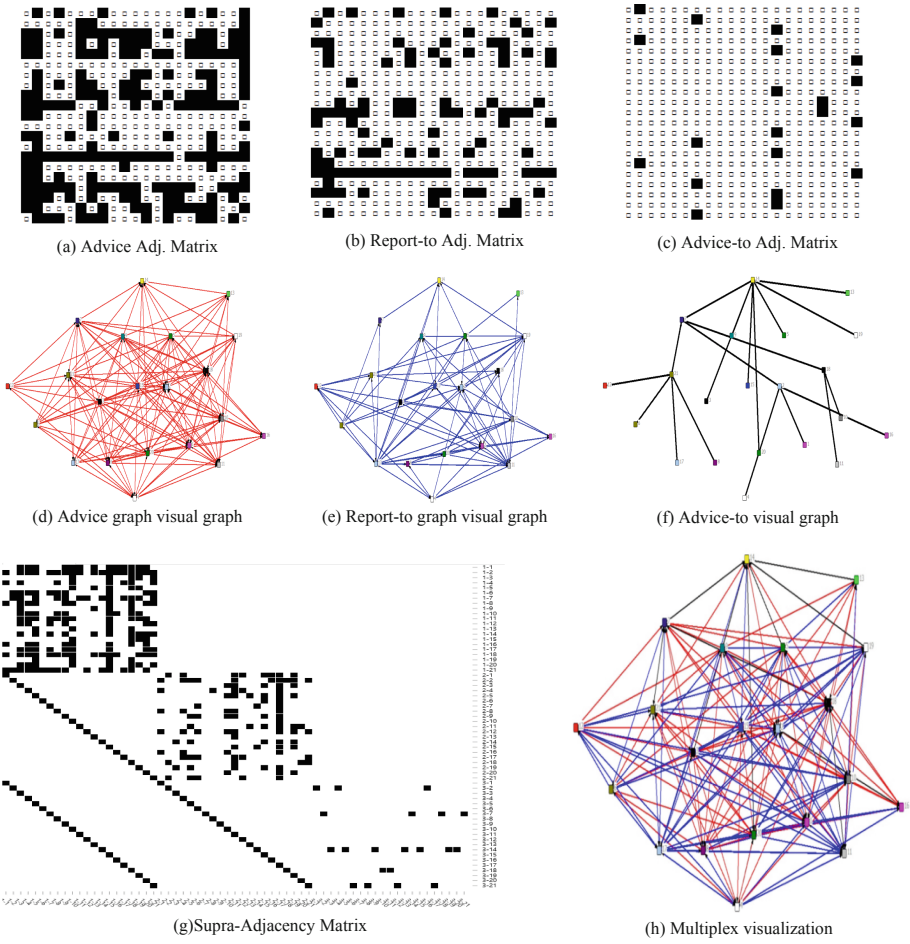


Fig. 1. (a,b,c) Adjacency matrix of three-layer using socNetV-2.1 (g) supra adjacency matrix using socNetV-2. (d) Supra adjacency matrix plot for CASE III-1 using MuxViz (d,e,f) Monoplex visualization of three layer using Gephi (h) Multiplex visualization using UCINET

fridaya@uef.fi

The rest of the paper unfolds as follows. The next section discusses the modelling of multi-dimensional data in a multi-layer paradigm while Sect. 3 discusses the multi-layer visualization with different perspectives. The last section summarizes the work.

2 Modeling of the Multi-dimension Dataset as Multi-layer

All the datasets used are modeled as a multi-layer network. The multi-layer model has L layers, nodes v represented in set V and edges e in set E, where each dimension is modeled as a single layer $l \in L = \{1 \ldots k\}$. Edges between nodes can be represented as $(u, l_\alpha), (v, l_\beta) \in E$ where the nodes u and v belong to the subset V and α and β are two layers of multi-dimensional set $(l_\alpha, l_\beta \in L)$. The edges between nodes will be intra-layer if α is equal to β or will be inter-layer if α is not equal β. All the cases covered in the current research have a categorical coupling (connectivity between replica nodes exists in all layers) to develop a fully interconnected node paradigm as shown in Fig. 2. But for the sake of simplicity, the inter-layer edges will not be shown in the succeeding developed visualizations but should be assumed to be existent in the network.

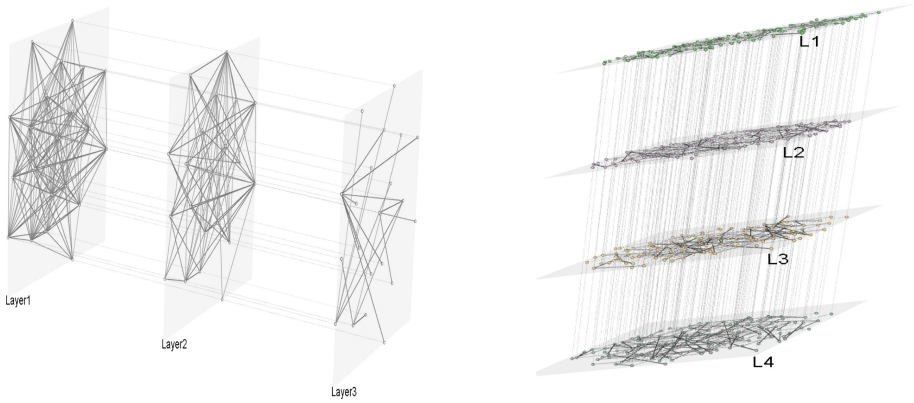


Fig. 2. Node aligned network with categorical coupling

3 Multi-layer Visualization

The multi-layer visualization and analysis of the network to study nodes characteristics and behavior can be carried out from different perspectives. A number of graphs are generated using different visualization methods, layouts, node measures, and edge measures to highlight network structure and node behavior in/across the network. In our current paper we have focused our point on studying nodes; therefore it discusses different perspectives used to analyses

fridaya@uef.fi

1. Studying nodes using centrality measures and characteristics as main descriptors of the nodes in monoplex network visualization.
2. Studying node behavior across the layers.
3. The multidimensional model can also be analyzed by using aggregated and multilayered co-analysis approach. Compare all the approaches to analyze multidimensional networks.
4. To analyze the layers similarities based on the nodes.

3.1 Studying Nodes in Monoplex Network Visualization

The visual representation of layers as a mono-plex network can be carried out with a goal to reveal the insight of structure and to represent nodes' characteristics at each dimension independently. Different layers in a model can depict different structural information. This structural information can be hidden if the layout used in visualization forces layers to be organized similarly to another layer. It is possible that two nodes connected on one layer are disconnected on another, making different intra-layer links of layers. Consequently, the independent layout should be used for each layer to study each layer independently.

The visual graphs enriched with analytical measures results in improved structural visualization of networks. The range of different colors and sizes of nodes represent the

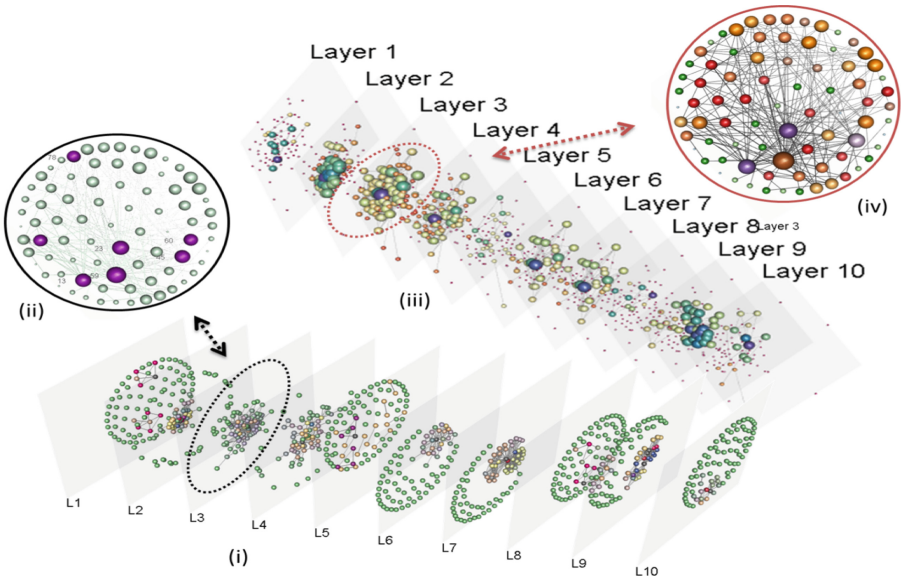
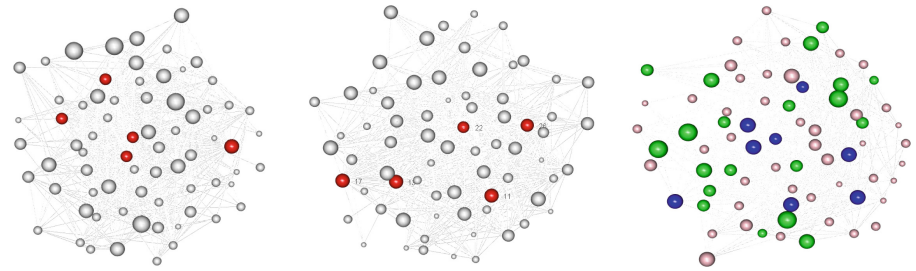


Fig. 3. Multilayer sliced visualization of Noordin Top Network (i) sliced independent visualization of each layer representing centrality of each node (ii) dual color coded Layer 1 visualization (purple = 7 top centrality individual, Light green = rest of member) the size of nodes = centrality value (iii) 2D visualization (nodes color = degree centrality, node size = Eigen centrality) (iv) nodes color 10 bins depending on centrality value and size of node is Log of centrality.

fridaya@uef.fi

value of the range of centrality measure. Each node can also represent two of its centrality measure values by using color to represent one and size to represent the other (shown in Fig. 3). Personal measures also play a significant role in unfolding any behavior of the node. Personal measures can also be merged with the structural measure to enrich visualization with advanced information about nodes. The mix and match of these attributes and matrix visualization can reveal the dependencies of the attributes on each other.

Figure 4 represents the Lazega law firm, where the size of the nodes represents the age of the employee using five bins. It reveals the interesting fact that the top key nodes using PageRank and Eigenvector centrality measure are not the most senior or aged employee, nor are they the youngest. In typical scenarios, mostly the employee has seniority according to their age, but in some cases, as seen in Fig. 4(c), the top-ranked nodes are the ones from middle age or middle seniority.



(a)Multi-layer EigenVector(size= age) (b) Multi-layer Pagerank(size = seniority) (c) Visualization using Size = age in 5 bins and colors = seniority in 3 bins

Fig. 4. Multilayer Visualization of Lazega law firm representing personal and structural attributes

3.2 Analyzing the Node Behavior Across Layers

Another basic exploration in analysis can be an investigation of node behavior change through different layers in a mono-plex sliced network architecture. If we have multiple relationships between the nodes, then change in nodes' behavior over the layers can be explored visually by using a colored representation of node for specific measure or property under study. Visualization using aggregated mode keeps each node position consistent throughout all layers to highlight similar pattern in all layers. Each node's consistent position in all layers facilitates analyzing the change in node behavior in terms of their presence, community, centrality measure, personal attributes, etc. However, it can hide the structures existing in any one specific layer, as it forces the layout suitable for the aggregated layer.

Figure 5 represents the degree centrality of nodes in Noordin Terrorist Network across layers using colors as a descriptor of degree centrality range, keeping the size of nodes constant. The behavior or role of any node cannot be consistent in all layers, as shown in Fig. 5, for a specific highlighted node 51. The node 51 has the highest degree of centrality in layer 4 and is one of the top centrality nodes of the layer, while it is least connected in layer 1. Hence, it is not necessary that any node plays a major role in all dimensions.

fridaya@uef.fi

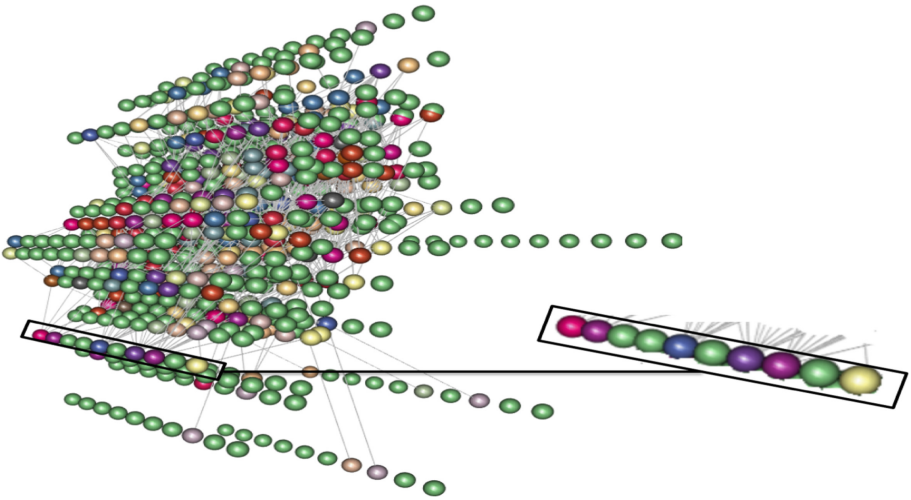


Fig. 5. Multilayer Visualization of Noordin Terrorist Network

Change of nodes behavior has also been explored by plotting node measures across many dimensions, where each dimension value corresponds to a node measure at a specific layer in a model. Moustafa [10] used parallel coordinates as a tool for visualizing multidimensional data and highlighted that the same node could show different behavior depending on the relation type on different layers. The role of different layers can be different depending on the perspective or measure used for analysis. The fact that a node is among the top-ranked nodes for degree centrality does not necessarily guarantee that it will also be top ranked for another centrality measure in the same layer.

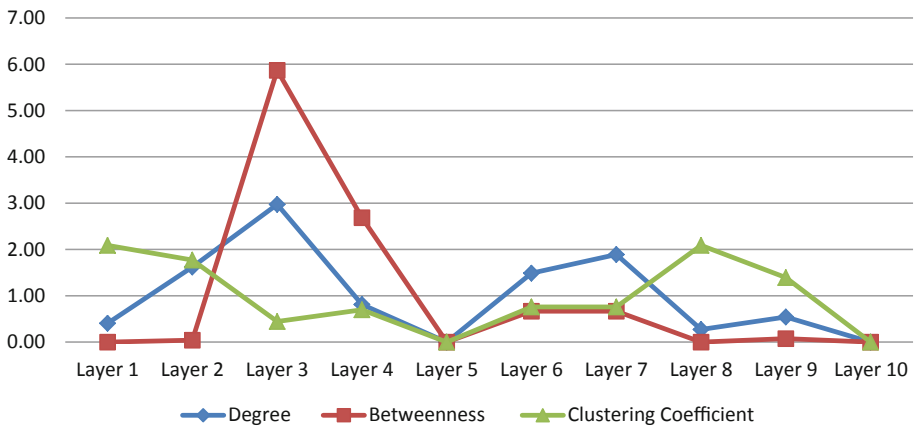


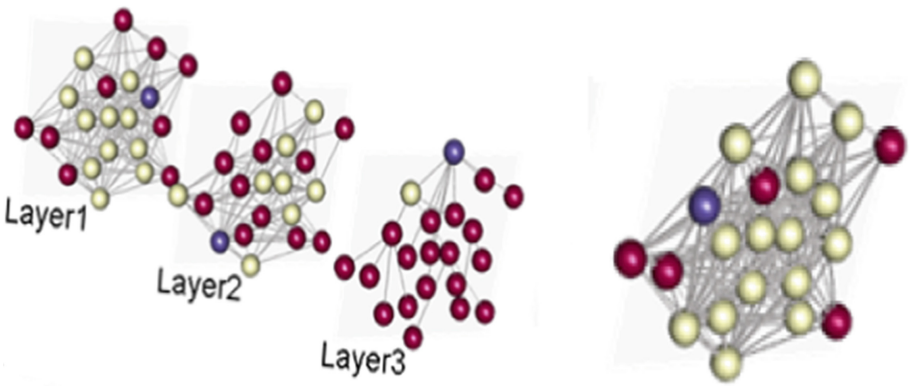
Fig. 6. Layer-specific centrality measures value for node 23 in Noordin Terrorist Network

fridaya@uef.fi

As shown in Fig. 6, the node 23 in Noordin Terrorist Network shows the highest number of connections and betweenness in layer three, but the triadic behaviors of friends are quite low comparatively. Similarly, on the contrary reason, the node 23 behavior in Layer 1 and Layer 9 is the inverse of behavior shown in layer 3 in terms of clustering ability. The node 23 shows the high tendency of clustering in layer 1 and layer 8. Hence, all layers may participate in the nodes' life from a different view, and it may vary. Each perspective should be considered in analysis and decision making.

3.3 Comparing the Three Network Architecture Approaches

The study of multiple layers' network modeled in a multilayer paradigm can be also done using aggregation and multilayer co-analysis approach along with the monoplex approach. The approach used for analysis effects the network and nodes characteristics directly. Figure 7 shows the PageRank centrality calculated using mono-plex networks and multi-layer co-analysis for Krackhardt network. From the color representation, it can be seen that the majority of nodes changes their color bin in mono-plex networks and multi-layer co-analysis. This means the ranking of nodes in the mono-plex networks and multilayered co-analysis approach changes. Furthermore, the node N19 in Krackhardt network has highest PageRank centrality in multi-layer co-analysis while it belongs to third (least PageRank bin) in two of the monoplex networks.



(a) using mono-plexed networks (b) using multilayer co-analysis

Fig. 7. Krackhardt network multilayer network

Furthermore, the Lazega law firm network is shown in Fig. 8 using sliced monoplex and flattened network approach, keeping nodes' position constant. In each approach, the graph is representing the nodes with dual color, representing the top five nodes in red color. If the top five ranked nodes are checked in both, we can see that only two nodes are common in all rankings but not at the same ranking position. The node N24 is among the top nodes ranked at second highest value, but becomes become

fridaya@uef.fi

less important in the sliced networks because its connections become distributed on different layers.

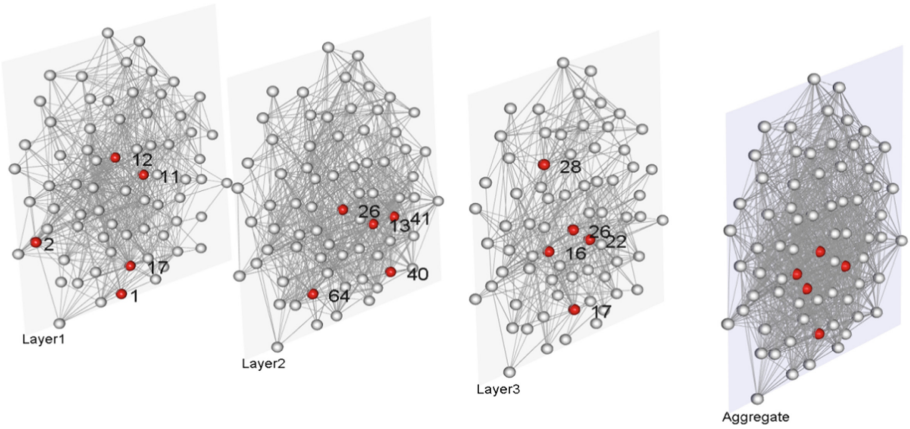


Fig. 8. Top 5 nodes based on degree centrality in Lazega law firm network

These examples indicate that calculating the nodes' centrality in each sliced network of the structure, in the single aggregated network or multilayer co-analysis may lead to a bit different results from each other. Each approach has its effects in the analysis, and choice of correct approach will affect the efficiency of the study as the characteristics of nodes vary greatly.

3.4 Analyzing the Layer Similarities Based on Nodes Across the Layers

The visualization is not only used to quantify the participation of nodes in the network structure but can also find the layers similarities based on nodes' behavior/participation. How many layers are contributing equally toward the multiplex structure can be roughly represented by a total number of nodes contributing to each layer. Additionally, if some layers can represent major or maximum node and edge pattern of another layer, then those layers can be considered to be more contributing significant layers, and can thus be given higher importance in the structure. The layer with the minimum contribution may get the least importance or weight; similarly, it can be neglected if only primary contributing layers should contribute.

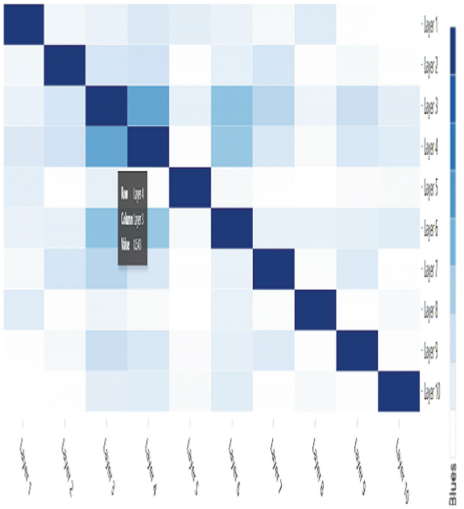
To explore the insight of layer significance and relationships in a network, the distributions of the number of active nodes and activity correlations between pairs of layers can be used. It is not compulsory that every node contributes to each layer. If the number of nodes' contribution varies in the layer, the structural properties of layers will also differ. Furthermore, if the two layers have the same 50% of node contribution, it is still possible that the nodes' contribution differs in both layers and two-layer differ in their structure, mostly due to different edge patterns. The inter-layer correlation using node similarity and edge similarity can capture this structural similarity of layers in multiplex networks.

fridaya@uef.fi

Furthermore, if the number of the node and the contributing nodes are different, the node and edge overlap percentage will also differ in different layers. Figure 9 shows the Noordin dataset inter-layer correlation using node overlapping and edge overlapping.

	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7	Layer 8	Layer 9	Layer 10
Layer 1	1.00	0.10	0.16	0.15	0.05	0.10	0.11	0.05	0.09	0.04
Layer 2	0.10	1.00	0.48	0.46	0.14	0.22	0.37	0.08	0.22	0.10
Layer 3	0.16	0.48	1.00	0.77	0.30	0.33	0.47	0.11	0.46	0.20
Layer 4	0.15	0.46	0.77	1.00	0.23	0.30	0.42	0.11	0.37	0.19
Layer 5	0.05	0.14	0.30	0.23	1.00	0.09	0.13	0.04	0.14	0.09
Layer 6	0.10	0.22	0.33	0.30	0.09	1.00	0.15	0.10	0.20	0.14
Layer 7	0.11	0.37	0.47	0.42	0.13	0.15	1.00	0.08	0.25	0.05
Layer 8	0.05	0.08	0.11	0.11	0.04	0.10	0.08	1.00	0.09	0.04
Layer 9	0.09	0.22	0.46	0.37	0.14	0.20	0.25	0.09	1.00	0.06
Layer 10	0.04	0.10	0.20	0.19	0.09	0.14	0.05	0.04	0.06	1.00

(a)



(b)

Fig. 9. Layer Similarities (a) Interlayer correlation using node overlapping (b) Edge overlapping for Noordin Dataset

Node and edge overlapping shows that the majority of the layers are different in their nodes behavior; only Layer 3 and Layer 4 have maximum similarity in their structure in terms of a quantity of matching nodes' edges, followed by Layer 3 and Layer 6. All the layers of Noordin network are independent in their pattern and none of the layers can independently replace other. The activity of nodes across layers is varied, and all layers do not have the same level of activity in the overall system. For instance, a maximum number of the layers have no more than 20% (of the total) active nodes. However, only a few layers contain up to 75% active nodes. This means that all layers do not have the same level of importance and the removal of layers participating at the minimum should cause only a minor effect on the system.

4 Summary

This paper explores the multi-dimensional networks to study the nodes using visualization as a tool. The visualization of the network can be carried out using different perspectives and different ways. The model used to study the nodes i.e. monoplex, aggregated or multi-layer coanalysis effects directly on the behaviour of the node. It can be very beneficial in the cases where focus of study is to identify key players or to perform covert network analysis. The visualization of networks can reveal the interesting network and node characteristics and results in improved analysis and decision making.

fridaya@uef.fi

References

1. Rossi, L., Magnani, M.: Towards effective visual analytics on multiplex and multilayer networks. Chaos, Solitons Fractals **72**, 68–76 (2015)
2. Ghoniem, M., Fekete, J., et al.: A Comparison of readability of graphs using nodelink and matrix representations. In: IEEE Symposium on Information Visualization, Austin, TX (2004)
3. Berlingerio, M., Coscia, M., Giannotti, F., Monreale, A., Pedreschi, D.: Foundations of multidimensional network analysis. In: International Conference on Social Network Analysis and Mining (ASONAM), pp. 485–489. IEEE (2011)
4. Freeman, L.C.: Visualizing social networks. J. Soc. Struct. **1**(1), 1–15 (2000)
5. Bastian, M., Heymann, S., Jacomy, M.: Gephi: an open source software for exploring and manipulating networks. In: International AAAI Conference on Weblogs and Social Media (2009)
6. Krackhardt, D.: Cognitive, social structures. Soc. Netw. **9**, 104–134 (1987)
7. De Domenico, M., Porter, M.A., Arenas, A.: Multilayer analysis and visualization of networks. J. Complex Netw. **3**, 159–176 (2015)
8. International Crisis Group: Terrorism in Indonesia: Noordin's Networks (No. Asia Report #114). International Crisis Group, Brussels, Belgium (2006)
9. Lazega, E.: The Collegial Phenomenon: The Social Mechanisms of Cooperation Among Peers in a Corporate Law Partnership. Oxford University Press, Oxford (2001)
10. Moustafa, R.E.: Parallel coordinate and parallel coordinate density plots. Comput. Stat. **3**(2), 134–148 (2011)

fridaya@uef.fi

The Nigerian E-Waste Problem: Way Forward

Victor Ndako Adama[1(✉)], Ibrahim Shehi Shehu[1], Solomon Adelowo Adepojur[1], and A. Fatima Sulayman[2]

[1] Federal University of Technology, Minna, Nigeria
{vnadama, ibrahim.shehu, solo.adepoju}@futminna.edu.ng
[2] Ibrahim Badamasi Babangida University, Lapai, Minna, Nigeria
s.fatima@ibbu.edu.ng

Abstract. Technology has constantly evolved, leading to the invention and manufacturing of enhanced electronic and ICT devices. This has fueled the desire for people, companies and organizations to continuously upgrade their electronic and ICT infrastructure. Basically the circulation of these devices is on the increase. However, if new devices continuously replace existing ones, what happens to the ones replaced? Where do they go? How are they managed today? Unfortunately, these old devices end up in under developed and developing countries as "Tokumbos" (second hand, handed down), cheap alternatives. Such countries clearly lack the technology and technical know-how to appropriately recycle or dispose of them when they break down and become irreparable. They then end up as electronic waste (e-waste) not properly managed. Local disposal techniques employed by scavengers in such countries lead to consequences detrimental to the environment, the economy, ecology and lives of their people. This paper looks into the lingering problem of e-waste in Nigeria. It traces its source, reviews its negative impacts and via a comparatives study on countries that have good waste management practices, vis-à-vis efforts so far by the Nigerian government, seeks to suggest how to improve on the countries e-waste management strategies.

Keywords: Assessing · E-Waste · Problem · Way-forward

1 Introduction

Electronic waste also known as e-waste can be defined from different perspectives. The Basel Convention defined e-waste as "substance or objects, which are disposed of or are intended to be disposed of or are required to be disposed of by the provisions of national law" [1, 2]. The United Nations Statistics Division defined e-waste as "materials that are produced for the market for which the generator and the user has no further use in terms of its purposes of production, transformation or consumption, and of which they want to dispose" [3]. Thus, e-waste can simply be viewed as any electronic device that is either of the following; damaged beyond repair, whose life cycle has expired, has become obsolete, can still be used, but consequently rejected by the user. All definitions of e-waste have a key point in common, which is that that the device is no longer needed or useful to the owner. The Organization for Economic Cooperation and Development (OECD) categorize waste into three categories namely,

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 368–385, 2019.
https://doi.org/10.1007/978-3-030-22868-2_28

fridaya@uef.fi

nuclear or radioactive waste, toxic waste and municipal waste [4]. E-waste was categorized as toxic waste [4]. The European Union directive on Waste Electrical and Electronic Equipment (WEEE) and African E-Waste Solution Group based in South Africa presents a more detailed category of e-waste shown in Tables 1 and 2, respectively.

Table 1. E-waste categories pursuant to the EU Directive 2002/96//EC [5]

SN	Category	Label
1	Large household appliances	Large HH
2	Small household appliances	Small HH
3	IT and telecommunications equipment	ICT
4	Consumer equipment	CE
5	Lighting equipment	Lighting
6	Electrical and electronic tools (with the exception of large-scale stationary industrial tools)	E&E tools
7	Toys, leisure and sports equipment	Toys
8	Medical devices (with the exception of all implanted and infected products)	Medical devices
9	Monitoring & control instruments	M&C
10	Automatic dispensers	Dispensers

As at 2013, E-waste was reported as the world's leading waste stream in terms of growth. It was also revealed e-waste's growth index was triple that of overall municipal solid waste [7]. According to Professor Ming Wong in 2013, about 50 million tons (Mt) of hazardous e-waste of various electronics is generated each year [7]. According to Statista.com in 2018, the amounts of e-waste generated worldwide from 2010 to 2014 in million metric tons were 33.8, 35.8, 37.8, 39.8 and 41.8, respectively [8]. The same source projection for 2015 to 2018 was 43.8, 44.7, 47.8 and 49.8 million metric tons respectively [8]. This is attributed to the decrease in lifespan of the devices [9]. Table 3 shows different devices and their respective approximate life spans.

Problem at Hand. In Nigeria, most of the growth in the ICT sector is via the inflow of second hand electronic products and devices from developed countries. They seem like cheap alternatives, but when examined extensively are best classified as e-waste. Before reaching the shores of Nigeria, a large percentage of them are obsolete, expired or broken down beyond repair. These e-wastes have flooded Nigeria, making her emerge as one of the topmost electronic waste dumping grounds in the world [14] as at 2012. In 2018, an article by Kid Krunk on E-Terra Technologies placed Nigeria 3rd in the top 20 Countries that have become e-waste dumping grounds in the world [15]. As a large percentage of the devices are either broken, unrepairable or pack up in a while, they end up in dump sites, landfills and worse of all burnt openly under precarious conditions by scavengers for valuable components. Sadly, e-waste contains hazardous materials detrimental to the environment, the economy, ecology and lives of those

fridaya@uef.fi

Table 2. Categories of E-Waste [6].

Category	Examples
Large House hold Appliances	Washing machines, Dryers Refrigerators, Air conditioners, etc.
Small House hold Appliances	Vacuum cleaners, Coffee Machines, Irons, Toasters, etc.
Office, Information & Communication Equipment	Telephones, Fax Machines, Computers, Copiers, Printers etc.
Entertainment, Consumer, Electronics, Toys, Leisure, Sports Recreational Equipment, and Automatic Issuing Machines	Televisions, VCR/DVD/DC players, Hi-Fi sets, Radios, Electric train sets, coin slot machines, treadmills, Vending machines, parking ticket equipment etc.
Lighting Equipment	Fluorescent tubes and lamps, sodium lamps etc. (Except Incandescent Bulbs, Halogen Bulbs) etc.
Electric and Electronic Tools	Drills, electric saws, Sewing Machines, Lawn Mowers etc.
Security & health care equipment	Surveillance and Control Equipment (e.g. CCTV cameras, scanning equipment), and Medical Instruments and Equipment (e.g. x-ray and heart lung machines) etc.

staying around such areas, yet Nigeria is yet to effectively address this problem. A few hazardous and yet valuable metals contained in e-waste are listed in Table 4.

2 Reviewed Literatures

Nigeria, one of the most populous countries in Africa, is home to one-sixth of African's population. The increase in population translates to an increase in demand for better technology. Unfortunately the weak financial strength and poverty level creates the market to import these cheap and affordable alternatives. A research work in 2013 stated that Nigeria was at risk of becoming a dumping ground for obsolete electronics [16]. Virtually every electronic device consumed in Nigeria (home appliances, computers, accessories, electronics etc.) are imported. It was assumed from statistics that the share between used electronic devices and the new at a point was about 70/30% going by an e-waste assessment report in 2012 [1]. Approximately 30% of second-hand products were deemed non-functional and best classified as e-waste going by the report. About 100 Mt of e-waste flooded their way into Nigeria illegally in 2010 [1]. Around 6.4 Mt of that were personal household appliances and ICT devices, while institutional and corporate consumer devices were estimated at 400 thousand tons (Tt) [1]. About 440 Tt ends up as e-waste in landfills, repair and refurbishment shops, or just about anywhere [1]. According to the Nigerian Communication Commission (NCC), Nigeria's tele-density grew from 91.1% in 2013 to 99.4% in 2014. In 2015 and 2016, it rose to about 107.9 and 110.4% respectively [17]. By the end of 2018, E-Terra

fridaya@uef.fi

Table 3. E-waste types and their estimated life cycle. Extracted from [10–13].

Item mass of	Category	Label
Personal Computer (PC)	25	3
Fax machine	3	5
High-fidelity system	10	10
Cell phone	0.1	2
Electronic games	3	5
Photocopier	60	8
Radio	2	10
Television (TV)	30	5
Video recorder/DVD Player	5	5
Air-conditioner	55	12
Dish washer	50	10
Food mixer	1	5
Freezer	35	10
Hair-dryer	1	10
Iron	1	10
Kettle	1	3
Microwave	15	7
Refrigerator	35	10
Telephone	1	5
Toaster	1	5
Tumble dryer	35	10
Vacuum cleaner	10	10
Washing machine	65	8

Table 4. Contents of selected valuable metals in the printed circuit boards of seven types of electronic products [1].

Products	Valuable metals contained in products
TV (CRT monitor)	Gold, silver, copper, platinum, antimony, nickel, yttrium, neodymium, iron, and aluminum
Washing machine, Air conditioner, Refrigerator	Gold, silver, copper, platinum, antimony, iron, and aluminum
TV (LCD, plasma)	Gold, silver, platinum, antimony, indium, yttrium, iron, aluminum

Technologies estimated that Nigeria generates around 1.1 million tons of e-waste yearly [18]. E-Terra Technologies also placed Nigeria 3rd in the top 20 Countries according to an article by Kid Krunk [15]. The overall resulting effect is the huge pile of e-waste accumulating around the country, as all available repair outlets lack the capacity and appropriate technology to safely repair, recycle or get rid of them. The

fridaya@uef.fi

following researches have in the past discussed the Nigeria's e-waste problem from different perspectives:

1. "Variation In Some Haematological Parameters, Iron And Lead Levels In Workers Exposed To Electronic Waste In Benin City, Nigeria" 2018 [19].
2. "Environmental and Health Impacts Of Informal Electronic Waste Recycling" 2018 [20].
3. "Impact of informal electronic waste recycling on metal concentrations in soils and dusts" 2018 [21].
4. "Health Risks Awareness Of Electronic Waste Workers In The Informal Sector In Nigeria" 2017 [22].
5. "Material And Substance Flow Analysis Of Mobile Phones In Nigeria: A Step For Progressing E-Waste Management Strategy" 2017 [23]
6. "Electronic Waste Scenario in Nigeria: Issues, Problems and Solutions" 2013 [24].
7. "E-Waste Country Assessment Nigeria" 2012 [1].
8. "An Overview of ICT Waste Management: Suggestions of Best Practices from Developed Countries to Developing Nations (Nigeria)" 2011 [25].

Research recommendations so far center on the following points:

1. The need to enforce existing laws by developing effective strategies.
2. The need for Nigeria to immediately ratify the Bamako Convention by enforcing local regulations, intensifying prohibition on importation of hazardous wastes, enforce proof of test and categorization on trans-boundary movement of electronic products.
3. Producer companies should be compelled to use raw materials that are more environmentally friendly in the manufacturing of electronic products.
4. Developing countries should intensify the ban and prohibition on importation of used products marked for repair or refurbishing.
5. Encourage the development of local capacity for the safe extraction of valuable metals from e-waste.
6. Enforcement across producing countries the "Extended Producer Responsibility Programme"; which states that manufactures should buy back the products they produced at the end of their life span.

Despite research recommendations over the years, the implementation of these recommendations have either not been established or intensified. Nigeria is still flooded with e-waste and more flow in by the day. It is therefore important at this time to revisit the Nigeria e-waste problem, trace its source, review its negative impacts then identify and suggest based on efforts so far, ways to practically resolve and effectively manage the countries e-waste problem.

Approach Employed. This paper looked into the Nigerian electronic waste problem, traced its source, reviewed its negative impacts, identified and suggested based on efforts made so far by the Nigerian Government ways to improve the countries e-waste management strategies. It explored for lessons that can be learnt from other countries with good waste management, thus then strike a balance with the reality on ground and suggests ways forward.

fridaya@uef.fi

3 Tracing the Source of the Electronic Waste Problem in Nigeria

The Basel Convention, European Union directives on WEEE established it is unlawful to export e-waste from developed nations to developing nations [2]. Stakeholders and several researches have over the years advised same. Yet, e-waste is still exported to developing countries [26–28]. According to a survey conducted at Lagos in 2013, the volume of imported electronics was 15 tons at Computer village, 100 tons at Alaba International Market, 15 tons at 30 tons at Oshodi Market, Lawanson Market and 40 tons at West Minister [29]. These figures as reported in 2016 increased annually because the collection of e-waste was not organized [30]. Going by an article in 2018 by Green Compass Recycling, it was discovered that 90% of illegally shipped in e-waste to Africa ended up in Nigeria [27]. The article was a based on report titled, "e-Waste Recycling's Contribution to a Circular Economy". The report was presented at the Nigerian Economic Summit Group stakeholders' workshop in Lagos. It identified Nigeria as the third largest generator of e-waste in Africa, contributing 12% to the total volume, with 288,000 metric tons in 2017.

Another report in 2018 revealed over 66,000 tons of used electronics and electrical equipment flooded Nigeria between 2015 and 2016 [28]. Around 16,900 were toxic illegal e-waste because they were nonfunctional. The study was led by the Basel Convention Coordinating Centre for Africa and United Nations University. More than 200 shipping containers with 2,100 vehicles heading to Nigeria were inspected. About 70% (41,500 tons) of the UEEE reaching Lagos each year arrived in vehicles exported to Nigeria's used car markets, and about 18,300 tons arrived in shipping containers. About 77% of the used electronics were shipped from the European Union and 7% were sent from the US. According to the article no one was prosecuted and held responsible upon investigations and findings, neither the exporters nor the importers. This simply means the countries e-waste problem is worsening. Figure 1 shows the flow of electronic products as they come into a country.

Quite a number of issues have militated against the successful actualization of these directives. The Osun State House of Assembly member and Chairman House Committee on Industry, Commerce, and Cooperative Empowerment, 'Tunde Olatunji', pointed out the need for state parliaments to domesticate national laws, review existing laws, and issue legally binding targets and deadlines towards waste reduction [27]. A research in 2015 also stated that "most African countries have ratified the convention directives, but they have not domesticated it into national laws" [1]. As earlier pointed out, an article in 2018 with respect to illegally shipped in e-waste observed no one was prosecuted and held responsible upon investigations and findings, neither the exporters nor the importers. Thus, Governments of most developing nations are the major challenge to the Basel Convention directives. This is why the dumping has continued. Results cannot be achieved if these directives are not embedded as part of national laws and enforced, as there will be no grounds for prosecuting any offender.

fridaya@uef.fi

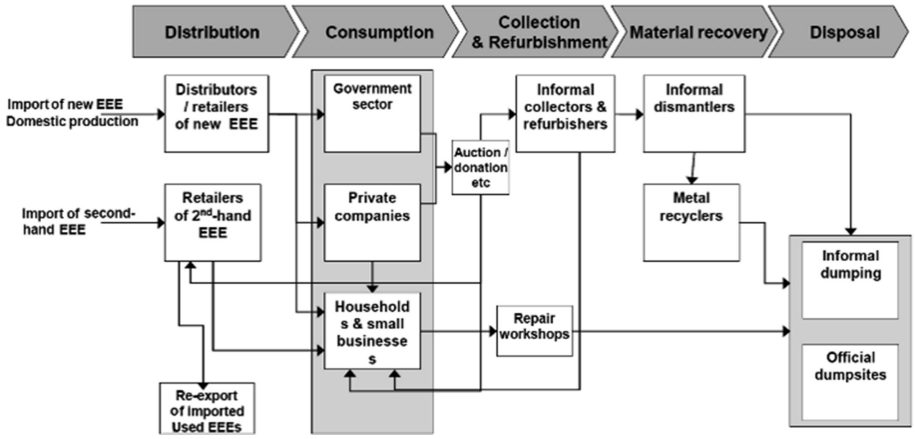


Fig. 1. A flow of products from importation to disposal [1].

4 Reviewing the Negative Impacts of E-Waste in Nigeria

E-waste procession of toxic substances was recognized only two decades ago, however it has been established that a large variety of such substances are contained in them [31]. These toxic substances pose health hazards to humans and are capable of contaminating the surrounding environment. It is thus paramount for this issue to be appropriately managed. One of the ways in which they are dangerous to humans involves the contamination of food. These toxic substances find their way into farmlands close to where e-waste are disposed or primitively recycled, consequently the farm produce affect human health when eaten. Another way is the direct impact on those involved or around primitive recycling sites. The result is being exposed through the air they breathe.

There also exists the risk of significant water contamination. On the long run these practices result in the poisoning of many locals within and around such sites. Over the years, incidences of excessive e-waste pollution were reportedly emitted from the e-waste recycling processes in Delta State, Nigeria [32], similarly in Accra, Ghana [33]. Puckett and Smith in 2002 revealed more than 1000 toxic substances that have been identified with improper e-waste recycling and disposal [34]. Some reported substances include: Manganese (Mn), Silver (Ag), Nickel (Ni), Molybdenum (Mo), Mercury (Hg), Lanthanum (La), Lithium (Li), Lead (Pb), Iron (Fe), Copper (Cu), Barium (Ba), Chromium (Cr), Cobalt (Co), Cadmium (Cd), Beryllium (Be), Hexavalent chromium Cr(VI). Also persistent organic pollutants (POPs) such as: Polychlorinated dibenzo-p-dioxins and Dibenzofurans (PCDD/Fs), DioxinPolyvinyl chloride (PVC), Polychlorinated biphenyls (PCBs), Brominated flame retardants (BFRs), Polycyclic aromatic hydrocarbons (PAHs), Polybrominated dibenzo-p-dioxins and dibenzofurans (PBDD/Fs). Table 5 shows common toxic substances associated with e-waste and their health impacts.

fridaya@uef.fi

Table 5. Effects of chemicals in e-waste on human health [1].

E-toxin	Source of e-toxin	Health effects
Arsenic	Found in computer chips and light emitting diodes	Arsenic is a known cancer-causing substance (carcinogen). It is known to cause skin and lung cancer
Brominated flame retardants	Added to plastic to prevent fires	Brominated flame retardants act as hormone disrupters. Children exposed to these substances show increased risk to thyroid disease and neurobehavioral disease
Cadmium	Cadmium coating of contacts and switches in the CPU and monitors is used to prevent corrosion. It is found in NiCd batteries	Breathing high levels of Cd can cause lung damage and death. Long term exposure to low levels of Cd can cause elevated blood pressure and kidney damage. Cadmium is a known carcinogen
Chromium	It is used as a hardener in plastics and a dye in pigments. It may be present in the coatings on some metal parts	Chromium has a variety of effects depending how it en-ters the body. Chromium is a carcinogen if inhaled. Chromium may also cause DNA damage
Halogens	Plastics and insulation	These substances are of concern because of the possibility that toxins such as dioxins and furans may be created and released burning
Lead	Cathode-ray tubes (about 5 lbs.) and solders	Initial symptoms of exposure are anorexia, muscle pain, malaise, and headache. Long-term exposure to lead de-creases the overall performance of the nervous system. High level exposure causes brain damage and death
Mercury	Cathode-ray tubes (about 5 lbs.) and solders	Initial symptoms of exposure are anorexia, muscle pain, malaise, and headache. Long-term exposure to lead de-creases the overall performance of the nervous system. High level exposure causes brain damage and death
Polyvinyl chloride (PVC)	PVC plastic is used in the insulation f some cables used in ICT equipment	When burnt it produces highly toxic dioxins; research is finding if PCV is a hormone disruptor

Various e-waste disposal methods exist detrimental to soil, water and air of those living around areas such is practiced. Table 6 identifies toxic substances related with e-waste recycling and the respective environmental mediums they could be disseminated through.

fridaya@uef.fi

Table 6. Extracted from: Toxic substances associated with recycling e-waste and their presence in the surrounding environment [7].

Environment	Toxic substances
Soil	PBDEs, PAHs, PCDD/Fs, As, Cu, Cr, Cd, Hg, Pb, Ag, Bi, Cd, Co, Cr, Cu, In, Hg, Mn, Mo, Pb, Sb, Sn, Tl, V and Zn
Water	As, Cd, Cr, Cu, F, Fe, Hg, Mn, Ni, Pb, Zn, Ag, Al, As, Be, Ca, Cd, Co, Cr, Cu, Fe, Li, Mg, Mn, Mo, Ni, Pb, Sb, Se, Sr, Ti, V and Zn
Air	PAHs, As, Cd, Cr, Cu, Mn, Ni, Pb, Zn, Ag, Bi, Cd, Co, Cr, Cu, In, Hg, Mn, Mo, Pb, Sb, Sn, Tl, and V

4.1 Landfills

Landfill sites also commonly referred to as tip, dump, rubbish dump or dumping ground, in this context are land sites allocated to disposal of e-waste materials and e-waste management processing (sorting, treatment, or recycling). Regardless of the global efforts put in place towards zero waste in recent times, more landfills can be observed across developing and under developed countries. Although owners of modern landfills are of the view that modern landfills are able to safeguarding the environment from e-waste pollutants [35], it has on the other hand been established they still cause groundwater contamination [36, 37]. Also it is certain that pollution tend to migrate via the soil and groundwater around and within landfill sites [38]. A decomposition of organic and putrescible material in landfills occurs via the soil through landfill leaching. Heavy metals, High concentrations of inorganic compounds and suspended/dissolved organic substances could be contained in leachates. The concentrations of the various toxic substances from leachates depend on the stages of waste decomposition and waste characteristics in a particular landfill [39]. Electronic devices are considered to be Toxicity Characteristic (TC) hazardous if they contain specific elements higher than TC regulated concentrations stipulated by the Resource Conservation and Recovery Act (RCRA). By this Act, they are considered TC hazardous if they contain specific elements higher, which are 1 mg/l of Se, 5 mg/l of Ag, 0.2 mg/l of Hg, 5 mg/l of Pb, 5 mg/l of Cr, 1 mg/l of Cd, 100 mg/l of Ba, 5 mg/l of As etc. [40]. Figure 2 shows leaching from electronic devices under laboratory-based TCLP conditions.

A study in 2018 at an e-waste site in Nigeria revealed the concentrations of 6 heavy metals [41]. Indigenous soil microbial communities were dominated by Planctomycetes, Acidobacteria, Firmicutes, Proteobacteria. Another study in 2018 investigated haematological indices, lead and serum iron levels in workers occupationally exposed to electronic wastes in South-South, Nigeria. The study enrolled participants from Benin City. Results obtained from this study show a significantly low serum iron (11.5 ± 2.14 μg/dL) in the EW compared with Control (168.9 ± 2.39 μg/dL); while blood lead level in EW (1.07 ± 0.02 μg/dL) was significantly elevated compared with Control (0.72 ± 0.12). The study concluded that chronic occupational e-waste exposure may increase blood lead level, lower serum iron status and may have adversely altered some important haematological parameters in the studied population.

fridaya@uef.fi

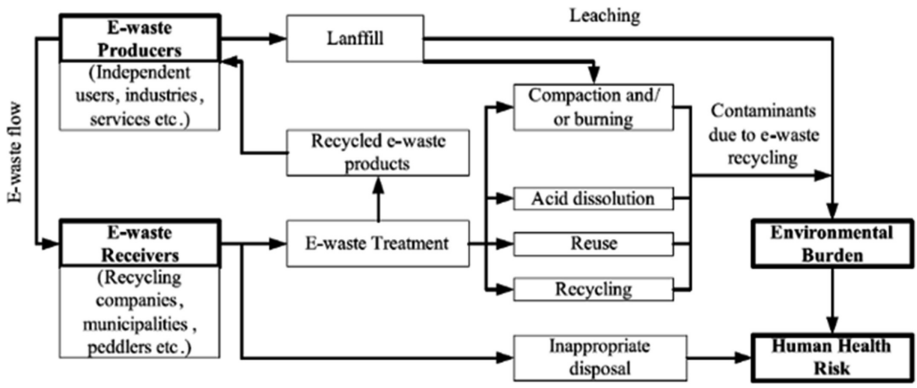


Fig. 2. Leaching process under TCLP [5].

4.2 Incineration and Local Recycling

There is an absence of the right set of recycling tools, technology and technical know-how in developing countries to properly perform incineration and recycling, Nigeria included. This leads to what is known as "backyard recycling" or "primitive recycling". This set of unsafe techniques end up having major impacts such as severe contamination of air from chlorinated compounds and brominated. Also from, heavy metals as detected at local recycling sites in developing countries. High concentrations of heavy metals such as Cu (483), Cr (1161) and Zn (1038) were detected in China via studies. These figures as revealed were between 4 and 33 times higher than those established across other Asian Countries. In Guiyu, China high pollution levels of PBDEs were 58–691 times higher than at other urban sites [42]. PCDD/Fs were detected at 64.9–2365 pg/m3, making them the highest concentrations in ambient air worldwide [43]. Environmental and human contamination at the e-waste recycling areas was also of great concern in Bangalore [44]. There, high levels of Ti, Sn, Mn, In, Cu, Cr, Co, Sb, Pb and Bi were detected in air around recycling areas as well as in India recycling sites [33]. Likewise the level of PPBDEs (BDE-17, 28, 47, 49, 66, 85, 99, 100, 153 and 154) in the indoor air of an e-waste storage facility were found at 46–350 pg/m3 in Thailand, on the other hand air pollution in outdoor locations, were estimated to be at about at 8 to 150 pg/m3 [45].

A Nigerian based study in 2017 assessed the practices, attitude and knowledge of e-waste workers association with the occupational health risk awareness via a control group in the informal sector [22]. The study sought establish the health risk awareness level across 279 e-waste dismantlers and repairers, 221 butchers from the informal sector in three locations. Results revealed a positive correlation existed between the workers' practice, attitude and knowledge. It was concluded that to decrease risky practices, there is a need to increasing workers' knowledge. Another study in 2018 sought establish the impact of various informal e-waste recycling activities (repairing, dismantling and burning) on metal concentrations in top soils and various dust samples [21]. In this study, comparative cross-sectional study design approach was employed to assess metal concentrations on samples from multiple e-waste recycling sites. Metal

fridaya@uef.fi

concentrations from a control were compared with that of Lagos, Ibadan, and Aba e-waste recycling sites. Results revealed that the Mean metal concentrations of the three study sites exceeded the concentrations at the control sites and the Nigerian standard guideline values by 100 s to 1000 s times. Burning sites showed the highest pollution level, followed by dismantling sites, lastly repair sites. Findings revealed serious environmental and public health concerns. The metal concentrations were also higher than levels reported in other studies at the same locations in Nigeria, indicating that the situation is worsening. Based on the results obtained, the study recommended the urgent development and implementation of effective support existing e-waste regulations in Nigeria. These findings and much more over the years have confirmed the significant levels of toxic substances that result from e-waste primitive recycling processes and their potential risk levels to the ecosystem and human health.

5 Suggestions on How to Improve the E-Waste Problem in Nigeria

In the comparative study of Switzerland, Japan and Greece, it was observed that one of the first country worldwide to establish an official e-waste management system was Switzerland [46]. They also as far back as 1998 introduced legislation for e-waste management via ordinance on "The Return, the Taking Back and the Disposal of Electrical and Electronic Appliances", ORDEA Law [47]. Consumers could either drop their e-waste at strategically located collection points, or take them directly to the recycling point where they are disassembled, shredded, sorted and disinfected to get rid of toxic substances. Manufacturers were fully made responsible for taking back their products from customers at no cost, and ensuring that they pose no threat to the environment [48].

In the case of Japan, the e-waste recycling system is based on social responsibility, environmental sensitivity and general discipline of Japanese people along with proper regulations put in place [5]. Their laws stipulate target rates and also penalties for violators [47]. Manufacturers in Japan are meant to be responsible for receiving products on return by last owners. Producers there have facilities and there exists collaborative partnerships between them and customers. They also have implemented in their business strategic e-waste management plans. Consumers return products to collection points where they proceeded to recycle facilities through distribution systems [12]. There also exists a withdrawal fees paid by consumers. Up till 2004 they had 41 e-waste recycling facilities partly funded by companies and ministries. It is also known that Japanese companies were the first to evolve welding without insulation and the electrical panel board connections without bromide compounds in accordance to the European guidelines on electronic products.

For a while Greece practiced an e-waste management system known as "grey recycling". This involved mixing e-waste with other materials such as metal waste for recycling. Approximately 170 kilotons (Kt), representing 3.8% of the total amount of domestic solid waste were processed using this method between 2003–2006. About 90% of Greece e-waste within that period were recycled or mixed with other materials. An alternative e-waste management system kicked off in 2004, with the responsibilities

fridaya@uef.fi

of collecting, transposing and processing e-waste in special facilities in other to effectively handle the developing problem increase in amounts of e-waste generated. The new approach recorded the collection of about 0.1 Kt of e-waste in 2005, also 31.5 Kt in 2007, 47 Kt in 2008 and 25 Kt during the first five months of 2009 [5]. This was achieved through the separate collection of at least 4 kg per resident per year of e-waste of domestic origin that is 44 Kt/year for Greece in total [5].

The first thing the Nigerian government should focus on now is raising proper awareness at all levels nationally. The National Environmental Standards and Regulations Enforcement Agency (NESREA) should also consider working with other stakeholders (both private and public) to create awareness on the e-waste problem, because efforts so far have grossly been insufficient. It can be recalled that NESREA in 2009, organized an International Conference titled "The Abuja Platform on E-Waste", the Environment Division of the Nigerian Society of Engineers also organized a Conference titled "Environmental Impact of Telecommunication Projects in Nigeria" in Abuja in 2010. Similarly in 2011, Nigeria held its 1st Eko International Summit on E-Waste at Lagos. These conferences aimed at e-waste control and to draw the attention of the Federal government to encourage, enforce collection, recovery, recycling and re-use of e-waste in Nigeria [49]. A research by Obaje in 2018 reported that while NESREA's efforts were commendable, country wide survey research revealed gross lack of awareness on the e-waste menace both within the literate and illiterate Nigerians [24]. As concluded by [22] in 2018, an increasing e-waste workers' knowledge may decrease risky practices.

It is believed when every stakeholder including the common man on the street down to school children are well aware and properly educated of the down sides associated with e-waste, no one would be interested in opting for products that neither may work let alone last nor keep them safe from the impending danger as a result of the toxic composition of its components. This would go a long way making it difficult for unscrupulous exporters making money from shipping in these junks to be patronized. It will thus be only a question of time before they begin considering channeling their investments into more environmentally friendly ventures.

The Federal Government of Nigeria in 2011 made efforts through agencies such as NESREA, to establish sound e-waste regulation starting with banning the importation of e-waste in Nigeria but allowing secondhand working devices to still scale through as cheap alternatives to those who cannot afford to buy new ones. Unfortunately smuggling of e-waste amongst second hand shipments started again. Today we are back to square one. Thus to address this lingering problem further steps have to be taken by the Federal Government and all agencies responsible, beyond efforts made so far to counter all improvised techniques developed over time in smuggling e-waste into the country.

5.1 How Can the Nigerian Government Counter All Improvised Techniques Developed Over Time in Smuggling E-Waste into the Country?

There exists a challenge as how to effectively classify each used electrical electronics shipped from outside Nigeria as second-hand or e-waste. The lack of clarity with regards to this issue makes it difficult for enforcement officers to distinguish between

fridaya@uef.fi

second-hand sound working devices and e-waste when screening shipments documents and during visual inspections. In most cases, exporters of these products deliberately do not declare these goods as waste to be able to smuggle them into developing countries and make money. In consideration to the fact that not every Nigerian citizen can afford new product couple with the goal of bridging the digital divide, the Federal government can place a total ban on the direct importation of second hand products by setting up another agency saddled with the responsibility of liaising directly with manufacturers in major countries and others across the world through whom cheap working alternatives can pass through. In other words this will narrow down the inflow of second hand products to just one source; a government agency that will ensure what it is shipping in are deemed sound and still in good working condition. With this, any further inflow of e-waste can then be blamed on that single established gateway. Once the inflow of e-waste has been taken care of properly, the country will then be faced with the next stage of getting rid of tons of e-waste already within Nigeria. According to [50], an encouraging success story was recorded of NESREA for successfully detecting and impounding two containers at TinCan Island, Lagos, Nigeria from France, flooded with electronic products that did not have exporter's test certification. Quite a laudable efforts and action it was from NESREA, however another set of questions still springs up. Has that effort been sustained since then? How do we deal with the excessive tons of e-waste that has already found its way and are currently within Nigeria?

5.2 How Can Efforts so Far Be Sustained? How Do We Deal with the Excessive Tons of E-Waste Already in Nigeria?

Agencies such as NESREA need to intensify efforts and actions to ensure the sustenance of such success stories. To successfully accomplish this task they continually have to strive to be steps ahead of smugglers in other to catch them. With regards to handling the excessive ton of e-waste already in the country; some measures are in place but does not cut across the entire nation. For example, e-waste collection exercises exist in Lagos handled by the Lagos Waste Management Authority (LAWMA). After gathering e-wastes, they are transferred to the available municipal dumpsites. It is there informal collectors sort out fractions valuable to them. Likewise, the Lagos State Environmental Protection Agency (LASEPA) is also saddled with e-waste collection focusing on e-waste from businesses. However this formal arrangement is not obtainable across the entire country. The only activities on ground elsewhere in Nigeria is the gathering and local recycling of e-waste by scavengers who lack the technical know-how for safe recycling let alone the right technology, using crude methodologies to recycle. They randomly go house to house, gathering metallic wastes steel, aluminum and copper, a times paying some money for each item. This is a good approach but must be restructured. Nigeria is blessed with land mass and collection centers can be strategically set up across all 36 states of the federation to facilitate collection. The government can provide incentive to these scavengers and engage them in gathering these e-waste to the nearest established collection site. When these pile up over time, measures can then be taken for proper scavenging of useful materials and recycling. The Nigerian Government could also provide incentive to Nigerians to encourage them let go of such e-waste to collectors for a token amount. All these will boost the rate at

fridaya@uef.fi

which the nation will be purged of e-waste lying around every home. On successful compilation of e-waste across the nation through strategically located collections centers, the next huddle to cross would be how to get them properly recycled or disposed of.

5.3 After the Successful Gathering of E-Waste Across Collection Centers, What Next?

An International Summit organized by Lagos State LASEPA in conjunction with NGOs, took place at Lagos in 2011. One major outcome of this event was the idea to form an alliance on African e-waste. Prominent companies such as Phillips, HP, Dell and Nokia were involved and part of the achievements so far was a pilot recycling program in Kenya. Nigeria can also get started on establishing a pilot recycling program as such to help when accumulated e-wastes pile up, to have it successfully moved from collection sites. If after due consideration this pilot program is not feasible, an alternative would be to saddle the earlier suggested agency (to handle the coordination of shipping in certified second hand working products), with an additional responsibility of establishing contacts with prospective interested recycling companies closest to Nigeria such as Kenya. Thus, as the agency helped coordinate the inflow of usable product, they should as well coordinate the exporting of now e-waste products to appropriate recycling plants, through the appropriate channels. Thus, if going the extra mile of dismantling, removing plastic, circuit board, smelting with the right technology to get the gold and silver is not feasible, they can simply be shipped to where it can be properly done. The cost of flushing out this toxic waste should be considered an investment and a sacrifice in comparison to the possible adverse effect they are capable of posing.

6 Recommendations

This paper makes the following recommendations that will ensure an effective management of the Nigerian e-waste problem:

1. There is need for an immediate exponential increase in every effort made so far in creating awareness. With social media and other advanced information dissemination mediums, the awareness campaigns can be overhauled with very little effort. It is suggested that no stone be left unturned, citizens from all works of life including children nationwide be carried along as this would drastically help establish the right mind set and attitude needed to facilitate the solution to our e-waste problem.
2. It is also recommended that collection centers across the nation be strategically established nationwide for the effective collection and accumulation of these e-wastes from the junk collectors where they can be properly sorted and stored as they await further processing.
3. Measures should be put in place to intensify the existing house-to-house collection of e-waste techniques by junk collectors alongside the voluntary take-back

fridaya@uef.fi

technique practiced in Europe, but restricting them to collection and gathering only. They should be prevented from making any attempts to recycle or extract any materials as they lack the technology and technical know-how. The junk collectors should be encouraged with some form of incentives to help gather these e-wastes to designated collection sites as it is believed that that this house-to-house collection technique achieves significantly higher and rapid collection rates compared to other methods and is more convenient for consumers.

4. Agencies responsible for health and occupational safety issues, environmental sanitation and national security must at this point come in and ensure effective enforcement of laws and regulations preventing junk collectors, scavengers from using crude methodologies to extract materials and recycle e-waste, thus rather ensuring they deliver all gathered waste at appropriate collection sites.
5. Having achieved proper general public awareness and the right set of measures in place to ensure effective e-waste gathering, some form of incentive to citizens will also go a long way in boosting their willingness to trade off as much e-waste in their possessions as possible.
6. An immediate ban should be placed on the importation of all used electronic products and devices, after which the government can take advantage of the situation, stream line and narrow down their importation to one single source; a government established agency saddled with the responsibility of liaising with foreign governments, organizations, manufacturers and companies. This will help properly scrutinize each and every used device to be shipped into the country as cheap alternative to those who cannot afford new ones, making sure they are still in good working condition.
7. Lastly, notable exporting countries of used electronic products should be urged to enforce existing laws, such as the proof of testing and categorization, and other directives of the Basel Convention.

It is therefore hopeful, if Nigeria puts all suggestions in this paper in place. There would barely be any avenue for inflow of e-waste into the country. It is time Nigeria and other developing nations realize they are capable of solving the e-waste problem on their own.

7 Conclusion

E-waste is clearly a growing problem especially amongst developing countries and possesses the potential of adversely affecting the water, air and lives of people who find themselves around such areas. Thus, as a matter of urgency the Nigerian government must swing into action and take extra measures aside depending on International Laws. This will ensure the situation is properly managed. This paper establishes a need to consistently ensure that every measure put in place is sustained to yield results, back up plans and strategies be continuously put in place in event of failed plans until the problem of e-waste is managed across the country.

fridaya@uef.fi

References

1. Ogungbuyi, O., Nnorom, I.C., Osibanjo, O., Schluep, M.: E-waste country assessment Nigeria. In: A Report, e-Waste Africa Project of the Secretariat of the Basel Convention (2012)
2. Terada, C.: Recycling electronic wastes in Nigeria: putting environmental and human rights at risk. Northwest. J. Int. Hum. Rights **10**(3), 154–172 (2011)
3. "Definition of Waste", United Nations Statistics Division (UNSD). http://www.informalwastesector.net/topics/view/51cbfc93f702fc2ba812dfcd/. Accessed 29 Nov 2018
4. "Resource Productivity and Waste", Organisation for Economic Co-operation and Development (OECD). http://www.oecd.org/env/waste/. Accessed 11 Dec 2018
5. Gaidajis, G., Angelakoglou, K., Aktsoglou, D.: E-waste: environmental problems and current management. J. Eng. Sci. Technol. Rev. **3**(1), 193–199 (2010)
6. "Categories of E-Waste", African E-Waste Solution. http://www.africaewaste.co.za/index.php/what-is-e-waste/categories-of-e-waste. Accessed 29 Nov 2018
7. Kiddee, P., Naidu, R., Wong, M.H.: Electronic waste management approaches: an overview. Waste Manag. **33**(5), 1237–1250 (2013)
8. Statista: Forecast of electronic waste generated worldwide from 2010 to 2018 (in million metric tons) (2018). https://www.statista.com/statistics/499891/projection-ewaste-generation-worldwide/. Accessed 27 Sept 2018
9. Hischier, R., Wäger, P., Gauglhofer, J.: Does WEEE recycling make sense from an environmental perspective? The environmental impacts of the swiss take-back and recycling systems for waste electrical and electronic equipment (WEEE). Environ. Impact Assess. Rev. **5**, 525–539 (2005)
10. Betts, K.: Producing usable materials from e-waste. Environ. Sci. Technol. **42**, 6782–6783 (2008)
11. Cobbing, M.: Toxic tech: not in our backyard: uncovering the hidden flows of e-waste. Report from Greenpeace International (2008)
12. Li, J., Gao, S., Duan, H., Liu, L.: Recovery of valuable materials from waste liquid crystal display panel. Waste Manag. **29**(7), 2033–2039 (2009)
13. Robinson, B.H.: E-waste: an assessment of global production and environmental impacts. Sci. Total Environ. **408**(2), 183–191 (2009)
14. Vos, S.: Electronic Waste Disposal. Lori Bennear Research Group, Duke University. http://sites.nicholas.duke.edu/loribennear/2012/11/15/electronic-waste-disposal/. Accessed 11 Dec 2018
15. E-Terra Technologies. Top 20 Countries That Are Used as Dumping Grounds of the World's Trash. http://www.eterra.com.ng/articles/top-20-countries-dumping-grounds-trash/. Accessed 27 Nov 2018
16. Agunwamba, J.C., Emenike, P.C., Tenebe, I.T.: Comparative analysis of hospital waste management in calabar metropolis and developed countries. Int. J. Struct. Civil Eng. Res. **2**(4), 154–167 (2013)
17. "Subscriber/Teledensity Data" Nigerian Communication Commission (NCC). http://www.ncc.gov.ng/stakeholder/statistics-reports/industry-overview#annual-2002-2016. Accessed 27 Nov 2017
18. E-Terra Technologies. E-Waste – Nigeria, West Africa on the Brink of a Disaster. http://www.eterra.com.ng/news/e-waste-nigeria-west-africa-brink-disaster/. Accessed 27 Nov 2018

fridaya@uef.fi

19. Igharo, O.G., Anetor, J.I., Osibanjo, O., Okungbowa, M.A., Idomeh, F.A., Aleoghena, T.O.: Variation in Some Haematological Parameters, Iron and Lead Levels in Workers Exposed to Electronic Waste in Benin City, Nigeria (2018)
20. Ohajinwa, C.M.: Environmental and health impacts of informal electronic waste recycling. PhD dissertation (2018)
21. Ohajinwa, C.M., van Bodegom, P.M., Vijver, M.G., Peijnenburg, W.J.: Impact of informal electronic waste recycling on metal concentrations in soils and dusts. Environ. Res. **164**, 385–394 (2018)
22. Ohajinwa, C., Van Bodegom, P., Vijver, M., Peijnenburg, W.: Health risks awareness of electronic waste workers in the informal sector in Nigeria. Int. J. Environ. Res. Public Health **14**(8), 911 (2017)
23. Babayemi, J.O., Osibanjo, O., Weber, R.: Material and substance flow analysis of mobile phones in Nigeria: a step for progressing e-waste management strategy. J. Mater. Cycles Waste Manag. **19**(2), 731–742 (2017)
24. Obaje, S.O.: Electronic waste scenario in Nigeria: issues, problems and solutions. Int. J. Eng. Sci. Invent. **2**(11), 31–36 (2013)
25. Adeola, A.M., Othman, M.: An overview of ICT waste management: suggestions of best practices from developed countries to developing nations (Nigeria). In: 7th International Conference on Network Computing, pp. 109–115 (2011)
26. "Undercover Operation Exposes Illegal Dumping of E-waste in Nigeria", Greenpeace. www.greenpeace.org/international/en/news/features/e-waste-nigeria. Accessed 09 Feb 2018
27. The Punch newspaper: 90% of Africa's illegally imported e-waste end in Nigeria—Report. https://punchng.com/90-of-africas-illegally-imported-e-waste-end-in-nigeria-report/. Accessed 11 Dec 2018
28. The World Is Sending Tons Of Illegal Electronic Waste To Nigeria. http://www.konbini.com/ng/lifestyle/the-world-is-sending-thousands-of-tonnes-of-electronic-waste-to-nigeria-illegally/. Accessed 23 Apr 2018
29. Amachree, M.: Update on e-waste management in Nigeria. In: 3rd Annual Meeting of the Global E-Waste Management Network, San Francisco (2013). https://www.epa.gov/sites/production/files/2014-05/documents/nigeria.pdf. Accessed 30 June 2017
30. The Daily Times. Dumping of e-waste in Nigeria. https://dailytimes.ng/editorial/dumping-e-waste-nigeria/. Accessed 11 Dec 2018
31. Zhang, S., Forssberg, E.: Mechanical separation-oriented characterization of electronic scrap. Resour. Conserv. Recycl. **21**(4), 247–269 (1997)
32. Puckett, J., Westervelt, S., Gutierrez, R., Takamiya, Y.: The Digital Dump: Exporting Re-use and Abuse to Africa. The Basel Action Network, Seattle, USA, pp. 85 (2005)
33. Asante, K.A., Agusa, T., Biney, C.A., Agyekum, W.A., Bello, M., Otsuka, M., Itai, T., Takahashi, S., Tanabe, S.: Multi-trace element levels and arsenic speciation in Urine of e-waste recycling workers from Agbogbloshie, Accra in Ghana. Sci. Total Environ. **424**, 63–73 (2012)
34. Puckett, J., Smith, T.: Exporting harm the high-tech trashing of Asia. In: Coalition, S.V.T. (ed.) (2002)
35. Swana: The Environmental Consequence of Disposing of Products Containing Heavy Metals in Municipal Solid Waste Landfills. Solid Waste Association of North America, Silver Spring, MD (2004)
36. Schmidt, C.W.: e-Junk explosion. Environ. Health Perspect. **110**(4), A188 (2002)
37. Yang, G.C.: Environmental threats of discarded picture tubes and printed circuit boards. J. Hazard. Mater. **34**(2), 235–243 (1993)

fridaya@uef.fi

38. Kasassi, A., Rakimbei, P., Karagiannidis, A., Zabaniotou, A., Tsiouvaras, K., Nastis, A., Tzafeiropoulou, K.: Soil contamination by heavy metals: measurements from a closed unlined landfill. Biores. Technol. **99**(18), 8578–8584 (2008)
39. Qasim, S.R., Chiang, W.: Sanitary Landfill Leachate: Generation, Control and Treatment. CRC Press, Boca Raton (1994)
40. Townsend, T.G., Vann, K., Mutha, S., Pearson, B., Jang, Y.C., Musson, S., Jordan, A.: RCRA toxicity characterization of electronic wastes. Hazard. Waste Consultant **23**, 16–18 (2005)
41. Jiang, B., Adebayo, A., Jia, J., Xing, Y., Deng, S., Guo, L., Liang, Y., Zhang, D.: Impacts of heavy metals and soil properties at a Nigerian e-waste site on soil microbial community. J. Hazard. Mater. **362**, 187–195 (2019)
42. Deng, W.J., Zheng, J.S., Bi, X.H., Fu, J.M., Wong, M.H.: Distribution of PBDEs in air particles from an electronic waste recycling site compared with Guangzhou and Hong Kong, South China. Environ. Int. **33**(8), 1063–1069 (2007)
43. Li, H., Yu, L., Sheng, G., Fu, J., Peng, P.A.: Severe PCDD/F and PBDD/F pollution in air around an electronic waste dismantling area in China. Environ. Sci. Technol. **41**(16), 5641–5646 (2007)
44. Ha, N.N., Agusa, T., Ramu, K., Tu, N.P.C., Murata, S., Bulbule, K.A., Parthasaraty, P., Takahashi, S., Subramanian, A., Tanabe, S.: Contamination by trace elements at e-waste recycling sites in Bangalore, India. Chemosphere **76**(1), 9–15 (2009)
45. Muenhor, D., Harrad, S., Ali, N., Covaci, A.: Brominated flame retardants (BFRs) in air and dust from electronic waste storage facilities in Thailand. Environ. Int. **36**(7), 690–698 (2010)
46. Liu, X., Tanaka, M., Matsui, Y.: Generation amount prediction and material flow analysis of electronic waste: a case study in Beijing, China. Waste Manag. Res. **24**(5), 434–445 (2006)
47. Widmer, R., Oswald-Krapf, H., Sinha-Khetriwal, D., Schnellmann, M., Böni, H.: Global perspectives on e-waste. Environ. Impact Assess. Rev. **25**(5), 436–458 (2005)
48. Fishbein, B.K.: Waste in the Wireless World: The Challenge of Cell Phones. Inform Inc., New York (2002)
49. Eko Declaration, Communiqué of the 1st Eko E-Waste Summit on Regulation and Management of E-Waste in Nigeria held in Lagos, Nigeria, 24–25 February 2011
50. Environment Media Group: "Agency Investigates E-Waste Shipment to Nigeria", Let's Recycle. https://www.letsrecycle.com/news/latest-news/agency-investigates-e-waste-shipment-to-nigeria/. Accessed 11 Dec 2018

fridaya@uef.fi

High-Throughput Electronic Band Structure Calculations for Hexaborides

Zhenxi Pan, Yong Pan, Jun Jiang, and Liutao Zhao(✉)

Beijing Computing Center, Blgd. 3, No.7 Fengxian Road, Haidian District, Beijing, China
{panzx, zhaolt}@bcc.ac.cn

Abstract. Several specific physical properties of alloys can be determined by first-principles (typically via density functional theory, DFT) due to the enormous gain in high-performance computing over the last decades. This article is devoted to the discussion of the high-throughput approach to band structures calculations. We present the computational key for the auto-generation band structure on the high-throughput calculating platform. This method is based on a standard creation of k-point paths for band. This standardization is automatic and fully integrated for the 14 Bravais lattices, the conventional and primitive unit cells, and the coordinates of the high symmetry k-path in the Brillouin zones. Some band structures for transition metal doped Hexaborides are given as example.

Keywords: High-throughput · Computational materials science · First-principles · Brillouin zones

1 Introduction

1.1 A Subsection Sample

Please note that the first paragraph of a section or subsection is not indented. The first paragraphs that follows a table, figure, equation etc. does not have an indent, either.

Materials science has developed under the influence of physics, chemistry, engineering mechanics, applied mathematics and many other disciplines. Since the 1990s, computational materials physics and quantum chemical method developed on top speed, computer technology and software development made continuous improvement. Based on these, calculation has played a more and more important role in the R&D of new materials. As a new subject [1–6], computational materials science has achieved unprecedented development in the last 20 years. Especially with the breakthrough of high-precision quantum mechanics and high performance parallel computing technology [7, 8], high-throughput material computing is gradually showing strong potential in accelerating the innovation and exploration of new materials and new phenomena [9].

High-throughput calculation of materials has been successfully applied and attempted in a broad range covering, such as systems like, forecast of energy materials [10], topological insulator discovery [11], thermoelectric materials [12], catalytic

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 386–395, 2019.
https://doi.org/10.1007/978-3-030-22868-2_29

fridaya@uef.fi

materials [13], lightweight magnesium alloy research [14], superconducting materials [15], magnetic materials [16, 17], complex multi-component surface design [18, 19], determinations about the stability of binary or ternary compounds [20], high-strength and high-temperature alloy research. In addition, a mass of open databases of properties of materials has been created in design and optimization of new materials.

With the increasing of mock objects, bridging multi-scale needs to be considered and calculation process is becoming more and more complicated. Therefore, it is necessary to design the automatic process to conveniently commit the design and optimization of the calculation materials. At the moment, known software articles of automatic process of high-throughput ab initio material design on a global scale are as follows: Automatic Flow (AFLOW) [21], Materials Project (MP) [22–25], Quantum Materials Informatics Project (QMIP) [26], Clean Energy Project (CEP) [27] and so on.

In calculation of materials, calculation of electronic structure is the most fundamental. Mainly based on ab initio density functional theory (DFT), uses plane wave or add local atomic orbitals as the basis function to calculate the electron density of the ground state by self-consistent iteration. Further, such physical and chemical properties of the system, such as band structure, state density (DOS), etc. Band structure shows the distribution of electronic energy in k-space, thus two-dimensional energy band diagram is closely related to selection of points arranged in the k-space layout. Conventionally, k-point for drawing energy band is chosen along the highly symmetric path of irreducible Brillouin zone (IBZ); whereas, the direction of k-point chosen is clearly artificial (random) and unfavorable for drawing of energy band of automatic process. Setyawan [21] et al. designed the automatic scheme of energy band k-point path generation and applied to AFLOW. But AFLOW is not open source or free, and it mainly for simple materials, therefore we build standardized automatic generation of all kinds of structures based on the scheme of document [21] and the related modules of Materials Project [22–25] which is open source. We will discuss how to confirm point group and space group by materials' atomic location in the second part of this article; Part 3 is for introduction to the standard k-point scheme in Literature [21]; Part 4 is for calculation of the energy band structure of CaB_6 and EuB_6 in an automatic process; and make the conclusion in the last part.

2 Standardized of K-Path and Point and Space Group

2.1 Standardized of K-Path

It is very convenient to sequence the k-point standardization scheme in Literature [21] by applying the process of standardizing primitive cells of a unit cell to a reciprocal space. According to DFT, calculation of electronic structure is generally required of relaxation calculation to get a stable ground-state unit cell structure and static calculation to get ground-state self-consistency electronic density; on that account, energy band calculation is completed. Therefore, upon completion of electronic density relaxation by static calculation, standardized k-point module is called to generate a k-path and complete energy band calculation to realize a "closed-loop" automatic process. In fact, the standardization of k-path helps automation of computing electronic

fridaya@uef.fi

structure of materials and greatly improves computing efficiency of high-throughput calculation. More importantly, this standard choice of k-path can prevent erroneous judgment of energy band calculation on the important properties such as band gap. Setyawan et al. indicated [21] that body-centered cubic GeF_4 has indirect gaps (Fig. 1); the conduction band minimum (CBM) is located at point Γ while the valence band maximum (VBM) is located at point H of the body-centered cubic; whereas, if the corresponding k-path does not include the point H (for instance, it is replaced by the k-path of cubic), GeF4 will have direct gaps in the point Γ only.

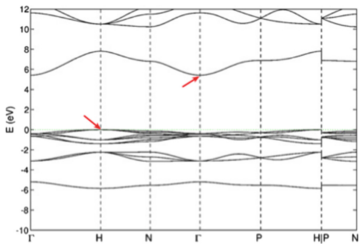


Fig. 1. Band structure of BCC-GeF_4 with the CBM and VBM indicated by red arrows (from [21]).

With regard to standardized reciprocal space, the scheme in Literature [21] enumerates k-paths of all kinds of energy band structure composed by piecewise traversal all high-symmetry point-to-point corresponding to 14 Bravais lattices. K-path schemes, such Bravais lattices as cubic, face-centered cubic (FCC), body-centered cubic (BCC), orthogonal, and orthorhombic, are illustrated in this article while k-path schemes of other Bravais lattices are shown in Literature [21].

High symmetry points of simple cubic (fractional coordinates) are taken as (Fig. 2):

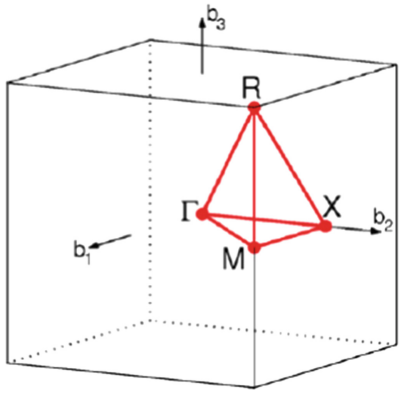


Fig. 2. Brillouin zone of Cubic lattice and the standard k-path (from [21]).

fridaya@uef.fi

$$\Gamma(0,0,0), M(1/2,1/2,0), R(1/2,1/2,1/2), X(0,1/2,0)$$

Its optional k-path sequence is as follows:

$$R-\Gamma-X-M-\Gamma-X-R|M-R$$

High symmetry points of FCC (fractional coordinates) are taken as (Fig. 3):

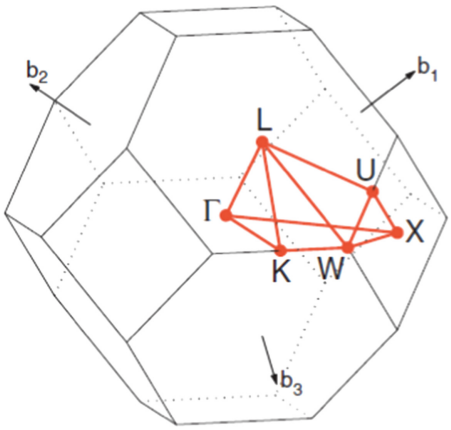


Fig. 3. Brillouin zone of Face-Centered Cubic lattice and the standard k-path (from [21]).

$$\Gamma(0,0,0), K(3/8,3/8,3/4), L(1/2,1/2,1/2),$$
$$U(5/8,1/4,5/8), W(1/2,1/4,3/4), X(1/2,0,1/2)$$

Its optional k-path sequence is as follows:

$$\Gamma-X-W--K-\Gamma-L-U-W-L-K|U-X$$

High symmetry points of BCC (fractional coordinates) are taken as (Fig. 4):

$$\Gamma(0,0,0), H(1/2,-1/2,1/2), P(1/4,1/4,1/4), N(0,0,1/2)$$

Its optional k-path sequence is as follows:

$$\Gamma-H-N-\Gamma-P-H|P-N$$

High symmetry points of simple tetragonal (fractional coordinates) are taken as (Fig. 5):

fridaya@uef.fi

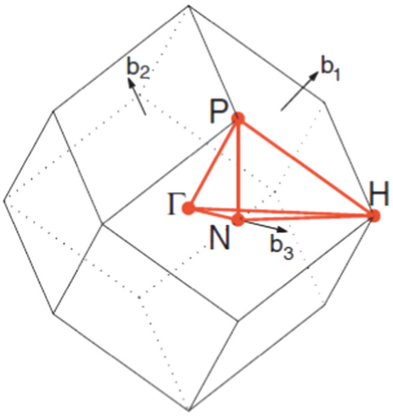


Fig. 4. Brillouin zone of Body-Centered Cubic lattice and the standard k-path (from [21]).

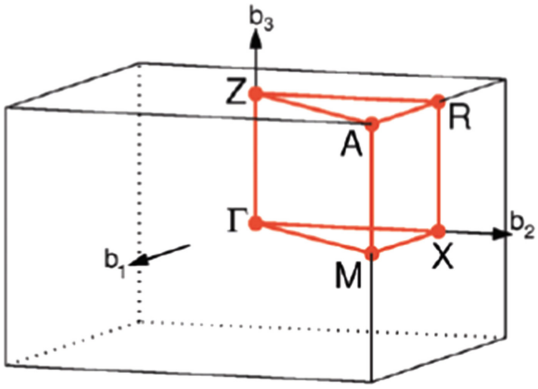


Fig. 5. Brillouin zone of Tetragonal lattice and the standard k-path (from [21]).

$$\Gamma(0,0,0), A(1/2,1/2,1/2), M(1/2,1/2,0), R(0,1/2,1/2), X(0,1/2,0), Z(0,0,1/2)$$

Its optional k-path sequence is as follows:

$$\Gamma - X - M - \Gamma - Z - R - A - Z|X - R|M - A$$

High symmetry points of simple orthogonal (fractional coordinates) are taken as (Fig. 6):

$$\Gamma(0,0,0), R(1/2,1/2,1/2), S(1/2,1/2,0), T(0,1/2,1/2),$$
$$U(1/2,0,1/2), X(1/2,0,0), Y(0,1/2,0), Z(0,0,1/2)$$

fridaya@uef.fi

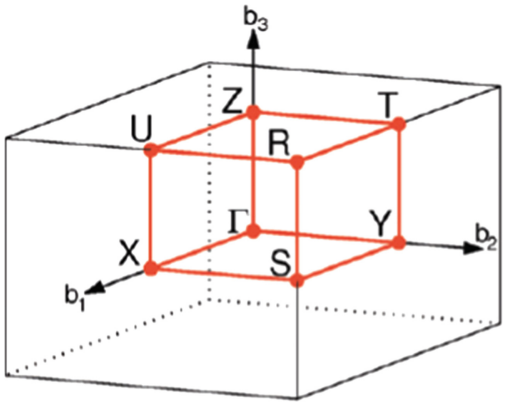


Fig. 6. Brillouin zone of orthorhombic lattice and the standard k-path (from [21]).

Its optional k-path sequence is as follows:

$$\Gamma - X - S - Y - \Gamma - Z - U - R - T - Z|Y - T|U - X|S - R$$

2.2 The Method of K-Path Standardizing in ABINIT

The electronic structure calculation is completed by open-source software of ABINIT [28–30]. ABINIT realizes ab initio calculation in the projector augmented wave (PAW) method of pseudo-potential-plane wave basis set [31] in spite of its built-in structural symmetry judgment module function. In order to better combine with Pymatgen to generate standardized automatic energy band structure, related structural symmetry judgment program is created according to the fundamental relationship between point group and space group in a translation period system in the group theory [32].

The most conveniently way of express symmetry k-point is fractions of the reciprocal lattice vectors. In order to achieve the symmetry k-point standardization, the normalization of the primitive cell representation by lattice vector must be achieved. The steps to achieve the standardization of unit cell structure are as follows:

(1) For a given unit cell structure input file, search and determine the shortest lattice vector based on its equivalent atom position:

Subject to given initial vectors $\vec{a}_1, \vec{a}_2, \vec{a}_3$ in the unit cell structure file, Bravais lattice which the unit cell belongs to and corresponding unit cell parameters $a, b, c; \alpha, \beta, \gamma$ are determined;

Vectors $\vec{a}_1, \vec{a}_2, \vec{a}_3$ are checked in iteration in order and the minimum iterative vectors $\vec{a}_1^{\min}, \vec{a}_2^{\min}, \vec{a}_3^{\min}$ are determined;

Detect and determine the minimum repeated vectors $\vec{a}_1^{\min}, \vec{a}_2^{\min}, \vec{a}_3^{\min}$ of vectors $\vec{a}_1, \vec{a}_2, \vec{a}_3$ in turn with iterative method;

fridaya@uef.fi

The ultimately confirmed minimum repeated vectors $\bar{a}_1^{\min}, \bar{a}_2^{\min}, \bar{a}_3^{\min}$, forms the initial vector of the primitive cell.

(2) Confirm standardized primitive cell:

Atomic coordinates in the unit cell are forcibly put in place of periodic condition coordinates (or atomic coordinates are transferred to the range [−0.5, 0.5]), all types of atom i are sorted in ascending order by coordinate value and recorded as array **TOR**[i] while **type and number of atoms of the lowest atomicity** are determined.

By reference to the first atom in the type of atoms of the lowest atomicity, the "translation vector" relationship between the atom and other atoms is separately determined and translation vectors determined are saved in the array **TRA**;

Atoms in each type in the unit cell are checked, the "translation vector" effects (**TOR**[i]+**TRA**) are used respectively to draw positions of atoms after translation and similarly sort such atoms in the ascending order and arrange such atoms in the array **TOR_TRA**[i].

The method of exhaustion (one-to-one) is used to compare atomic coordinates in two lattices (**TOR**[i]. vs. **TOR_TRA**[i]) and guarantee that atomic coordinates in such two lattices coincide with each other (criterion for judgment: Difference $<10^{-6}$);

All the translation vectors in the above process are summed up and recorded and minimum vectors in directions $\vec{a}_1, \vec{a}_2, \vec{a}_3$ are determined and saved in the array ***PTRAN[i]*** (expressed as internal coordinates), which in combination with $\bar{a}_1^{\min}, \bar{a}_2^{\min}, \bar{a}_3^{\min}$, are used to ultimately determine initial vectors $\vec{p}_1, \vec{p}_2, \vec{p}_3$ of primitive cell.

(3) The number of multiple cell *NCELL* is judged subject to volume determined according to vectors of unit cell and vectors of primitive cell:

$$NCELL = \frac{V_\Omega}{V_{\Omega_0}} = \frac{(\vec{a}_1 \times \vec{a}_2) \cdot \vec{a}_3}{(\vec{p}_1 \times \vec{p}_2) \cdot \vec{p}_3} \tag{1}$$

The number of primitive cells in each direction in the unit cell is rechecked and licensed number of standard primitive cells in each direction is determined.

(4) Located symmetry operation point group which unit cell belongs, and traverse all point groups in turn:

All the point group's basic operating matrices **Point_M**[j] are reacted in order to atomic coordinates **TOR**[i]. If coordinates drawn are corresponding to positions of equivalent coordinates of atom, it shall be determined as a licensed point group operation **Point_M**[j].

Upon reaction of all the point group's basic operating matrices **Point_M**[j] in order to atomic coordinates **TOR**[i], if extra translation operation **PTRAN**[j] (slip) is required to get positions of equivalent coordinates of the corresponding atom, it shall be determined as an accompanying translation group to the point group.

fridaya@uef.fi

What must be indicated is that the operation of traversing point groups must be executed in each type of atoms, whose results must be applicable to all atoms; otherwise, it cannot constitute an effective symmetric operation.

Special attention must be attached to: since a unit cell is probably primitive, non-primitive slip (trivial translations of generating cell) must exist. In order to avoid the non-uniqueness of definition of slip, it is stipulated to take the minimum vector of the coordinate value from all the licensed slip vectors.

(5) Combine point group and slip information to ensure the space group to which the unit cell belongs.

3 Band Structure of EuB6 for Automatic Process Calculation

Divalent metal Hexaborides EuB_6 has many special properties. For instance, it has two ferromagnetic phase transitions in temperature $(T_{C1} = 15.1K, T_{C2} = 12.7K)$, followed by significant change in optical, electronic and transport properties [33–35]. In order to expound the micro mechanism of such phenomena, it is necessary to have an in-depth understanding about the electronic structure of EuB_6. Jiang et al. indicated in calculation [36] that the energy band of EuB_6 was typical semi-metal (Fig. 7). K-path auto-generation module is used to generate k-points required for drawing an energy band diagram, realize automation of the whole process of electronic calculation and get the energy band of EuB_6 in view of spin-polarization. Calculations show that the spin-up

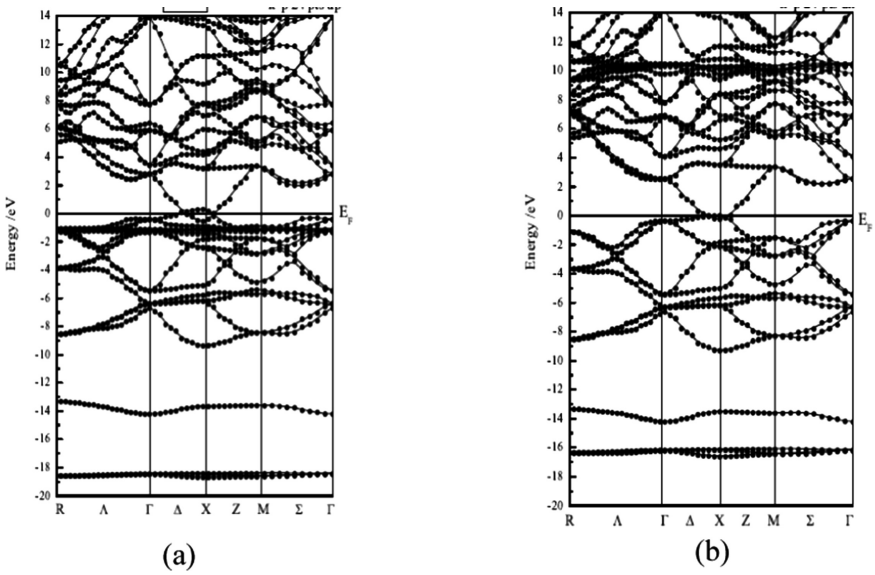


Fig. 7. The spin-polarization band structure of EuB_6 calculated by ABINIT with the k-path generated automatically. (a) for spin-up, (b) for spin-down.

fridaya@uef.fi

energy band structure has more significant overlapping near the X point of the Fermi-plane while spin-down energy band structure approaches 0 near X point of Fermi-plane.

4 Conclusion

Standardization of unit cell input file in electronic structure calculations can reduce primitive cells, get standardized primitive cells of unit cell and k-path of energy band in the Brillouin area, and realize automation throughout the whole "relax-static" high-throughput electronic structure computing process. Generation of standardized k-path ahead of static calculation of electronic energy band can prevent incorrect identification of some band gaps. This module can be conveniently integrated to all kinds of ab initio electronic structure computing software articles. We would like to express our gratitude to Professor Chongyu Wang, Department of Physics, Tsinghua University for helpful discussion. This research is financed by key special fund of "materials genome project"–a national key research and development project (2017YFB0701500).

References

1. Koinuma, H., Takeuchi, I.: Nat. Mater **3**, 429 (2004)
2. Spivack, J.L., Cawse, J.N., Whisenhunt, D.W., Johnson, B.F., Shalyaev, K.V., Male, J., Pressman, E.J., Ofori, J.Y., Solovrichik, G.L., Patel, B.P., Chuck, T.L., Smith, D.J., Jordan, T.M., Brennan, M.R., Kilmer, R.J., Williams, E.D.: App. Cata. A **5**, 254 (2003)
3. Boussie, T.R., Diamond, G.M., Goh, C., Hall, K.A., LaPointe, A.M., Leclerc, M., Lund, C., Murphy, V., Shoemaker, J.A.W., Tracht, U., Turner, H., Zhang, J., Uno, T., Rosen, R.K., Stevens, J.C.: J. Am. Chem. Soc. **125**, 4306 (2003)
4. Potyrailo, R.A., Chisholm, B.J., Morris, W.G., Cawse, J.N., Flanagan, W.P., Hassib, L., Molaison, C.A., Ezbiansky, K., Medford, G., Reitz, H.: J. Comb. Chem. **5**, 472 (2003)
5. Potyrailo, R.A., Takeuchi, I.: Measure. Sci. Tech. **16**, 1 (2005)
6. Chiang, Y.-M., Sadoway, D.R., Aydinol, M.K., Jang, Y.-I., Huang, B., Ceder, G.: Nature **392**, 472 (1998)
7. Jóhannesson, G.H., Bligaard, T., Ruban, A.V., Skriver, H.L., Jacobsen, K.W., Nørskov, J.K.: Phys. Rev. Lett. **88**, 255506 (2002)
8. Stucke, D.P., Crespi, V.H.: Nano Lett. **3**, 1183 (2003)
9. http://www.materialsgenome.org
10. Yu, L., Zunger, A.: Phys. Rev. Lett. **108**, 068701 (2012)
11. Hasan, M.Z., Kane, C.L.: Rev. Mod. Phys. **82**, 3045 (2010)
12. Madsen, G.K.H.: J. Am. Chem. Soc. **128**, 12140 (2006)
13. Maier, W.F., Stöwe, K., Sieg, S.: Angew. Chem. Int. Ed. **46**, 6016 (2007)
14. Taylor, R.H., Curtarolo, S., Hart, G.L.W.: Phys. Rev. B **84**, 084101 (2011)
15. Kolmogorov, A.N., Shah, S., Margine, E.R., Bialon, A.F., Hammerschmidt, T., Drautz, R.: Phys. Rev. Lett. **105**, 217003 (2010)
16. Nakamura, E., Sato, K.: Nat. Mater. **10**, 158 (2011)
17. Yuasa, S., Djayaprawira, D.D.: J. Phys. D **40**, R337 (2007)
18. Tian, N., Zhou, Z.Y., Sun, S.G., Ding, Y., Wang, Z.L.: Science **316**, 732 (2007)
19. Chepulskii, R.V., Curtarolo, S.: ACS Nano **5**, 247 (2011)

fridaya@uef.fi

20. Niu, H., Wang, J., Chen, X., Li, D., Li, Y., Lazar, P., Podloucky, R., Kolmogorov, A.N.: Phys. Rev. B **85**, 144116 (2012)
21. Setyawan, W., Curtarolo, S.: Comput. Mater. Sci. **49**, 299 (2010)
22. Jain, A., Hautier, G., Moore, C.J., Ong, S.P., Fischer, C.C., Mueller, T., Persson, K.A., Ceder, G.: Comput. Mater. Sci. **50**, 2295 (2011)
23. Ong, S.P., Jain, A., Hautier, G., Kang, B., Ceder, G.: Electrochem. Commun. **12**, 427 (2010)
24. Kim, J.C., Moore, C.J., Kang, B., Hautier, G., Jain, A., Ceder, G.: J. Electrochem. Soc. **158**, A309 (2011)
25. Ong, S.P., Chlia, S., Jain, A., Brafman, M., Gunter, D., Ceder, G., Persson, K.A.: Comput. Mater. Sci. **97**, 209 (2015)
26. http://www.qmip.org/qmip.org/Welcome.html
27. Hachmann, J., Olivares-Amaya, R., Atahan-Evrenk, S., Amador-Bedolla, C., Sánchez-Carrera, R.S., Gold-Parker, A., Vogt, L., Brockway, A.M., Aspuru-Guzik, A.: J. Phys. Chem. Lett. **2**, 2241 (2011)
28. Gonze, X., Beuken, J.-M., Caracas, R., Detraux, F., Fuchs, M., Rignanese, G.-M., Sindic, L., Verstaete, M., Zérah, G., Jollet, F.: Comput. Mater. Sci. **25**, 478 (2002)
29. Torrent, M., Jollet, F., Bottin, F., ZérahX, G., Gonze, X.: Comput. Mater. Sci. **42**, 337 (2008)
30. http://www.abinit.org
31. Blöchl, P.: Phys. Rev. B **50**, 17953 (1994)
32. Dresselhaus, M.S., Dresselhaus, G., Jorio, A.: Group Theory: Applications to the Physics of Condensed Matter. Springer, Berlin (2008)
33. Degiorgi, L., Felder, E., Ott, H.R., Sarrao, J.L., Fisk, Z.: Phys. Rev. Lett. **79**, 5134 (1997)
34. Süllow, S., Prasad, I., Aronson, M.C., Bogdanovich, S., Sarrao, J.L., Fisk, Z.: Phys. Rev. B **62**, 11626 (2000)
35. Broderick, S., Ruzicka, B., Degiorgi, L., Ott, H.R.: Phys. Rev. B **65**, 121102(R) (2002)
36. Jun, J., Jiang, B., Lemin, L.: ACTA Phys-Chem. Sinic **24**, 1 (2008)

fridaya@uef.fi

Check for updates

Switched-Current Sampled and Hold Circuit with Digital Noise Cancellation Circuit for 2+2 MASH ΣΔ Modulator

Guo-Ming Sung(✉), Leenendra Chowdary Gunnam, and Shan-Hao Sung

National Taipei University of Technology, 1, Section 3, Zhongxiao East Road, Taipei 10608, Taiwan
gmsung@mail.ntut.edu.tw

Abstract. This paper proposes a switched-current (SI) sampled and hold circuit (SH) for the 2+2 multi-stage noise shaping (MASH) delta-sigma (ΣΔ) modulator, which is fabricated in TSMC 0.18 μm 1P6 M CMOS process. In the proposed design, the input impedance can be reduced by a factor of (1+A) using the modified feedback memory cell to obtain low transmission error and the circuit stability can be improved with the cross-connected common-mode feedforward circuit (CMFF). A differential cross-connected CMFF circuit can speed up a fast-response circuit and bring a stabilized output current. The proposed digital noise-cancellation circuit not only eliminates the higher-order quantisation noise from the last stage but also cancels the lower-order quantisation noise from the earlier stage of the modulator. The simulation results show that the signal-to-noise and distortion ratio (SNDR) is approximately 87.1 dB, and the effective number of bits (ENOBs) is roughly 14.18 bits at a sampling rate of 10.24 MHz with an oversampling ratio (OSR) of 256 and a signal bandwidth of 20 kHz. The designed chip draws 18.19 mW from the supply voltage of 1.8 V and occupies a core area of 0.13 mm^2.

Keywords: Switched-current · Delta-sigma modulator · Sample and hold circuit · Multi-stage noise shaping · Common-mode feedforward circuit

1 Introduction

ΣΔ modulator is a popular type of oversampling analog to digital converter (ADC), which is widely used in various consumer electronics such as high performance data converters, dc-dc converters, wireless sensor networks, and motor controllers [1–3]. Higher-order ΣΔ modulators can be obtained by either single-loop or cascade architectures. The cascade architecture popularly known as MASH architecture is the finest architecture meant for high SNDR with inherent stability [4]. One of the foremost concerned MASH architecture is the noise leakage due to mismatch of analog circuit elements [5], as they are sensitive to process, voltage, and temperature variations [6]. In order to overcome the mismatch problem, careful design and layout are essential for analog circuit.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 396–405, 2019.
https://doi.org/10.1007/978-3-030-22868-2_30

fridaya@uef.fi

Compared with switched-capacitor circuit, the switched-current circuit is more attractive due to its high speed, low supply voltage and small chip area [7]. Besides, switched-current circuit suffered with the transmission error, clock feedthrough error (CFT), non-linearity, and so on [8]. Different techniques had been developed to improve those aforementioned disadvantages of the switched-current SH circuits. The CFT error is minimized by using the current minimum function instead of a sampling switch [9]. The input impedance is reduced by a factor of two and the transmission error can be improved with the help of feedback memory cell (FMC). Furthermore, the CMFF is considered to eliminate the offset current and consequently non-linearity [10].

In the following section, the design of the 2+2 SI ΣΔ modulator with digital noise cancellation circuit is described. Section 2 elucidates the circuit implementation of the proposed 2+2 SI MASH ΣΔ modulator and its relational digital noise cancellation circuit. Simulation results are presented in Sect. 3, and conclusions are drawn in Sect. 4.

2 2+2 MASH Delta-Sigma Modulator with Digital Noise Cancellation Circuit

Figure 1 shows a block diagram of 2−2 MASH ΣΔ modulator with a digital noise cancellation circuit. The proposed modulator is composed of a cascading of two second-order modulators, which are also named as primary modulator and secondary modulator. Both modulators provide an absolute stability in view of low-order modulators (≤ 2) [4]. As shown in Fig. 1, the analog coefficients of the ΣΔ modulator are listed in the following Table 1, which are valued with zero quantization error in primary modulator.

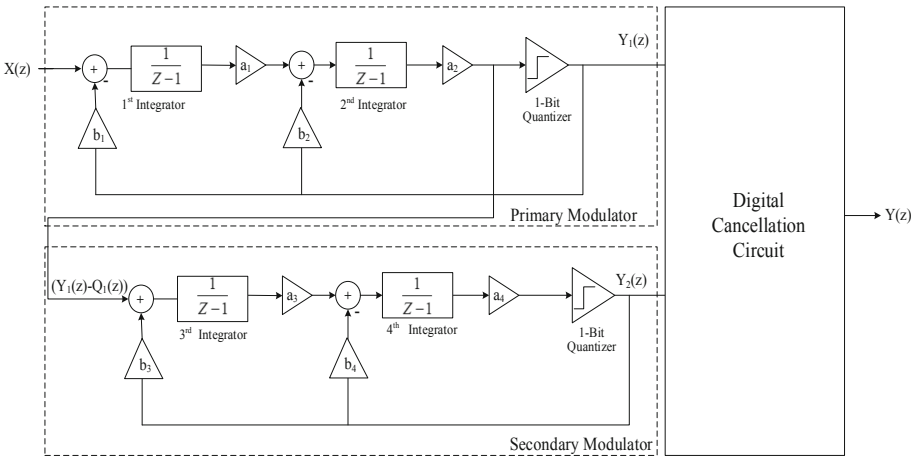


Fig. 1. Block diagram of the 2−2 MASH ΣΔ modulator with a digital cancellation circuit.

fridaya@uef.fi

Table 1. Analog coefficients of the proposed ΣΔ modulator.

a_1	b_1	a_2	b_2	a_3	b_3	a_4	b_4
0.5	1	0.5	1	0.5	1	2	1

The essential components of the designed modulator include integrators, one-bit quantizer, one-bit DAC and non-overlapping clock generator. The performance of the entire modulator is determined by two cascading integrators, primary modulator and secondary modulator. Each modulator is composed of two integrators; and that two SHs, SH-1 and SH-2, are considered to establish an integrator, which is also called a delay unit (z^{-1}). Note that the non-overlapping clock is designed to control the operation of SH, as shown in Fig. 2.

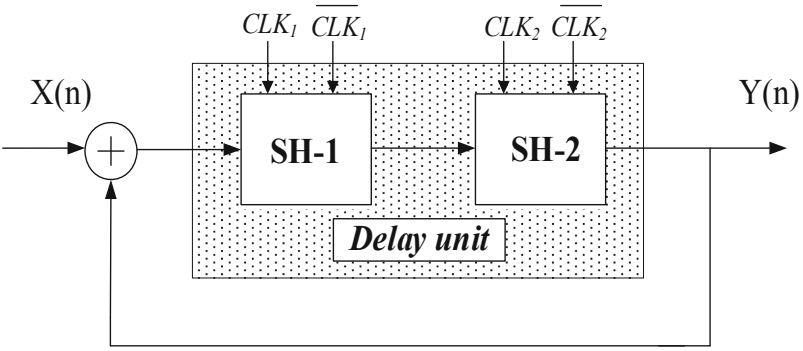


Fig. 2. Proposed integrator with two SHs, SH-1 and SH-2. The non-overlapping clock CLK is designed to control the operation of SH.

2.1 Proposed Switched-Current Feedback Memory Cell

A traditional SI FMC is published in [10] with a low input impedance of $(2g_{m1})^{-1}$. Even though this FMC performs with low input impedance and small transmission error, the aforementioned non-ideal effects affect the performance of the SI FMC seriously. Figure 3 depicts the proposed modified SI FMC and its small-signal model. According to the small-signal model, the current transfer function is given by

$$A_I = \frac{I_{out}}{I_{in}} = \frac{g_{m5}}{g_{m1} + (g_{m6}g_{m4}A/g_{m7})} \tag{1}$$

$$A = \frac{V_1}{V_X} = \frac{1 - g_{m2}R_f}{1 + g'_{m3}R_f} \tag{2}$$

where g_{mi} is the transconductance of i^{th} MOSFET, g'_{m3} is equal to $[g_{m3} + (r_{o2})^{-1} + (r_{o3})^{-1}]$, and R_f is the electronically controlled feedback resistance. A is the transfer

fridaya@uef.fi

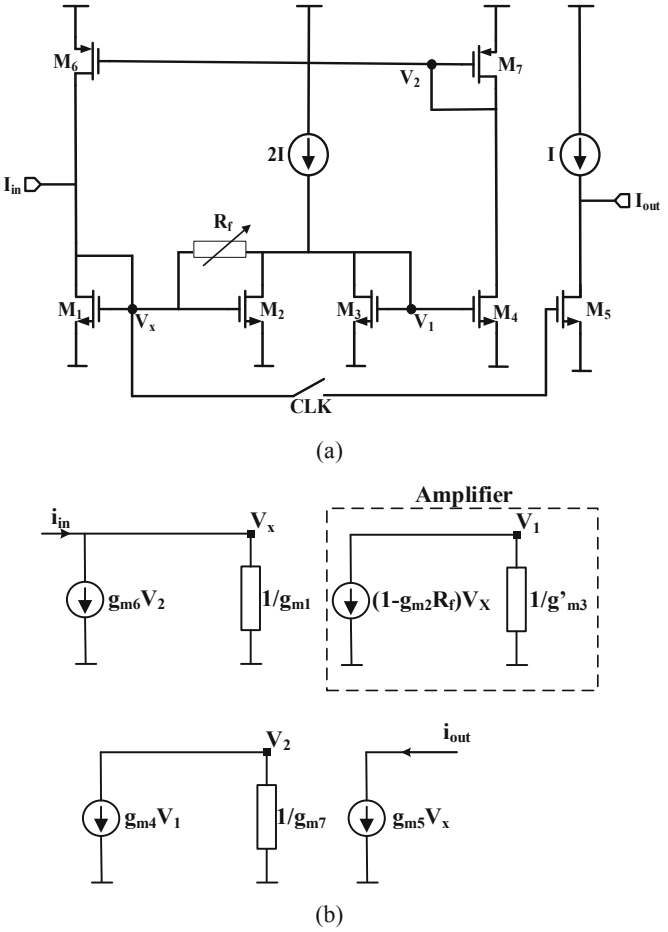


Fig. 3. Proposed modified SI FMC and its small-signal model. (a) Modified FMC, (b) Small-signal model.

function of V_1 with respect to V_X. Note that the output current i_{out} is equal to $g_{m5}V_X$. Then the input impedance R_{in} can be given by

$$R_{in} = \frac{V_X}{I_{in}} = \frac{1}{g_{m1} + (g_{m6}g_{m4}A/g_{m7})} \tag{3}$$

By properly adjusting the metric ratio, we obtain $g_{m1} = g_{m6}$ and $g_{m4} = g_{m7}$. Then the input impedance of the proposed modified FMC can be simplified as

$$R_{in} = \frac{1}{g_{m1}(1+A)} \tag{4}$$

fridaya@uef.fi

As shown in (4), the input impedance is decreased by a factor of (1+A) and the transmission error is greatly reduced compared to the traditional FMC. According to the modified SI FMC, a new current-mode sample-and-hold circuit is proposed with a modified differential SI FMC, a coupled differential replicate (CDR) with CMFF, and a differential cross-connected CMFF circuit. Figure 4 shows the proposed current-mode SH with two modified FMCs, a CDR with CMFF circuit, and a cross-connected CMFF circuit. Cumulating the CDR circuit with modified FMC and differential cross-connected CMFF circuit, the amendment speeds up a fast-response circuit and brings a stabilized output current.

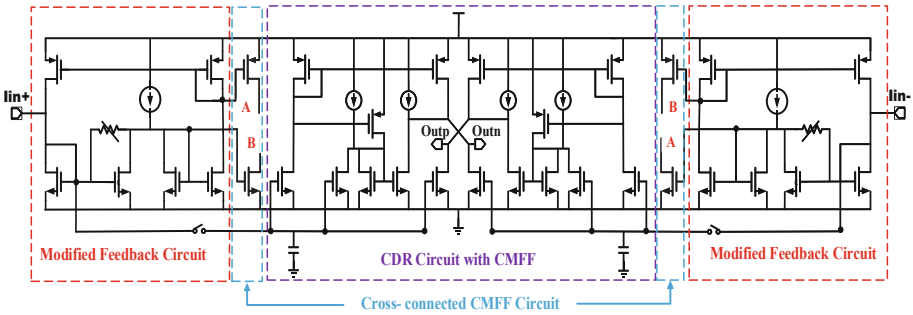


Fig. 4. Proposed current-mode SH circuit with two modified FMCs, a CDR circuit with CMFF, and a cross-connected CMFF circuit.

2.2 Proposed Digital Noise Cancellation Circuit with Karnaugh Map

Figure 5 shows the mathematical model of the digital noise cancellation circuit (DNCC). The output Y(z) can be expressed as follows [11].

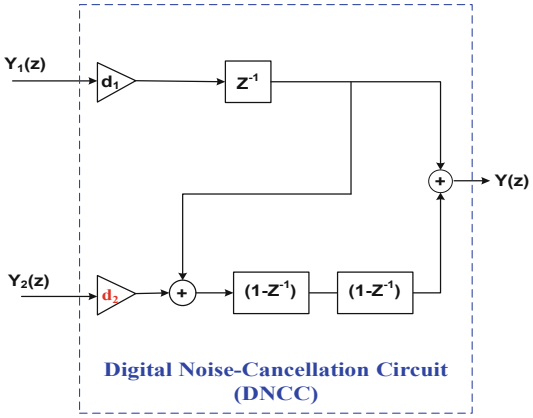


Fig. 5. Proposed mathematical model of digital noise cancellation circuit (DNCC).

fridaya@uef.fi

$$Y(z) = d_1 z^{-1}\left[1 - \left(1 - z^{-1}\right)^2\right] Y_1(z) + d_2\left(1 - z^{-1}\right)^2 Y_2(z) \tag{5}$$

where those coefficients, d_1 and d_2, can be expressed with 1.0 and 4.0, respectively. Then Eq. (5) can be rewritten as

$$\begin{aligned} \mathrm{Y}(z) &= \left(2z^{-2} - z^{-3}\right) Y_1(z) + \left(4 - 8z^{-1} + 4z^{-2}\right) Y_2(z) \\ &= (2A - B) + (4C - 8D + 4E) = F + G \end{aligned} \tag{6}$$

where A, B, C, D, and E are $z^{-2}Y_1(z)$, $z^{-3}Y_1(z)$, $Y_2(z)$, $z^{-1}Y_2(z)$, and $z^{-2}Y_2(z)$, respectively. The output numerals are between −19 and +19, which is a function of inputs, A, B, C, D, and E. A 6-bit digital word, from S_1 to S_6 in 2's complement form, is used to represent the output numerals. Where S_6 is the most significant bit (MSB) and S_1 is the least significant bit (LSB) of the output digital word. As shown in (6), five delay blocks are required to implement the DNCC. One-unit delay (z^{-1}) can be produced using a pair of D-type flip-flops as shown in Fig. 6; Ck_{1d} and Ck_{2d} are nonoverlapping clock signals. Note that the output is represented as a digital word with minimum and maximum values of −19 and +19 in 2's complement form with 6 bits, S_6 to S_1. The output digital bits can be written as

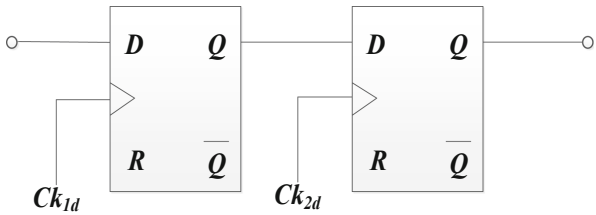


Fig. 6. One-unit delay circuit (z^{-1}) using D-type flip-flops.

$$S_6 = \sum m\,(0, 2, 3, 6, 7, 8, 10, 11, 14, 15, 18, 19, 22, 26, 27, 30) \tag{7}$$

$$S_5 = \sum m\,(0, 3, 6, 7, 8, 11, 14, 15, 18, 19, 21, 22, 26, 27, 29, 30) \tag{8}$$

$$S_4 = \sum m\,(0, 2, 5, 7, 8, 10, 13, 15, 17, 19, 20, 22, 25, 27, 28, 30) \tag{9}$$

$$S_3 = \sum m\,(0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15) \tag{10}$$

$$S_2 = \sum m\,(0, 1, 2, 3, 4, 5, 6, 7, 16, 17, 18, 19, 20, 21, 22, 23) \tag{11}$$

$$S_1 = 1 \tag{12}$$

fridaya@uef.fi

3 Simulation Results

In Fig. 7, we show the simulated percentage of transmission error of a feedback memory cell and modified feedback memory cell, indicates the modified feedback memory cell offers around 0.1% of transmission error for +35μA. Furthermore, the error is virtually linear fashion, while feedback memory cell offers oscillatory. Thus, the modified feedback memory cell provides better stabilization than that of feedback memory cell. A Transfer characteristics of differential input current and output current as shown in Fig. 8, there is a linear relationship between the input current and output current. Figure 9 show the simulated current-mode sample-and-hold circuit with sinusoidal input signal, when the frequency of 500 kHz and amplitude of 25 μA applied at a sampling frequency of 10.24 MHz and an oversampling ratio of 256. These simulated waveforms show that the accuracy and feasibleness of the sample and hold circuit are decent.

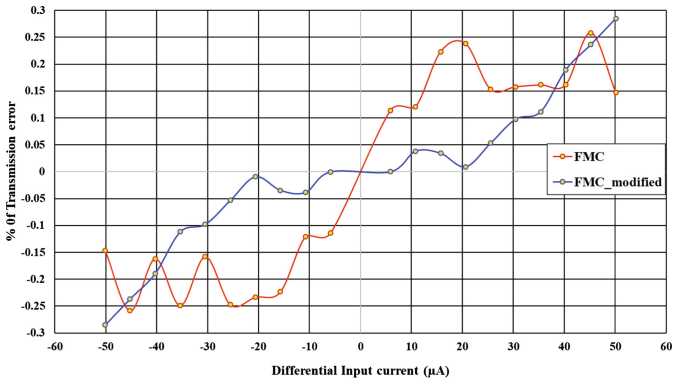


Fig. 7. Simulated percentage of Transmission error of a Feedback memory cell and modified feedback memory cell.

Figure 10 illustrates the simulated input-output waveforms of the proposed digital noise cancellation circuit (DNCC), which is implemented with proposed algorithm. As shown in Eq. (6), the input $Y_1(z)$ will be delayed 2 units (2T) and 3 units (3T) to generate A and B, respectively; and that the input $Y_2(z)$ is copied to generate C (0T) and is delayed 1 unit (1T) and 2 units (2T) to generate D and E, respectively. For certain inputs such as A = −1, B = +1, C = −1, D = −1, and E = +1, Y = (2A − B) + (4C − 8D + 4E) = 5, and its 2's complement value (S6 to S1) was 000101 at the output (Y).

Figure 11 shows the simulation results of the proposed 2+2 switched-current delta-sigma modulator with digital noise cancellation circuit. The signal to noise and distortion ratio (SNDR), effective number of bits (ENOB), and power dissipation are 87.1 dB, 14.18 bits, and 18.19 mW, respectively.

fridaya@uef.fi

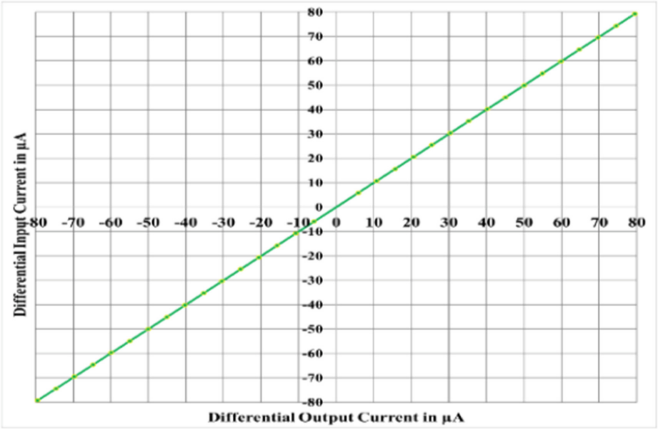


Fig. 8. Transfer characteristics of differential input current and output current.

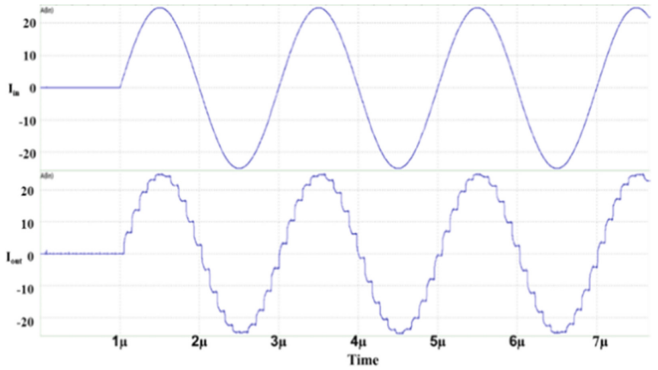


Fig. 9. Simulated current-mode sample-and-hold circuit with sinusoidal input signal.

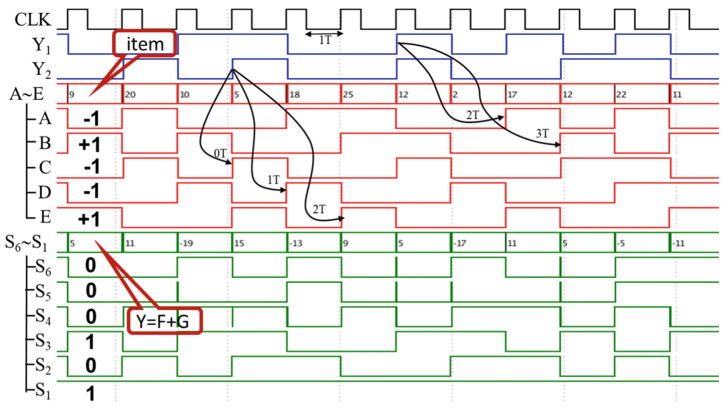


Fig. 10. Simulated waveforms of the designed DNCC using Kamaugh map.

fridaya@uef.fi

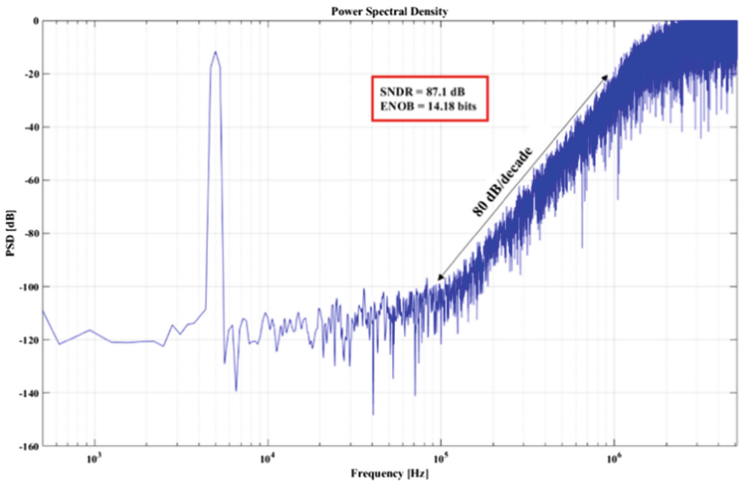


Fig. 11. Output spectrum of the proposed 2+2 switched-current delta-sigma modulator with digital noise cancellation circuit.

4 Conclusions

This paper presents the design of current-mode SH with DNCC for 2+2 SI MASH ΣΔ modulator, which is fabricated in TSMC 0.18 μm 1P6 M CMOS process. The input impedance of the modified FMC is reduced by (1+A) times and the transmission error was diminished particularly intended for low input current. Not only the CDR circuit with modified FMC but also the differential cross-connected CMFF circuit could speed up the circuit response and generate a stabilized output current. Besides, the proposed DNCC not only eliminates the higher-order quantisation noise from the last stage but also cancels the lower-order quantisation noise from the earlier stage of the modulator. By the way, the DNCC is easy to be implemented with Karnaugh map using five delay blocks. A one-unit delay (z^{-1}) can be produced with a pair of D-type flip-flops using two nonoverlapping clocks to have a small delay. The designed chip draws 18.19 mW at the supply voltage of 1.8 V and occupies a core area of 0.359 × 0.370 mm^2.

References

1. Mooney, J., Halton, M., Iordanov, P., O'Brien, V.: Dithered multi-bit sigma-delta modulator based DPWM for dc-dc converters. In: IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 2835–2839. Charlotte, NC (2015)
2. Lai, W., Huang, J., Huang, S.: Sigma-delta modulator with near infrared charging for IOT wireless power transfer application. In: 11th IEEE International Conference on Anti-counterfeiting, Security, and Identification (ASID), pp. 134–137. Xiamen, China (2017)
3. Sung, G.M., Yu, C.P., Hung, t.W., Hsieh, H.Y.: Mixed-mode chip implementation of digital space SVPWM with simplified-cpu and 12-bit 2.56 Ms/s switched-current delta-sigma ADC in motor drive. IEEE Trans. Power Electron. **27**(2), 916–930 (2012)

fridaya@uef.fi

4. Rosa, J.M.: Sigma-delta modulators: tutorial overview, design guide, and state-of-the-art survey. IEEE Trans. Circuits Syst. I **5**(1), 1–21 (2011)
5. Rio, R.: CMOS Cascade Modulators for Sensors and Telecom: Error Analysis and Practical Design. Springer, New York (2006)
6. Yang, W., Lin, Y., Wang, C., Chang, K., Chen, C., Lo, Y.: Analysis and design considerations of static CMOS logics under process, voltage and temperature variation in UMC 0.18 μm CMOS process. In: International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), pp. 57–61. Nusa Dua (2015)
7. Sung, G.M., Yu, C.P., Yao, D.A.: A comparison of second-order sigma-delta modulator between switched-capacitor and switched-current techniques. In: IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), pp. 1172–1175. Macao, China (2008)
8. Jonsson, B.E.: Switched-current signal processing and A/D conversion circuits design and implementation. The Kluwer International Series in Engineering and Computer Science (2000)
9. Luangpol, A., Petchmaneelumka, W., Kamsri, T., Riewruja, V.: A current signal CMOS sample-and-hold circuit. In: International Conference on Control, Automation and Systems, pp. 171–174. Seoul, Korea (2007)
10. Sung, G.M., Gunnam, L.C., Lin, W.S., Lai, Y.T.: A third-order multibit switched-current delta-sigma modulator with switched-capacitor flash ADC and IDWA. IEICE Trans. Electron. E100.C, 1–10 (2017)
11. Gunnam, L.C., Sung, G.M., Weng, L.W.: 2+1-order switched-current MASH delta-sigma ADC with the digital cancellation circuit. In: International Conference on Applied System Innovation (ICASI), pp. 1801–1804. Sapporo, Japan (2017)

fridaya@uef.fi

Towards the Identification of Context in 5G Infrastructures

Chrysostomos Symvoulidis(✉), Ilias Tsoumas, and Dimosthenis Kyriazis

University of Piraeus, 18532 Piraeus, Attica, Greece
{simvoul,itsoum,dimos}@unipi.gr

Abstract. The evolution of communication networks, bringing the fifth generation (5G) of mobile communications in the foreground, gives the vertical industries opportunities that were not possible until now. Flexible network and computing infrastructure management can be achieved, hence bringing more freedom to the service providers, to maximize the performance of their computing resources. Still, challenges regarding the orchestration of these resources may arise. For this reason, an engine that can recognize possible factors that might affect the use of these resources and come up with solutions when needed in real-time, is required. In this paper, we present a novel Complex Event Processing engine that is enriched with Machine Learning capabilities in order to be fully adaptive to its environment, as a solution for monitoring application components deployed in 5G infrastructures. The proposed engine utilizes Incremental DBSCAN to identify the normal behavior of the deployed services and adjust the rules accordingly.

Keywords: 5G · Complex Event Processing · Dynamic complex event processing · Anomaly detection · Event identification · Context awareness

1 Introduction

Unlike previous generations, 5G will be more than just a mobile network. According to Soldani et al. [2] and Moyse [3] 5G will enable a reliable and efficient way of deploying several services in multiple distant facilities. The utilization of such diverse infrastructures should lead us to considerable advantages. Numerous paradigms and methodologies such as Software-as-a-Service (SaaS) and Infrastructure-as-a-Service (IaaS) can be now coupled [3], thus leading to greater performance of the services per se. Services will be substantially more efficient, since the service providers will have the ability to orchestrate them, even in the cases where they are deployed in heterogeneous infrastructures. Monitoring would present a more complete image to the service providers, showing aggregated results on all their resources on all infrastructures [4,5,7].

As already mentioned, 5G is an emerging technology that will bring many advantages both for the end-users, and the service providers, due to its increased

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 406–418, 2019.
https://doi.org/10.1007/978-3-030-22868-2_31

fridaya@uef.fi

bandwidth, high throughput and particularly high reliability [1]. But in order to take full advantage of a 5G infrastructure's capabilities and its functionalities, an intelligent way to orchestrate both the network and the infrastructure is necessary. Anomaly detection in such diverse environments consisting of several factors is a demanding process, hence an intelligent and interactive way to bypass such issues is required.

Complex Event Processing (CEP) is a technology, mainly used for the analysis and the process series of correlated events, using a set of techniques and tools [10]. CEP is a perfect fit for the aforementioned issues, since it can be used as a reactive alternative to process incoming information in order to identify complex states and propose solutions to problems instantaneously.

Given the fact that in 5G infrastructures anomaly detection is a complex process, common anomaly detection solutions might not work effectively. Hence, we present a novel CEP engine for anomaly detection for such environments. This engine focuses mainly on the identification of events that affect the deployment and the performance of both the applications and the network per se. The engine has been developed using the Drools Fusion framework [12], a tool created by RedHat JBoss with respect to operating CEP jobs. The proposed approach also includes an intelligent mechanism that can make the CEP engine more adaptive to the current environment, via the exploitation of Machine Learning techniques, in order to provide more customized results based on its environment every time, thus making the CEP engine context aware [6].

The rest of the paper is organized as follows: Sect. 2 focuses on the delivery of the state of the art in the area of network anomaly focusing mainly in the new era networks, and the CEP architectures that are currently being used. Section 3 depicts the overall architecture of the proposed engine. Section 4 regards the experimental evaluation of the engine and depicts the results. Finally, Sect. 5 concludes the paper and suggests future work that can be done based on the current engine.

2 Related Work

This section provides an overview of the current State of the Art in the CEP and Anomaly Detection areas.

2.1 Complex Event Processing

As already declared, CEP is a technique that is primarily used for the identification of certain situations (events), through the process of streams of information from various sources. Its purpose is to identify such meaningful events and respond quickly. CEP is part of the Rule Engines area [13]. CEP can be used along other technologies, such as Business Process Management (BPM) or Business Activity monitoring, but its distinguished due to its agility [14] in comparison with the rest. CEP is widely used in a great amount of information

fridaya@uef.fi

systems, including network monitoring and anomaly detection, business process automation (BPA) and business intelligence (BI), among others. Complex Event Processors can identify events immediately, thus speeding up the decision-making process significantly. CEP engines can identify events from numerous data streams that are often correlated, or possible event patterns and situations, as well as alert when necessary and trigger actions that have to be done [11]. An integral element of the CEP engines is the exploitation of potential incidents, namely the *events*, that are in general, things that happen in a certain environment, or changes of state in a current situation. Yet, the creation of events is not a simple process, on the ground that complex events can exist through the combination of several individual events.

Another highly important term of the CEP engines regards the *Sliding Windows*. The Sliding Windows are a way of grouping (correlating) complex events together based on some factors. A widely used technique to combine such events is done through the notion of time (e.g. a number of events are occurred in a specified time window), or by the number of times they are occurred (e.g. an event X has occurred Y times).

One of the first known systems in the CEP area is Aurora [15]. Aurora supported a number of basic operations including Filtering, Resampling and Tumbling that could be applied into streams of information in a window time, or filters such as joins, maps and GroupBys that could be applied into tuples of data, once a time. Borealis [15] came as the second generation of Aurora that could operate as distributed stream processing engine. Through Borealis dynamical revision of query results and modification of the query specifications was achieved.

Gyllstrom et al. [16] proposed SASE+ witch was eventually developed in the University of Massachusets, a new approach which described a temporal sequence of events that are correlated based on their attribute values. Endler et al. [22] presented a semantic model for streaming data as an extension of a CEP engine in order to provide Semantic reasoning. Brenna et al. [18] in their work presented CAYUGA, a CEP engine that uses non-finite automation to match event patterns to queries, through an SQL-like language. Bhargavi et al. [17], proposed an adaptive CEP engine that can update its Knowledge Base at run-time in order to make it more dynamic and interactive. Petersen et al. [21] in their work have presented a CEP rule alteration engine using Online Learning techniques, thus making the engine more efficient over time, since the rules would adapt better to the current state of the environment.

In terms of production CEP tools, WSO2 created a CEP engine called WSO2 Complex Event Processor [8,9]. In its core, the WSO2 CEP uses the Siddhi query engine [19] which describes each event as a tuple. It provides a lot of features including time windows, filters, joins, event patterns, event partitions and it focuses mostly on the real-time analyzation of the identified events. A robust and versatile known CEP engine is Esper [20]. Esper describes an attribute of an event as a composition of other events, divides each event in fragments and offers domain-specific information per event. It uses an extension of SQL, the

fridaya@uef.fi

Esper Query Language (EQL), which offers filters such as sliding windows and patterns over the streams of information. Other systems that could operate CEP processes include Cordies [38], LogCEP [37], and SPA [32], along with Apache Flink [34,45], Apache Storm [33,44] and Apache Spark [35,43]; which regard general purpose frameworks, that could also work as CEP engines.

2.2 Anomaly Detection

Anomaly detection is used broadly in communication networks. Depending on the situation there are numerous ways to analyze and detect anomalies [24]; classification based, clustering-based, or using statistical methods and information theory. Deljac et al. in [23,24] used a Bayesian network for the outage detection of telecommunication networks. Clustering-based techniques regard unsupervised learning engines, mainly for unlabeled data. Ahmed et al. in [25] used regular clustering techniques to cluster the data and detect outliers to improve the accuracy of anomaly detection engines. Another interesting approach regards the one presented in [26] where Ester et al. used the DBSCAN algorithm to discover noise in databases. Statistical-oriented methods are widely used in anomaly detection, including [27], that suggests using signal processing techniques in order to identify anomalies in IP networks. Shyu et al. [28] proposed a novel paradigm to handle and analyze multidimensional network traffic datasets that could identify whether an instance in the dataset regarded an anomaly or not.

Yang et al. [29] presented a rule-based Intrusion Detection System (IDS) for Supervisory Control and Data Acquisition (SCADA) networks, using an in-depth analysis and deep packet inspection techniques. Balabine and Velednitsky in [31] published a patent that uses Support Vector Machines (SVM) for network traffic anomaly detection. Other than SVMs, Neural networks are also used for such purposes, like the one presented by Poojitha et al. in [30] where they tried to detect anomalies in a given dataset with labeled information.

Taking under consideration all the aforementioned approaches, an issue that has not been answered yet, regards the ability of a given CEP engine to be fully adaptive to the environment it concerns and provide customized insights. For this reason, our approach regards a CEP engine in order to combine these two areas and provide better results in diverse ecosystems such as a 5G infrastructure. Our approach aims at providing an engine that can be adaptive and support the deployment of applications and services in such environments trouble-free with customized events, taking under consideration the current conditions.

3 Overall Architecture

This section describes the architecture of the developed engine. More specifically, it describes the mechanisms responsible for the collection of the data from the various components, along with the way the data are fed to the CEP engine, how this engine matches the real-time data with the rules in predefined the Knowledge Base (KB), and finally what is our proposed solution to the creation of an adaptive CEP engine.

fridaya@uef.fi

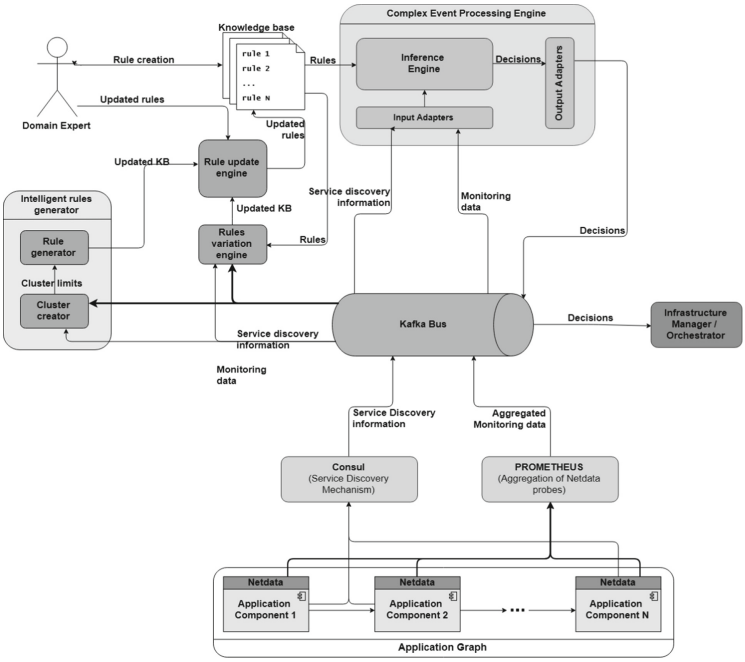


Fig. 1. CEP Engine overall architecture

The proposed engine is divided in six major steps which are also depicted in Fig. 1: (i) the creation of the rules that are matched to the events by the domain expert, (ii) the collection of the data by the various monitoring engines, (iii) the data cleaning part along with the transformation of the data into CEP-utilizable information, (vi) the event processing part where the streams of information are matched to some rules stored in the KB, (v) the intelligent rules generator; meaning the engine responsible for adapting to the environment and update the rules accordingly, and finally (vi) the propagation of the decisions made by the CEP engine to the interested mechanisms. These steps are described in more detail in the sections below.

3.1 Creation of the Events

The whole process starts with the creation of the events by a domain expert, in 5G networks and infrastructures. Events related to the performance of the network and the resource infrastructures. Most of these events have to do with the resource usage of the computing resources and the network usage. These events are then "translated" into rules and are used for the identification of the events by the engine. This translation regards the creation of rules based on the decisions of the domain expert.

fridaya@uef.fi

3.2 Data Collection

The process continues with the collection of the data by the various monitoring engines. The architecture suggests using Netdata [39] probes in each application component, that collect information related to the usage of the machine, including CPU, Memory and Disk usage percentage, currently running applications, system up-time, IPv4/IPv6 received, sent and aborted packets and Kbps, etc. These application components regard Virtual Machines, Containers, or Hardware resources.

As depicted in Fig. 1, the collected data are scraped by the Prometheus monitoring engine [40]. Its job is to provide aggregated results with respect to the whole infrastructure and specific results per deployed application component, in order for both the CEP engine and the Intelligent rules generator to have an overall image of the current state of the facility.

Additionally, data are also collected by the Consul [41] service discovery mechanism. Its purpose is to collect valuable information related to the services, including health checking, service communication, check whether the APIs of the services are functional, etc.

All the data that are collected by both the Prometheus and Consul services are then sent to a Kafka Bus [42], as the proposed Publish/Subscribe (Pub-Sub) platform. From there anyone that is interested in retrieving that data, can subscribe to each topic and collect them.

3.3 Data Cleaning

As soon as the data are collected from the Prometheus engine and the Consul mechanism from the input adapters of the CEP engine, the data cleaning process is instantiated. This process regards mainly the data collected from the Prometheus system. Netdata and Prometheus collect a great number of information that only a just a small amount of data are used by the CEP engine. In addition, the data are received in a form that the CEP engine cannot handle, so they are transformed into CEP engine - utilizable information, and are finally fed to the engine.

3.4 Complex Event Processing

After the collection of the data collection and data cleaning processes are complete, the CEP engine is now able to initiate the event identification process. Drools Fusion [12] is the preferred CEP engine for the reason that it has several advantages on the way the events are handled, its ease of creation rules and has numerous useful features, including Sliding Time and Length Windows, support of both Stream and Batch Processing and several ways of grouping events together apart from Sliding Windows.

Before the collection of the data, the engine collects the rules from the KB. These rules are the representation of the events and are in the form of WHEN ... THEN ..., as also presented in Fig. 2.

fridaya@uef.fi

After the streaming data are collected and cleaned they are forwarded to the CEP engine. Here starts the process of the identification of complex events. The engine collects the rules from the KB and when the incoming data match a rule, that specific rule gets triggered.

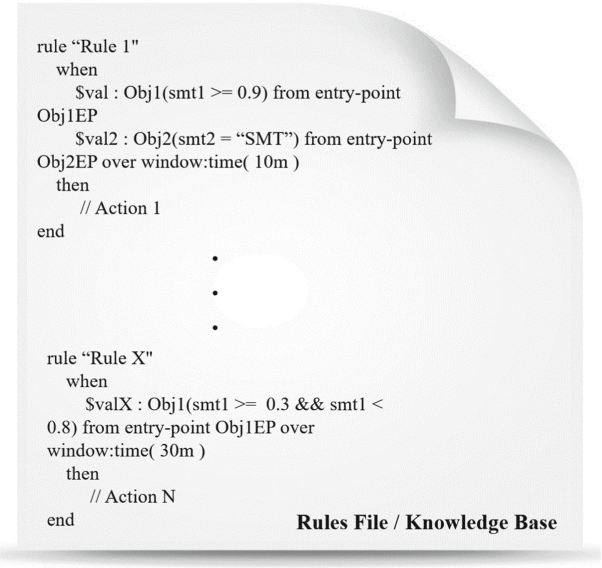


Fig. 2. Rules file example

3.5 Personalized Events on the Fly

This is the most important part of the engine. The proposed engine includes a *dual mode of operations* regarding the update of the KB.

The first mode regards the *Rules variation engine*. Its purpose is to check the numerical difference between the streaming data and the expressions in the rules (i.e the expression in the WHEN part, Fig. 2, that triggers the rule). If that difference is greater that a given amount, the rules are updated in order to reduce that void, in cases where it is considered necessary. Also, through this engine, the domain expert can update the KB personally, by modifying, adding, or deleting rules without the engine having to shut-down or restart. Then a listener notifies the inference engine, in order to act based on the updated KB.

The second mode, regards a more *sophisticated* way of updating the KB in a more personalized manner, with respect to the deployed application graphs, using Machine Learning techniques. The intention of this module, is to identify the normal behavior of the deployed applications and services, based on the data streams provided by the monitoring and the service discovery mechanisms, in order to create customized rules. In such diverse environments like a 5G infrastructure, even the domain experts cannot be fully aware of the current state

fridaya@uef.fi

of the infrastructure, so the events are more generalized. Hence, the existence of a mechanism that can adapt on the current state of the environment is necessary.

Therefore, we developed an engine that uses *Incremental Density Based Spatial Clustering of Applications with Noise (Incremental DBSCAN)* [36], an enhancement of the known DBSCAN algorithm that can work with continuously incoming data.

Using DBSCAN the engine can identify the normal behavior of the application graph and the corresponding application components and adjust the rules accordingly, and group it into clusters. In more detail, the *Intelligent rules generator* component, as depicted in Fig. 1, is comprised of two major components; the *Cluster creator* and the *Rule generator*.

The Cluster creator collects the data from the Kafka bus and with the use of Incremental DBSCAN, it creates clusters that present the behavior of the networking and computing resources of the application components. It then discovers the limits of the major cluster that regards the normal behavior of the component. The cluster that is considered to be the one that depicts the normal behavior is the one with the most core elements in it.

The limits of the clusters are then forwarded to the Rule Generator. Its job is to update the rules in the KB the same way the Rules Variation engine does, but it is now based on the outcomes of the Cluster creator. After the update of the KB is complete, the listener notifies the CEP engine so that it can now act based on the updated KB.

3.6 Propagation of the Decisions Made by the CEP Engine

When an event is identified, a rule is triggered. After such a rule is triggered, the process of the communication of the outcomes begins. The outcome of every rule which is triggered is forwarded to the Kafka Bus that is described above and the corresponding services and engines that are subscribed to the topic related to the CEP, in order to get notified.

4 Experimental Evaluation

This section presents the experiments that the developed CEP is put under. Due to the absence of a real 5G infrastructure, an experiment was executed in a comparable environment.

4.1 Working Environments

For the testing of the developed CEP engine we used two Virtual Machines (VM) that are comprised of 2 vCPUs, 8 GB of memory and 20 GB of Hard Disk. They both ran the Ubuntu 14.04 operating system. During the test we put these VMs under pressure using benchmarks. Another VM with the same computing

fridaya@uef.fi

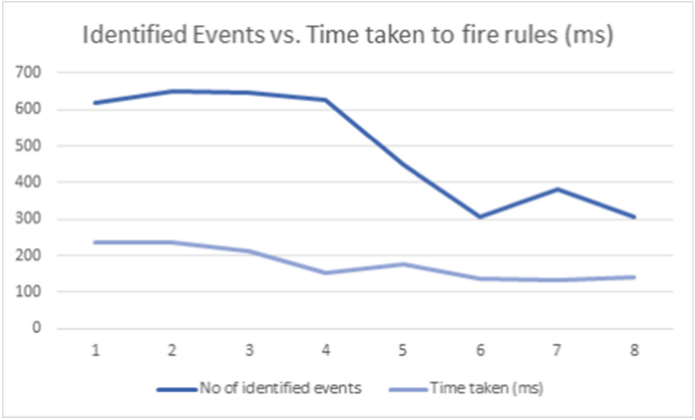


Fig. 3. Number of Identified events vs. Time taken to fire rules (ms) in a 8-h period

resources had the responsibility to collect the monitoring data and the service discovery information from Prometheus and Consul and forward them to the Kafka bus.

The CEP engine, as well as the rule generator engine ran on a desktop PC with Windows 10 operating system, with Intel dual-core i3 CPU, 8 GB of RAM and 500 GB of Hard Disk Drive.

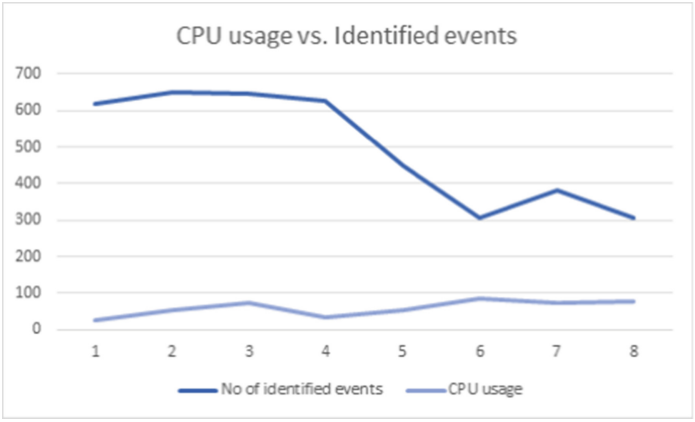


Fig. 4. CPU usage vs. Identified events in a 8-h period

4.2 Results

For the evaluation part, we are based on metrics that can describe the overall performance of the engine. More specifically, these metrics regard the number of

fridaya@uef.fi

the rules that are fired, the total time taken to fire the corresponding rules from the time the data arrived in the engine, and if that changed when the amount of the rules triggered was increased.

The CEP engine presented indeed the current situation of the resources with minimum delay, since the events were being identified in less than 300 ms from the moment the data were received in the engine, regardless of the number of the number of the rules that were being triggered at the same time, as also depicted in Fig. 3.

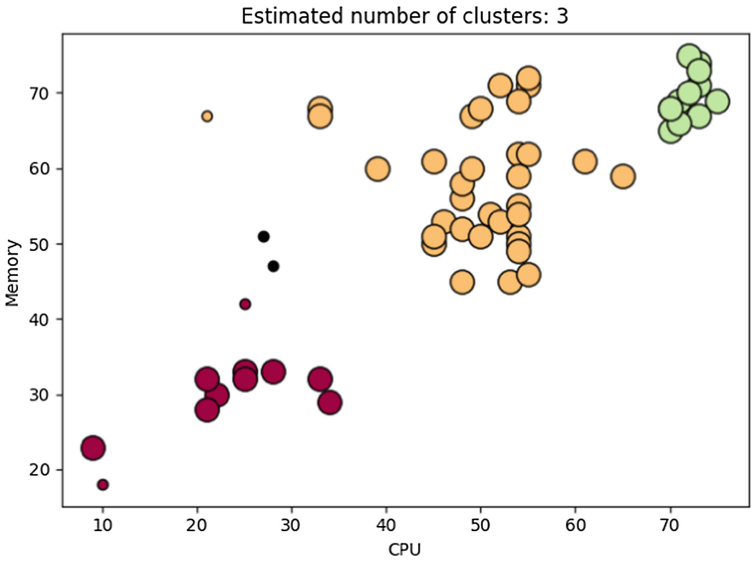


Fig. 5. Clusters created by the Intelligent rules generator engine

In Fig. 4 the number of the rules that were being triggered is depicted, in comparison to the CPU usage. At first, the number of the identified events was quite large. This was due to the fact that the rules that were used at first were the predefined ones. After the Rule generator engine was activated, the number of the identified events dropped significantly. Still though, events that were essential and described a change of state in the environment were identified.

Figure 5 presents the clusters that were created by the DBSCAN algorithm. The cluster with the orange color was identified as the normal behavior of the deployed VMs. The limits of the cluster (i.e minimum CPU, Memory, Disk usage, etc.) were forwarded to the Rule generator, in order to update the KB. When the process was complete the CEP engine identified as an event anything outside these boundaries. This led to the creation of events that were not general, but were adapted to the specific situation.

fridaya@uef.fi

5 Conclusion and Future Work

The identification of factors that affect the use of such diverse infrastructures and networks, such as 5G is indeed a difficult and complex process. In this paper, we have presented a novel CEP engine that can overcome issues, when dealing with multidimensional data and be considered a trustworthy alternative for a more interactive network anomaly detection. The use of the Incremental DBSCAN algorithm makes the current engine more adaptive and the events identified are customized to the current situation.

It is within our future plans to extend the present approach, by making the Incremental DBSCAN more precise and continue our research in Machine Learning and CEP engines in order to make the engine more generic and flexible to changes in its environment, thus being capable of providing insights to other areas of interest including Cloud computing.

Acknowledgment. This work has received funding from the European Unions Horizon 2020 research and innovation program under grant agreement No 761898 project MATILDA.

References

1. Suriano, D.: The Future of Networking Is 5G: Businesses Must Prepare Now, 23 September 2018. https://www.forbes.com/sites/oracle/2018/09/24/the-future-of-networking-is-5g-businesses-must-prepare-now
2. Soldani, D., Manzalini, A.: A 5G infrastructure for "anything-as-a-service". J. Telecommun. Syst. Manag. **3**(2), 1 (2014)
3. Moyse, I.: 5 Reasons for a Multi-Cloud Infrastructure—Dyn Blog (n.d.). https://dyn.com/blog/5-reasons-for-a-multi-cloud-infrastructure/
4. Uittenbogaard, T.: What are the advantages of a multi-site/multi-domain solution and who's to benefit from it? (2015). https://oneshoe.com/news/what-are-advantages-multi-sitemulti-domain-solution-and-whos-benefit-it
5. Hume, A.C., Al-Hazmi, Y., Belter, B., Campowsky, K., Carril, L. M., Carrozzo, G., Engen, V., Garcia-Perez, D., Ponsati, J.J., Kubert, R., Rohr, C., van Seghbroeck, G., Liang, Y.: Bonfire: a multi-cloud test facility for internet of services experimentation. In: International Conference on Testbeds and Research Infrastructures, pp. 81–96. Springer, Heidelberg (2012)
6. Schilit, B., Adams N., Want R.: Context-aware computing applications. In: Workshop on Mobile Computing Systems and Applications, pp. 85–90. IEEE, Santa Cruz (1994)
7. Baldin, I., Chase, J., Xin, Y., Mandal, A., Ruth, P., Castillo, C., Orlikowski, V., Heermann, C., Mills, J.: ExoGENI: a multi-domain infrastructure-as-a-service testbed. In: The GENI Book, pp. 279–315. Springer, Cham (2016)
8. Perera, S., Sriskandarajah, S., Vivekanandalingam, M., Fremantle, P., Weerawarana, S.: Solving the grand challenge using an opensource CEP engine. In: Proceedings of the 8th ACM International Conference on Distributed Event-Based Systems, pp. 288–293. ACM, May 2014
9. WSO2 Complex Event Processor. https://wso2.com/products/complex-event-processor

fridaya@uef.fi

10. Luckham, D.: The Power of Events, vol. 204. Addison-Wesley, Reading (2002)
11. Garcia, J.: A complex event processing system for monitoring manufacturing systems. Tampere University of Technology (2012). https://dspace.cc.tut.fi/dpub/bitstream/handle/123456789/20958/garcia_izaguirre.pdf
12. Drools Fusion 6.2.0 documentation. https://docs.jboss.org/drools/release/6.2.0.CR3/drools-docs/html/
13. Fulop, L.J., Toth, G., Rcz, R., Panczel, J., Gergely, T., Beszedes, A., Farkas, L.: Survey on complex event processing and predictive analytics. In: Proceedings of the Fifth Balkan Conference in Informatics, pp. 26–31, July 2010
14. Saboor, M., Rengasamy, R.: Designing and developing complex event processing applications. Sapient Global Markets (2013)
15. Carney, D., Cetintemel, U., Cherniack, M., Convey, C., Lee, S., Seidman, G., Stonebraker, M., Tatbul, N., Zdonik, S.: Monitoring streams: a new class of data management applications. In: Proceedings of the 28th International Conference on Very Large Data Bases, pp. 215–226. VLDB Endowment, August 2002
16. Diao, Y., Immerman, N., Gyllstrom, D.: SASE+: an agile language for kleene closure over event streams. UMass Technical Report (2007)
17. Bhargavi, R., Pathak, R., Vaidehi, V.: Dynamic complex event processing—adaptive rule engine. In: 2013 International Conference on Recent Trends in Information Technology (ICRTIT), pp. 189–194. IEEE, July 2013
18. Brenna, L., Gehrke, J., Hong, M., Johansen, D.: Distributed event stream processing with non-deterministic finite automata. In: Proceedings of the Third ACM International Conference on Distributed Event-Based Systems, p. 3. ACM, July 2009
19. Suhothayan, S., Gajasinghe, K., Loku Narangoda, I., Chaturanga, S., Perera, S., Nanayakkara, V.: Siddhi: a second look at complex event processing architectures. In: Proceedings of the 2011 ACM Workshop on Gateway Computing Environments, pp. 43–50. ACM, November 2011
20. ETES Intelligence - Esper and NEsper: Where Complex Event Processing meets Open Source. Esper & NEsper, 2, 2006-2013 (2006)
21. Petersen, E., To, M.A., Maag, S.: An online learning based approach for CEP rule generation. In: 2016 8th IEEE Latin-American Conference on Communications (LATINCOM), pp. 1–6. IEEE, November 2016
22. Endler, M., Briot, J.P., e Silva, F.S., de Almeida, V.P., Haeusler, E.H.: Towards stream-based reasoning and machine learning for IoT applications. In: Intelligent Systems Conference (IntelliSys), pp. 202–209. IEEE, September 2017
23. Deljac, Z., Randic, M., Krcelic, G.: Early detection of network element outages based on customer trouble calls. Decis. Support Syst. **73**, 57–73 (2015)
24. Ahmed, M., Mahmood, A.N., Hu, J.: A survey of network anomaly detection techniques. J. Netw. Comput. Appl. **60**, 19–31 (2016)
25. Ahmed, M., Mahmood, A.N.: A novel approach for outlier detection and clustering improvement. In: 2013 8th IEEE conference on Industrial electronics and applications (ICIEA), pp. 577–582. IEEE, June 2013
26. Ester, M., Kriegel, H. P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: KDD, vol. 96, no. 34, pp. 226–231, August 1996
27. Thottan, M., Ji, C.: Anomaly detection in IP networks. IEEE Trans. Signal Process. **51**(8), 2191–2204 (2003)
28. Shyu, M.L., Chen, S.C., Sarinnapakorn, K., Chang, L.: A novel anomaly detection scheme based on principal component classifier. Department of Electrical and Computer Engineering, Miami University Coral Gables, FL (2003)

fridaya@uef.fi

29. Yang, Y., McLaughlin, K., Littler, T., Sezer, S., Wang, H.F.: Rule-based intrusion detection system for SCADA networks (2013)
30. Poojitha, G., Kumar, K.N., Reddy, P.J.: Intrusion detection using artificial neural network. In: 2010 International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–7. IEEE, July 2010
31. Balabine, I., Velednitsky, A.: U.S. Patent No. 9,843,488. U.S. Patent and Trademark Office Washington, DC (2017)
32. Dijkman, R., Peters, S., ter Hofstede, A.: A toolkit for streaming process data analysis. In: 2016 IEEE 20th International Enterprise Distributed Object Computing Workshop (EDOCW), pp. 1–9. IEEE, September 2016
33. Gaunitz, B., Roth, M., Franczyk, B.: Dynamic and scalable real-time analytics in logistics combining Apache Storm with complex event processing for enabling new business models in logistics. In: 2015 International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE), pp. 289–294. IEEE, April 2015
34. Bansod, R., Kadarkar, S., Virk, R., Raval, M., Rashinkar, R., Nambiar, M.: High performance distributed in-memory architectures for trade surveillance system. In: 2018 17th International Symposium on Parallel and Distributed Computing (ISPDC), pp. 101–108. IEEE, June 2018
35. Liu, G., Zhu, W., Saunders, C., Gao, F., Yu, Y.: Real-time complex event processing and analytics for smart grid. Proc. Comput. Sci. **61**, 113–119 (2015)
36. Chakraborty, S., Nagwani, N.K.: Analysis and study of Incremental DBSCAN clustering algorithm. arXiv preprint arXiv:1406.4754 (2014)
37. Cao, J., Wei, X., Liu, Y.Q., Mao, D., Cai, Q.: LogCEP-Complex event processing based on pushdown automaton. Int. J. Hybrid Inform. Technol. **7**(6), 71–82 (2014)
38. Koch, G. G., Koldehofe, B., Rothermel, K.: Cordies: expressive event correlation in distributed systems. In: Proceedings of the Fourth ACM International Conference on Distributed Event-Based Systems, pp. 26–37. ACM, July 2010
39. NetData. https://my-netdata.io/
40. Prometheus monitoring system. https://prometheus.io/
41. Consul by HashiCorp. https://consul.io/
42. Apache Kafka. https://apache.kafka.org/
43. Apache Spark: Lightning-fast cluster computing. http://spark.apache.org
44. Marz, N.: Storm: distributed and fault-tolerant realtime computation (2013). https://www.infoq.com/presentations/Storm-Introduction. Accessed 21 Oct 2011
45. Flink, A.: Scalable batch and stream data processing (2016). https://flink.apache.org

fridaya@uef.fi

Researches on Time-Domain Relay Feedback Identification Approaches for High-Acceleration Linear Servo Systems

Xiaoli Shi, Yong Han, Jianhua Wu, and Chao Liu(✉)

State Key Laboratory of Mechanical System and Vibration,
School of Mechanical Engineering, Shanghai Jiao Tong University,
Shanghai 200240, China
{jxgc-shixiaoli, ramoflaple, wujh, aalon}@sjtu.edu.cn

Abstract. The amplitude and frequency of linear servo system output responses are affected by the main harmonic approximation of description function in frequency-domain identification approaches. To deal with this issue, the time-domain analysis of continuous oscillation signal is implemented in this paper to establish accurate response expressions, and the velocity/position-based time-domain relay feedback identification method is derived in detail. Simulations show that compared with the frequency-domain methods, the velocity/position-based time-domain identification method improves the identification accuracy of model parameters with the ability of effectively avoiding the approximation problem. Experiments are carried out on the high-acceleration PMLSM-driven system and the results demonstrate that the tracking performance of high-acceleration linear servo system with the position-based time-domain identification method performs better than that with the velocity-based time-domain identification method and the maximum tracking error is reduced by 35.48%.

Keywords: Relay feedback technology · Time-domain identification · Ideal relay · Linear servo system

1 Introduction

With the development of motor technique and drive control, permanent magnet linear synchronous motor (PMLSM), which has the advantages of simple structure, high thrust-weight ratio, high stiffness and little friction, has been gradually become the key and indispensable component for high-end precision machines [1–3]. In order to accurately and effectively reflect the characteristics of PMLSM-driven system, such as nonlinear, high-order, fast time-varying, strong-coupling and so on, many researchers dedicated to system modeling and proposed a vast amount of approaches. Liu et al. [4] established a reduced-order model based on the frequency response test data, which consists of first-order low-frequency model, time delay and three main resonance models. Within the wide frequency range, a multi-mode vibration coupled model was proposed for high-performance motion systems by analyzing the kinematics and dynamic features [5]. Iwasaki et al. [6] proposed a high-order model according to the multi-order natural frequency of system and analyzed the servo performances affected

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 419–434, 2019.
https://doi.org/10.1007/978-3-030-22868-2_32

fridaya@uef.fi

by the first natural frequency fluctuation such as overshoot, tracking error, convergence time and so on. Chen et al. [7] proposed a novel physical model for linear servo system based on the analysis of nonlinearities and high-frequency structural flexible modes caused by the ball bearings flexibility. However, the above-mentioned high-order system models cannot be directly applied into high-performance control algorithms, such as variable structure control [8], adaptive robust control [9], iterative learning control [10], H_∞-based robust control [11] and disturbance observer [12]. The high-order system model should be reduced to control-oriented model for implementing the various high-performance control algorithms and the identified parameters of control-oriented model largely determine servo system performance. Therefore, the control-oriented model design and parameters identification techniques for high-performance controllers need to be urgently explored.

By virtue of its strong robustness and easy-to-implement characteristic, relay feedback technique becomes the concerning focus and has been widely used in servo mechanical systems [13, 14]. Relay feedback technique was first proposed by Åström and Hägglund [15] and its basic idea is to identify the model parameters from the oscillations excited by the relay module. As a common frequency-domain identification technique, the describing function (DF) method, which approximates the oscillating signal by using the main harmonic of Fourier series of ideal relay module output, inevitably deteriorates the identification accuracy. To solve the problem of main harmonic approximation, Shen et al. [16] employed a saturation relay instead of the conventional ideal relay for improving the identification accuracy of oscillation gain and frequency. Panda and Yu [17] studied the shape factor of relay feedback and analyzed its influence to the identification accuracy. Liu et al. [18] proposed a relay shaping method of ideal relay module to identify the model of servo mechanical system, which regulates the shape of ideal relay by bringing in a low-pass filter and hence make the response of the relay module more sin-wave-like. Chen et al. [19] proposed a time-domain based approach with a triple relay feedback apparatus to identify friction-impeded servo mechanical systems, but this method increases the difficulty of model identification, since it requires not only the oscillation amplitude and frequency but also the time-domain experimental information. To deal with the issue, two time-domain relay feedback identification approaches are discussed in this paper and simulations are carried out to verify the simplicity and effectiveness of the proposed time-domain identification approaches.

The remainder of this paper is organized as follows. Section 2 gives the dynamic model of high-acceleration PMLSM-driven system. In Sect. 3, two different time-domain relay feedback identification approaches are discussed and identification algorithms are formulated in detail. In Sect. 4, compared with the frequency-domain approaches, simulations show the validity of the proposed time-domain identification methods. In Sect. 5, experiments are carried out on the high-acceleration PMLSM-driven system, and the results show a better tracking performance with the position-based time-domain relay feedback identification approach. Section 6 gives the conclusion.

fridaya@uef.fi

2 Model of Linear Servo Systems

The dynamic model of high-acceleration PMLSM-driven system can be written as:

$$m\ddot{x} = F_e - F_d \tag{1}$$

where m is the mass of the slider, $\ddot{x}$ is the acceleration, F_e is the generated force which can be expressed as $F_e = k_t u$, k_t is the thrust coefficient, u is the input voltage. F_d is the external disturbances. Since viscous friction is only extracted from F_d, the dynamic model can be expressed as:

$$m\ddot{x} + B\dot{x} = k_t u - F_{dd} \tag{2}$$

where $\dot{x}$ is the velocity, B is the viscous coefficient and F_{dd} is the external disturbances besides viscous friction. Equation (2) can be rewritten as:

$$\alpha\ddot{x} + \beta\dot{x} = u - f_{dd} \tag{3}$$

where $\alpha = m/k_t, \beta = B/k_t$ and $f_{dd} = F_{dd}/k_t$.

Substitute $\tau = \alpha/\beta$ and $k = 1/\beta$ into Eq. (3), and the transfer function $G_v(\mathrm{s})$ from u to $\dot{x}$ can be obtained as:

$$G_v(\mathrm{s}) = \frac{\dot{X}(\mathrm{s})}{U(\mathrm{s})} = \frac{k}{\tau \mathrm{s} + 1} \tag{4}$$

where τ is the time constant, k is the system gain, which is a relatively large constant in servo mechanical systems, $\dot{X}(\mathrm{s})$ is the Laplace transformation of $\dot{x}$ and $U(\mathrm{s})$ is the Laplace transformation of u.

Similarly, the transfer function $G_p(\mathrm{s})$ from u to x is:

$$G_p(s) = \frac{X(s)}{U(s)} = \frac{k}{\tau s^2 + s} \tag{5}$$

3 Time-Domain Relay Feedback Identification Approaches

3.1 Velocity-Based Time-Domain Identification Approach

Figure 1 shows the block diagram of the velocity-based ideal relay feedback test. The linear servo system generates a linear motion under the excitation of initial disturbance, and the actual displacement signal of the motion is reflected by linear optical encoder. Then the actual displacement is converted to velocity signal through the difference method. After that, the difference $e_v(t)$ between the actual velocity signal and the reference input $r(t) = 0$ is obtained and delivered to ideal relay module to generate the corresponding voltage $u_R(t)$. Finally, $u_R(t)$ generates $u(t)$ after a time-delay module, and the voltage $u(t)$ drives the PMLSM-driven system to realize movement.

fridaya@uef.fi

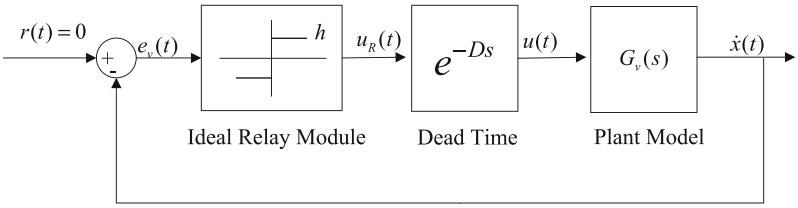


Fig. 1. Block diagram of velocity-based ideal relay feedback test

The input of ideal relay module is a sinusoid with expression shown below:

$$e_v(t) = a_{fv} \sin(\omega_{fv} t) \tag{6}$$

where a_{fv} and ω_{fv} are the amplitude and frequency of input signal of ideal relay module, t is the time variable.

Under the excitation of input signal, the output of ideal relay module shown in Fig. 1 keeps square-wave-like, and its Fourier series can be expressed as:

$$u_R = A + \sum_{i=1}^{+\infty} B_i \sin(i\omega_{fv} t) + \sum_{i=1}^{+\infty} C_i \cos(i\omega_{fv} t) \tag{7}$$

where

$$A = \frac{1}{2\pi} \int_0^{2\pi} u_R(t) d(\omega_{fv} t)$$
$$B = \frac{1}{\pi} \int_0^{2\pi} u_R(t) \sin(i\omega_{fv} t) d(\omega_{fv} t)$$
$$C_i = \frac{1}{\pi} \int_0^{2\pi} u_R(t) \cos(i\omega_{fv} t) d(\omega_{fv} t)$$

Since it is an odd function, $A = 0$. Similarly, $C_i = 0$. So the Fourier series of output square wave $u_R(t)$ can now be expressed as:

$$u_R = \sum_{i=1}^{+\infty} B_i \sin(i\omega_{fv} t) \tag{8}$$

In frequency-domain method, the main harmonic of right-part expression of (8), i.e. $B_1 \sin(\omega_{fv} t)$ with $i = 1$, is usually utilized to approximate the output square wave. Meanwhile, $B_1 = 4h/\pi$, where h is the amplitude of output response. And then, the description function of ideal relay model can be expressed as:

$$N_v(a_{fv}, \omega_{fv}) = \frac{u_R(t)}{e_v(t)} = \frac{B_1 \sin(\omega_{fv} t)}{a_{fv} \sin(\omega_{fv} t)} = \frac{4h}{\pi a_{fv}} \tag{9}$$

fridaya@uef.fi

According to the block diagram of velocity-based ideal relay feedback test shown in Fig. 1, the open-loop transfer function of the system can be written as:

$$-1 = G_v(j\omega_{fv})N_v(a_{fv}, \omega_{fv})e^{-Ds} \tag{10}$$

Substitute Eqs. (4) and (9) into Eqs. (10) and (11) can be obtained:

$$\frac{4h}{\pi a_{fv}}\cos(D\omega_{fv}) - \frac{4h}{\pi a_{fv}}\sin(D\omega_{fv})j = -\frac{1}{k} - \frac{\tau}{k}\omega_{fv}j \tag{11}$$

Then the model parameters of $G_v(s)$, i.e. τ and k, can be computed by comparing the real and imaginary parts of the right-side expression of Eq. (11) to those of the left side, and the calculation results can be expressed as:

$$\hat{\tau}_{fiv} = \frac{\tan(\pi - D\omega_{fv})}{\omega_{fv}} \tag{12}$$

$$\hat{k}_{fiv} = \frac{\pi a_{fv}\sqrt{(\hat{\tau}_{fiv}\omega_{fv})^2 + 1}}{4h} \tag{13}$$

where $\hat{\tau}_{fiv}$ and $\hat{k}_{fiv}$ are the identified values by using the velocity-based frequency-domain relay feedback identification approach. This method approximates the oscillating signal by using the main harmonic of Fourier series of ideal relay module output, inevitably deteriorates the identification accuracy. To solve the problem of main harmonic approximation, a velocity-based time-domain relay feedback identification approach is given as follows.

Since the identified system can be assumed to be symmetrical, only the first half of oscillation cycle is analyzed to identify the model parameters of $G_v(s)$. According to Fig. 1, the relationship between the input $u(t)$ and output $\dot{x}(t)$ of plant model $G_v(s)$ is obtained and shown in Fig. 2. The first half of oscillation cycle can be divided into two intervals: $[0, D]$ and $(D, T_{hv}]$. And the output response in the first-half oscillation cycle can be expressed as:

$[0, D]$:

$$\dot{x}(t) = k\mu(1 - e^{-t/\tau}) \tag{14}$$

$(D, T_{hv}]$:

$$\dot{x}(t) = k\mu(1 - e^{-t/\tau}) - 2k\mu(1 - e^{-(t-D)/\tau}) \tag{15}$$

where μ represents the actual thrust of linear motor. The value of μ and h are equal without regard to coulomb friction or other disturbances.

fridaya@uef.fi

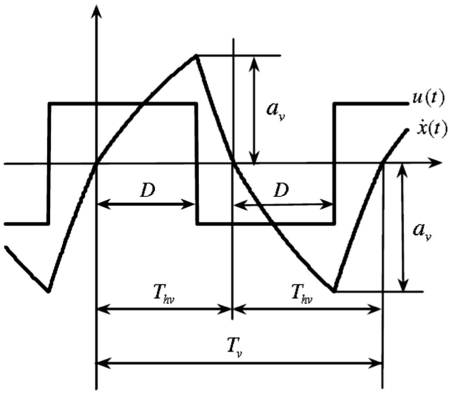


Fig. 2. Input-output relationship of plant model $G_v(s)$

From Fig. 2, the boundary conditions of oscillation can be obtained as:

$$\dot{x}(D) = a_v, \dot{x}(T_{hv}) = 0 \tag{16}$$

Take Eq. (16) into Eqs. (14) and (15), the time-domain characteristics of output response $\dot{x}(t)$ can be expressed as:

$$kh(1 - e^{-D/\tau}) = a_v \tag{17}$$

$$e^{-T_{hv}/\tau} - 2e^{(D-T_{hv})/\tau} + 1 = 0 \tag{18}$$

$\hat{\tau}_{tiv}$ can be computed through numerical calculation method from:

$$T_{hv} = \tau \ln(2e^{D/\tau} - 1) \tag{19}$$

And then, the model parameter k of $G_v(s)$ can be obtained as:

$$\hat{k}_{tiv} = \frac{a_v}{h(1 - e^{-D/\hat{\tau}_{tiv}})} \tag{20}$$

Till now, $\hat{\tau}_{tiv}$ and $\hat{k}_{tiv}$ are separately calculated out by using the velocity-based time-domain relay feedback identification approach.

3.2 Position-Based Time-Domain Identification Approach

Figure 3 shows the block diagram of ideal relay position feedback, and the ideal relay description function can be expressed as:

$$N_p(a_{fp}, \omega_{fp}) = \frac{B_1 \sin(\omega_{fp} t)}{a_{fp} \sin(\omega_{fp} t)} = \frac{4h}{\pi a_{fp}} \tag{21}$$

fridaya@uef.fi

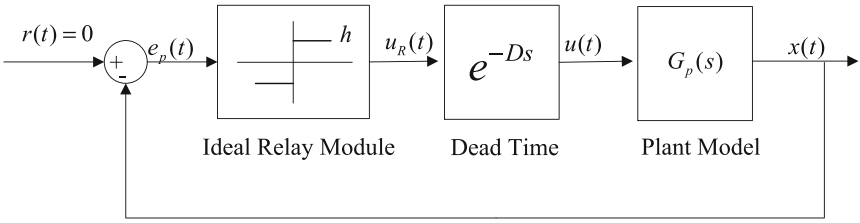


Fig. 3. Block diagram of position-based ideal relay feedback test

where a_{fp} and ω_{fp} are the magnitude and frequency of input signal of ideal relay module, respectively.

The open-loop transfer function is:

$$-1 = G_p(j\omega_{fp})N_p(a_{fp}, \omega_{fp})e^{-Ds} \tag{22}$$

Take Eqs. (5) and (21) into Eq. (22), it can be obtained as:

$$\frac{4h}{\pi a_{fp}}\cos(D\omega_{fp}) - \frac{4h}{\pi a_{fp}}\sin(D\omega_{fp})j = \frac{\tau}{k}\omega_{fp}^2 - \frac{1}{k}\omega_{fp}j \tag{23}$$

The model parameters τ and k of plant model $G_p(s)$ can be calculated as:

$$\hat{\tau}_{fip} = \frac{1}{\omega_{fp}\tan(D\omega_{fp})} \tag{24}$$

$$\hat{k}_{fip} = \frac{\pi a_{fp}\omega_{fp}\sqrt{(\hat{\tau}_{fip}\omega_{fp})^2 + 1}}{4h} \tag{25}$$

where $\hat{\tau}_{fip}$ and $\hat{k}_{fip}$ are the identified values by using the position-based frequency-domain relay feedback identification approach. Velocity is obtained from actual displacement signal by using the difference method, which inevitably amplifies the quantization errors during velocity detection. Therefore, a position-based time-domain relay feedback identification approach is detailed here to avoid the noise and further improve the identification accuracy of model parameters.

According to the block diagram of position-based ideal relay feedback test shown in Fig. 3, the input-velocity and input-displacement relationships of plant model $G_p(s)$ are given in Fig. 4. It can be seen that the oscillation in position feedback test is more complicated than that in velocity feedback test. After the control signal reverses, the position signal keeps rising (falling) till the velocity passes through 0, but not following the relay input.

fridaya@uef.fi

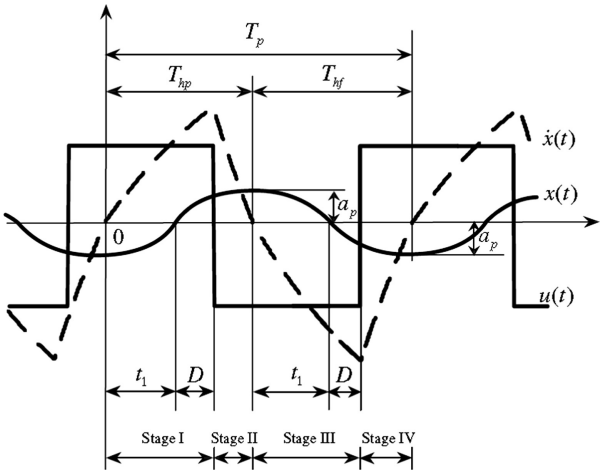


Fig. 4. Input-velocity and input-displacement relationships of plant model $G_p(s)$

According to the input-velocity and input-displacement relationships shown in Fig. 4, the displacement in the first-half cycle T_{hp} can be expressed as:

$[0, t_1 + D]$:

$$x(t) = k\mu'(t - \tau + \tau e^{-t/\tau}) - a_p \tag{26}$$

Where μ' is the actual thrust of linear motor, t_1 is the intermediate variable in Fig. 4.

$(t_1 + D, T_{hp}]$:

$$x(t) = k\mu'(t - \tau + \tau e^{-t/\tau}) - 2k\mu'(t - t_1) + 2k\mu'(D + \tau + \tau e^{-(t-t_1-D/\tau)}) - a_p \tag{27}$$

The boundary conditions of the oscillation are:

$$x(t_1) = 0, x(T_{hp}) = a_p \tag{28}$$

Substitute Eq. (28) into Eqs. (26) and (27), the time-domain characteristics of oscillation displacement $x(t)$ can be expressed as:

$$k\mu'(t_1 - \tau + \tau e^{-t_1/\tau}) = a_p \tag{29}$$

$$k\mu'(T_{hp} - \tau + \tau e^{-T_{hp}/\tau}) - 2k\mu'(T_{hp} - t_1) + 2k\mu'(D + \tau + \tau e^{-\left(T_{hp}-t_1-D/\tau\right)}) = 2a_p \tag{30}$$

The velocity $\dot{x}(t)$ in $(t_1 + D, T_{hp}]$ can be expressed as:

$$\dot{x}(t) = k\mu'(1 - e^{-t/\tau}) - 2k\mu'(1 - e^{-(t-t_1-D)/\tau}) \tag{31}$$

fridaya@uef.fi

Under the condition of $\dot{x}(T_{hp}) = 0$, Eq. (31) can be rewritten as:

$$k\mu'(1 - e^{-T_{hp}/\tau}) - 2k\mu'(1 - e^{-(T_{hp}-t_1-D)/\tau}) = 0 \tag{32}$$

Substitute Eqs. (30) and (32) into Eq. (29), the expression of T_{hp} is:

$$T_{hp} = 2(D + \tau) - \frac{4\tau e^{D/\tau}}{e^{T_{hp}/\tau} + 1} \tag{33}$$

The identified value of τ, i.e. $\hat{\tau}_{tip}$, can be obtained from Eq. (33) by numerical calculation. And then, the calculation formula of the intermediate variable t_1 can be obtained as:

$$\hat{t}_1 = \hat{\tau}_{tip} \ln(2\hat{\tau}_{tip}/(2D - T_{hp} + 2\hat{\tau}_{tip})) \tag{34}$$

Take Eq. (34) into Eq. (29), the identification result of k, i.e. $\hat{k}_{tip}$, is expressed as:

$$\hat{k}_{tip} = \frac{a_p}{\mu'(\hat{t}_1 - \hat{\tau}_{tip} + \hat{\tau}_{tip} e^{-\hat{t}_1/\hat{\tau}_{tip}})} \tag{35}$$

Till now, $\hat{\tau}_{tip}$ and $\hat{k}_{tip}$ are separately calculated out by using the position-based time-domain relay feedback identification approach.

4 Simulations

Based on the equivalent simulation system of high-acceleration linear servo system [20], the simulation analysis of different relay feedback identifications are carried out with four plant models. The values of ideal relay module parameters are $h = 0.1$ v and $D = 0.05$ s, and the sample period is set to 0.1 ms.

With reference to Fig. 1, the velocity-based relay feedback tests are carried out. The period and amplitude of velocity response are recorded. With the velocity-based frequency/time-domain relay feedback identification methods, $\hat{\tau}_{fiv}, \hat{k}_{fiv}, \hat{\tau}_{tiv}$ and $\hat{k}_{tiv}$ are presented in Table 1. Similarly, the position-based relay feedback tests are also implemented based on the schematic of Fig. 3 and the identified values of $\hat{\tau}_{fip}, \hat{k}_{fip}, \hat{\tau}_{tip}$ and $\hat{k}_{tip}$ calculated out by using the position-based frequency/time-domain relay feedback identification methods are given in Table 1.

Table 1. Identification results with different identification methods.

Num.	τ	k	$\hat{\tau}_{fiv}$	$\hat{k}_{fiv}$	$\hat{\tau}_{tiv}$	$\hat{k}_{tiv}$	$\hat{\tau}_{fip}$	$\hat{k}_{fip}$	$\hat{\tau}_{tip}$	$\hat{k}_{tip}$
1	0.2	5	0.167	5.104	0.207	5.119	0.223	5.396	0.199988	5.002
2	0.2	10	0.167	10.208	0.207	10.239	0.223	10.792	0.199988	10.005
3	0.1	5	0.082	5.107	0.102	5.160	0.109	5.334	0.099956	4.998
4	0.1	10	0.082	10.214	0.102	10.319	0.109	10.667	0.099956	9.996

fridaya@uef.fi

Table 2. Identification relative errors with different identification methods.

Num.	$\hat{\tau}_{fiv}$	$\hat{k}_{fiv}$	$\hat{\tau}_{tiv}$	$\hat{k}_{tiv}$	$\hat{\tau}_{fip}$	$\hat{k}_{fip}$	$\hat{\tau}_{tip}$	$\hat{k}_{tip}$
1	16.50%	2.08%	3.50%	2.38%	11.50%	7.92%	0.006%	0.04%
2	16.50%	2.08%	3.50%	2.39%	11.50%	7.92%	0.006%	0.05%
3	18.00%	2.14%	2.00%	3.20%	9.00%	6.68%	0.044%	0.04%
4	18.00%	2.14%	2.00%	3.19%	9.00%	6.67%	0.044%	0.04%

Table 2 shows the identification accuracy of model parameters obtained by different identification methods. With the velocity-based frequency-domain relay feedback identification method, the average relative errors of $\hat{\tau}_{fiv}$ and $\hat{k}_{fiv}$ are 17.25% and 2.11%, respectively. With the position-based frequency-domain relay feedback identification method, the average relative errors of $\hat{\tau}_{fip}$ and $\hat{k}_{fip}$ are 10.25% and 7.30%, respectively. By avoiding the main harmonic of Fourier series of ideal relay module output, the time-domain identification methods improve the identification accuracy of model parameters. The average relative errors of $\hat{\tau}_{tiv}, \hat{\tau}_{tip}, \hat{k}_{tiv}$ and $\hat{k}_{tip}$ are 2.75%, 0.025%, 2.79% and 0.043%, respectively. Since the quantization errors are effectively suppressed, the identification accuracy of model parameters is further improved by using the position-based time-domain relay feedback identification method and the maximum relative errors of $\hat{\tau}_{tip}$ and $\hat{k}_{tip}$ are both within 0.05%.

5 Experimental Studies

5.1 Experimental Setup

The experimental setup is shown in Fig. 5, which is composed of an X-Y positioning stage, a PC-based dSPACE controller, actuating amplifiers, cooling devices, power supplies and so on. Experiments are conducted on the Y-axis of high-acceleration PMLSMs-driven positioning stage, in which the sampling time is set to 0.1 ms. A non-contact linear optical encoder with hardware resolution of 0.4 μm (0.1 μm after quadruplication) is used as the feedback element of Y-axis. The PC-based dSPACE control system is used here for the computational intensive tasks associated with execution of the referred algorithms.

5.2 System Identification

From the above-mention simulation results, the time-domain identification methods are implemented here for system identification. According to Figs. 1 and 3, the relay feedback experiments are carried out with the parameters $h = 0.2$ v and $D = 0.02$ s. Figures 6 and 7 show the output of ideal relay module and velocity response of high-acceleration linear servo system obtained by the velocity-based relay feedback test. The period and amplitude of velocity response are 0.0777 s and 57 mm/s, respectively. The model parameters of $G_v(s)$ obtained by the velocity-based time-domain relay feedback identification method are $\hat{\tau}_{tiv}$ = 0.3333 s and $\hat{k}_{tiv}$ = 4893.38 Vs/mm.

fridaya@uef.fi

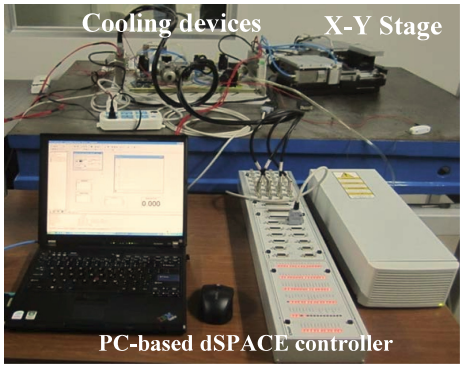


Fig. 5. Experimental setup

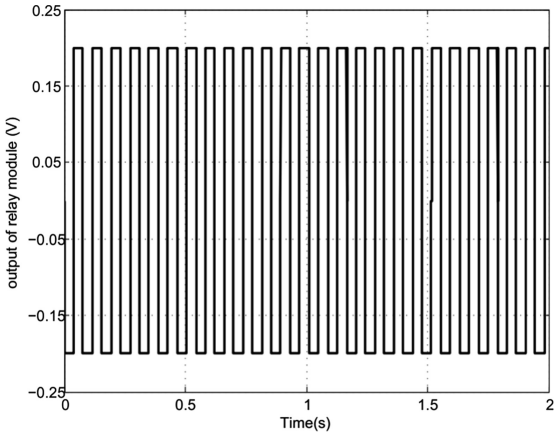


Fig. 6. Output of ideal relay module based on velocity feedback

Figures 8 and 9 show the output of ideal relay module and position response of high-acceleration linear servo system obtained by the position-based relay feedback test. The period and amplitude of displacement response are 0.2942 s and 0.8887 mm, respectively. The model parameters of $G_p(s)$ obtained by the position-based time-domain relay feedback identification method are $\hat{\tau}_{tip} = 0.0922$ and $\hat{k}_{tip} = 1662.95$ Vs/mm.

fridaya@uef.fi

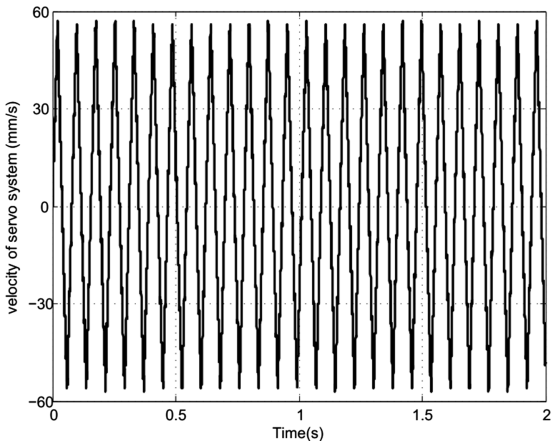


Fig. 7. Velocity response of servo system based on velocity feedback

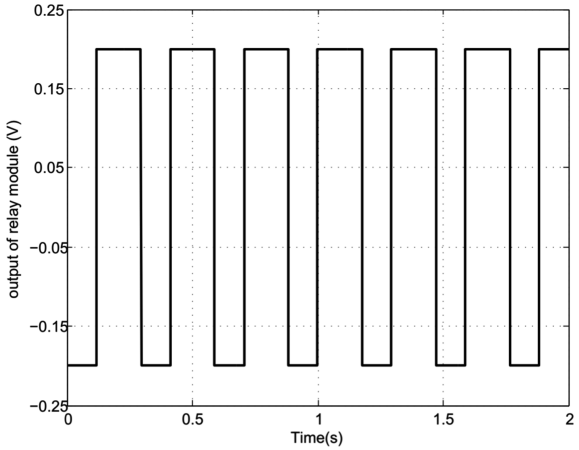


Fig. 8. Output of ideal relay module based on position feedback

5.3 Tracking Performance

Under the hybrid control scheme shown in Fig. 10, which consists of a PD feedback controller, a feedforward compensator and a disturbance observer, high-acceleration short-stroke motions are carried out to analyze the system tracking performance under the identified value of model parameters mentioned in Sect. 5.2. The PD feedback controller is designed for achieving the desired tracking performance as well as stabilizing the closed-loop system. The feedforward compensator is used for improving the system dynamics. The disturbance observer is utilized to estimate and compensate the disturbances for improving the performance of the high-acceleration linear servo system.

fridaya@uef.fi

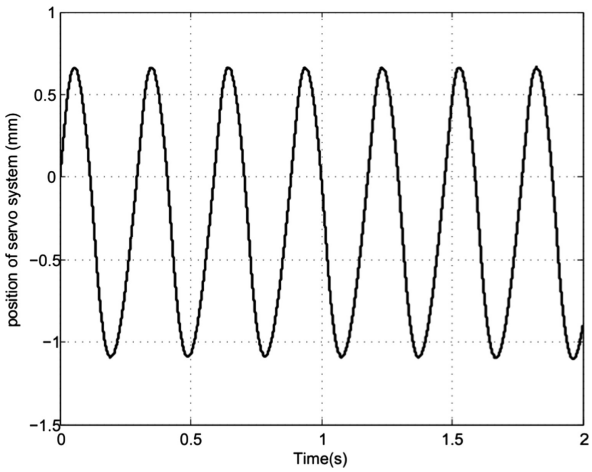


Fig. 9. Position response of servo system based on position feedback

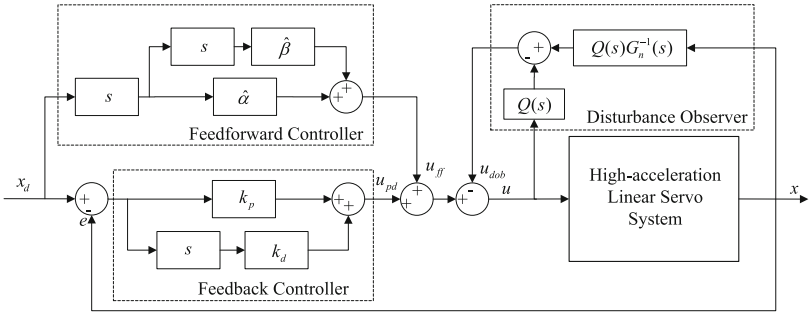


Fig. 10. The employed hybrid control scheme

Based on the identified value of model parameters mentioned in Sect. 5.2, the coefficients $\hat{\alpha}$ and $\hat{\beta}$ of the feedforward controller can be directly calculated from the model-inverse method and the coefficients k_p and k_d can also be obtained by the pole-placement method. The poles are $\rho_1 = -400$ rad/s and $\rho_2 = -400$ rad/s. The coefficient τ_1 of the disturbance observer is chosen as 0.0005 s by fully taking into account of the system characteristics.

Figure 11 shows the tracking error curve with velocity-based time-domain relay feedback identification method and the maximum value of tracking errors is 3.1 μm. Figure 12 shows the tracking error curve with position-based time-domain relay feedback identification method and the maximum value of tracking errors is 2.0 μm. Compared with the velocity-based time-domain identification method, the position-based time-domain identification method makes high-acceleration linear servo systems have a better tracking performance and the maximum tracking error is reduced by 35.48%.

fridaya@uef.fi

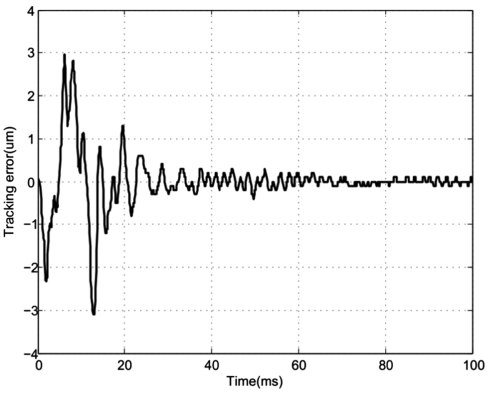


Fig. 11. Tracking errors with velocity-based time-domain identification method

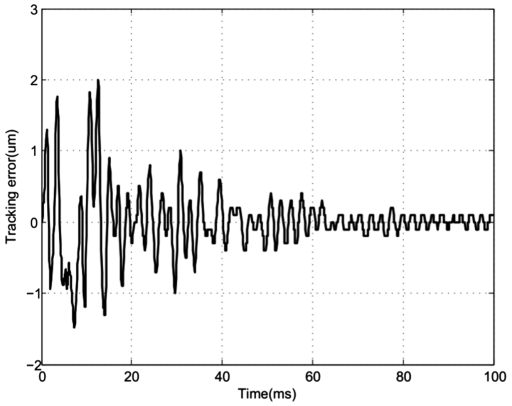


Fig. 12. Tracking errors with position-based time-domain identification method

6 Conclusion

In the frequency-domain relay feedback identification process, the describing function is widely used for identifying the model parameters. However, this method approximates the oscillating signal by using the main harmonic of Fourier series of ideal relay module output, and then inevitably deteriorates the identification accuracy. To tackle with the issue, the time-domain analysis of continuous oscillation signal is detailed here to establish accurate response expressions. The velocity-based time-domain relay feedback identification method is first derived and then position-based time-domain relay feedback identification method is given to further suppress the quantization errors for the identification accuracy improvement. Simulations show that the velocity/position-based time-domain identification method can both improve the identification accuracy of model parameters. Experiments are implemented on the high-acceleration linear servo system and their results verify that the system tracking

fridaya@uef.fi

performance with the position-based time-domain identification method performs better than that with the velocity-based time-domain identification method. The maximum tracking error is reduced from 3.1 μm to 2.0 μm.

Acknowledgments. This work supported by the National Natural Science Foundation of China (Grant No. 51805325) and the China Postdoctoral Science Foundation (Grant No. 2015M580325).

References

1. Hama, T., Sato, K.: High-speed and high-precision tracking control of ultrahigh-acceleration moving-permanent-magnet linear synchronous motor. Precis. Eng. **40**(1), 151–159 (2015)
2. Ding, H., Xiong, Z.: Motion stages for electronic packaging design and control. IEEE Robot. Autom. Mag. **13**(4), 51–61 (2006)
3. Kim, J., Choi, S., Cho, K., et al.: Position estimation using linear hall sensors for permanent magnet linear motor systems. IEEE Trans. Industr. Electron. **63**(12), 7644–7652 (2016)
4. Liu, Z.Z., Luo, F.L., Rahman, M.A.: Robust and precision motion control system of linear-motor direct drive for high-speed X-Y table positioning mechanism. IEEE Trans. Industr. Electron. **52**(5), 1357–1363 (2005)
5. Liu, Q., Qi, C., Yuan, S., et al.: Modeling and control of hi-performance motion system. In: World Congress on Intelligent Control and Automation (WCICA 2008), pp. 455–460. IEEE, Piscataway, NJ, Chongqing, China (2008)
6. Iwasaki, M., Seki, K., Maeda, Y.: High-precision motion control techniques: a promising approach to improving motion performance. IEEE Ind. Electron. Mag. **6**(1), 32–40 (2012)
7. Chen, Z., Yao, B., Wang, Q.: μ-synthesis-based adaptive robust control of linear motor driven stages with high-frequency dynamics: a case study. IEEE/ASME Trans. Mechatron. **20**(3), 1482–1490 (2015)
8. Cheema, M.A.M., Fletcher, J.E., Farshadnia, M., et al.: Combined speed and direct thrust force control of linear permanent-magnet synchronous motors with sensorless speed estimation using a sliding-mode control with integral action. IEEE Trans. Ind. Electron. **64**(5), 3489–3501 (2017)
9. Xu, L., Yao, B.: Adaptive robust precision motion control of linear motors with negligible electrical dynamics: theory and experiments. IEEE/ASME Trans. Mechatron. **6**(4), 444–452 (2000)
10. Freeman, C.T., Tan, Y.: Iterative learning control with mixed constraints for point-to-point tracking. IEEE Trans. Control Syst. Technol. **21**(3), 604–616 (2013)
11. Moon, S.J., Chung, T.Y., Lim, C.W., et al.: A linear motor damper for vibration control of steel structures. Mechatronics **14**(10), 1157–1181 (2004)
12. Cho, K., Nam, K.: Periodic learning disturbance observer based precision motion control in PMLSM motion systems considering long-term instability problem. Int. J. Precis. Eng. Manufact. **17**(9), 1101–1112 (2016)
13. Pu, D.L., Wu, J.H., Xiong, Z.H., et al.: Nonlinear analysis and parameters identification of servo mechanism with relay feedback. Assembly Autom. **30**(3), 221–227 (2010)
14. Han, Y., Liu, C., Wu, J.: Backlash identification for PMSM servo system based on relay feedback. Nonlinear Dyn. **84**(4), 1–13 (2016)
15. Åström, K.J., Hägglund, T.: Automatic tuning of simple regulators with specifications on phase and amplitude margins. Automatica **20**(5), 645–651 (1984)
16. Shen, S.H., Yu, H.D., Yu, C.C.: Use of saturation-relay feedback for autotune identification. Chem. Eng. Sci. **51**(8), 1187–1198 (1996)

fridaya@uef.fi

17. Panda, R.C., Yu, C.C.: Shape factor of relay response curves and its use in autotuning. J. Process Control **15**(8), 893–906 (2005)
18. Liu, C., Liu, J., Wu, J.H., et al.: A relay shaping method for servo mechanical system identification. In: Intelligent Robotics and Applications, pp. 353–364. Springer, Berlin, Heidelberg (2013)
19. Chen, S.L., Tan, K.K., Huang, S.: Identification of coulomb friction-impeded systems with a triple-relay feedback apparatus. IEEE Trans. Control Syst. Technol. **20**(3), 726–737 (2012)
20. Liu, C., Wang, H., Wu, J., et al.: Design of an equivalent simulation platform for PMLSM-driven positioning system. In: Intelligent Robotics and Applications, pp. 210–220. Springer International Publishing (2015)

fridaya@uef.fi

Multimodal Biometrics Using Features from Face, Iris, and Conjunctival Vasculature

Noorjahan Khatoon[1,2(✉)] and Mrinal Kanti Ghose[3]

[1] Sikkim Manipal Institute of Technology, Majhitar 737136, East Sikkim, India
njk_gtk@yahoo.com
[2] Advanced Technical Training Centre, Bardang 737134, East Sikkim, India
[3] Sikkim University, Samdur 737102, East Sikkim, India

Abstract. The performance deficiencies of unimodal biometrics systems arising out of background noise, signal noise and distortions, and environment or device variations can be overcome by using a multimodal biometric system. In this study, an attempt has been made to implement a multimodal biometric fusion using the modalities of face, iris, and conjunctival vasculature. The features from each modality are first extracted using a local ternary pattern-based texture descriptor, viz., Steady Illumination colour Local Ternary Pattern, and then, these are fused at the feature level. The fusion is implemented using a concatenation of the extracted features from these modalities and similarity matching is conducted using Zero-Mean Sum of Squared Differences. Furthermore, the homogeneity of the features is maintained by applying the same feature extractor for all the modalities. The feature extractor is sufficiently efficient in extracting all important features to be fused, which leads to better accuracy. Moreover, the increased number of feature sets because of the fusion at the feature level is reduced using a genetic algorithm, which further improves the accuracy. Experimental results show the effectiveness of the proposed method and results reveal that multimodal biometrics verification is considerably more reliable and precise than the single biometric approach.

Keywords: Multimodal biometrics · Face · Iris · Conjunctival vasculature · Fusion · Concatenation

1 Introduction

In the past few decades, biometrics has drawn significant attention for its potentiality in automatic identification (or verification) of an individual (or a claimed identity) using certain physiological or behavioural traits associated with the person. Traditional passwords (knowledge-based security) and identification (ID) cards (token-based security) used to restrict access to locations or applications are not so secured. Nevertheless, Biometrics can play a significant role to ensure better security, as indicated by its definition. Most biometric systems used in real-world situations are unimodal in nature, i.e. they use a single biometric modality for authentication; this results in noisy sensor data, poor illumination, intra-class variations, and inter-class similarities [1]. Many studies have and are being conducted to improve the overall accuracy and robustness of

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 435–445, 2019.
https://doi.org/10.1007/978-3-030-22868-2_33

fridaya@uef.fi

the techniques used for feature extraction, data acquisition, and image pre-processing. Nonetheless, the use of a single biometric modality cannot overcome all the limitations posed by such unimodal systems. Therefore, multimodal biometrics has emerged, which is the current trend in the field of biometrics. Many studies and algorithms have been presented in literature regarding multimodal biometric fusion [2–4].

The sources for modality fusion might include multiple sensors for the same biometric (e.g. optical- and solid-state fingerprint sensors), multiple instances of the same biometric (e.g. fingerprints from different fingers of a person), multiple snapshots of the same biometric (e.g. four impressions of a user's right index finger), multiple representations and matching algorithms for the same biometric (e.g. multiple face matchers like principal component analysis and linear discriminant analysis), or multiple biometric traits (e.g. face and fingerprint). The use of multiple sensors addresses the problem of noisy sensor data; however, all other potential problems associated with unimodal biometric systems remain. A recognition system that works on multiple instances of the same biometric can ensure the presence of a live user by asking the user to provide a random subset of the biometric measurements (e.g. Fingerprint of left index finger followed by right middle finger). Multiple snapshots of the same biometric, or multiple representations and matching algorithms for the same biometric might also be used to improve the recognition performance of such a system. However, all these methods still suffer from many problems faced by unimodal systems [5].

In recent time, the development of biometric systems based on multiple biometric traits has received considerable attention from researchers due to its robustness, non-universality and better matching accuracy etc. In a multimodal biometric system various levels of fusion are possible:

- fusion at the feature level; and
- fusion at the matching score level or decision level.

Fusion at the decision level is effectively regarded as rigid because only a limited amount of information is available at this level. The only type of information available for fusion at the decision level is 'Accept' or 'Reject' [2]. In contrast, integration at the matching score level is generally preferred owing to the ease of access and combination of matching scores [5].

Sanderson and Paliwal [6] classified information fusion in biometric systems into two broad categories: pre-classification fusion and post-classification fusion. Pre-classification fusion refers to combining information prior to the application of any classifier or matching algorithm, whereas, in post-classification fusion, the information is combined after the decisions from the classifiers have been obtained.

In the context of verification, both these approaches have been widely studied in the literature. Ross and Jain [4] showed that the combination approach yields better results than some classification methods such as decision tree and linear discriminant analysis. However, no single classification or combination scheme works universally well in all circumstances. Because the matching scores generated by different modalities are heterogeneous, normalisation is required to transform these scores to a common domain before combining them [5].

fridaya@uef.fi

Despite the abundance of research related to multi-biometrics, fusion at the feature level is not as well studied in comparison to fusion at the matching score level [7–9]. Feature-level fusion is performed by a simple concatenation of the feature sets obtained from different sources [10]. Before performing the concatenation, the features are normalised to ensure the same range and scales of values. In addition, to reduce the dimensionality of the feature space, a feature selection/reduction technique is applied to the resultant feature sets [11].

Among the many biometric traits, the face and iris provide several useful features for recognition. Face sensing is non-invasive, and it is relatively easy to acquire good-quality face images [1]. Nevertheless, iris-based identification systems have proved to have the best performance in many application domains. The accuracy of face recognition is often affected by illumination, pose, obstructions, and facial expression, whereas current iris recognition systems require the active cooperation of the user. Thus, these two modalities are quite complementary. Moreover, as these two modalities belong to the same biological trait (i.e. the head), they can easily be sampled simultaneously with the same sensing technology [10]. Another complementary modality for the face and iris could be the Conjunctival Vasculature (CV), which could be examined when the iris is not centrally aligned.

Therefore, considering this, the motivation of our study is to fuse modalities using the face, iris, and CV. In this light, in our study, the approach for concatenation of the extracted features for fusion has been considered as the features are kept homogenous by extracting them from the face, iris, and CV using the same feature extractor.

The features of these modalities have been extracted using the descriptor that we previously proposed and implemented in [12], called the Steady Illumination colour Local Ternary Pattern (SIcLTP). In this study, the fusion has been considered at the feature level, and the homogeneity of the extracted features is maintained by using SIcLTP as the feature extractor for all three modalities. The obtained results are validated using Local Binary Patterns (LBPs).

The remainder of the paper is organised as follows. In Sect. 2, the fusion/integration strategy used for the feature sets extracted from the unimodalities, i.e. the face, iris, and CV, using SIcLTP [12] is described. Section 3 presents the multimodal fusion and its results. Section 4 discuss the feature reduction using a genetic algorithm (GA) and present the comparative results, respectively. Finally, Sect. 5 provides our conclusion and directions for future works.

2 Fusion of Features

The datasets used in the experiment for feature extraction and fusion are as follows: for the face, the Indian face database by Vidit and Amitabha [13] is used; for the iris, the images are partially obtained from the database by Michal Dobes and Libor Machala [14]. In addition, we captured additional images using a Nikon digital camera model D5200. We prepared the entire dataset for the CV using the same camera. Some sample images used from the datasets are shown in Figs. 1, 2, and 3.

fridaya@uef.fi

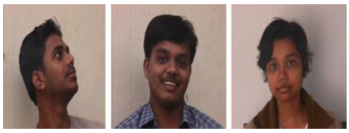


Fig. 1. Sample face images

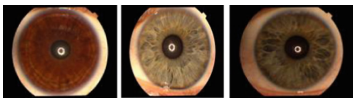


Fig. 2. Sample iris images

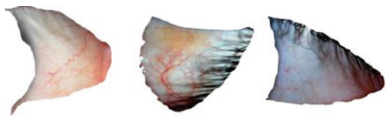


Fig. 3. Sample CV images

In general, the strategy adopted for integration depends on the level at which fusion is performed. Because the experiment is conducted by fusing the important features at the feature level, concatenation of the three compatible/homogenous SIcLTP extracted features sets—obtained in the form of ternary patterns of the texture—is performed. Features extracted using the LBP [15] as a feature extractor are also concatenated for validation.

Because the images are typically represented in the form of a matrix, it is easy to append/concatenate two or more images to create a single matrix. The concatenation of matrices can be performed either horizontally or vertically. Therefore, two or more matrices are simply appended either horizontally or vertically. One of the advantages of this method is that all the extracted features are retained in the images, which is, in fact, an important requirement for multimodal biometrics.

2.1 Steps for Fusion of Extracted Features

The fusion of the extracted features is performed as follows:

Step 1: Input the features extracted from the three modalities obtained using SIcLTP after converting the space of the RGB images into the YIQ colour space.
Step 2: Perform concatenation for the extracted features.
Step 3: Save the concatenated output in the form of an image.
Step 4: Match the images using zero-mean sum of squared differences (ZSSD) [16].
Step 5: Analyse the obtained result to determine the recognition accuracy.

fridaya@uef.fi

When the accuracy obtained after applying ZSSD to the fused modalities is compared with the accuracies of the individual modalities of the face, iris, and CV that we obtained previously in [12, 17, 18], which are tabulated in Tables 1 and 2, there are significant improvements for the fused case compared with the individual/unimodal cases. Accordingly, it can be strongly inferred that multimodal biometrics is more accurate and robust than unimodal biometrics. The results are analysed using the receiver operating characteristic (ROC) curve.

Table 1. Accuracies obtained for face, iris, and CV using SIcLTP

Modality	Recognition accuracy %
Face	82.2
Iris	82.1
CV	67.86

Table 2. Accuracies obtained for face, iris, and CV using LBP

Modality	Recognition accuracy %
Face	51.1
Iris	53.6
CV	21.43

In this study, multimodal performed to see the effect of fusion with respect to multimodality. Experiments are conducted to examine the effect of fusion of three modalities namely, face iris, and CV.

3 Experimental Results for Multimodal Fusion

Experiments were conducted to study the results of multimodal fusion; in addition, the obtained results were validated using LBP as a feature extractor. Table 3 lists the recognition accuracies obtained by the fusion of combined modalities.

3.1 Steady Illumination Colour Local Ternary Pattern (SIcLTP)

The mathematical expression of the SIcLTP operator is given below [12]:

Given any pixel location, (x_c, y_c), SIcLTP encodes it as

$$SIcLTP^{t}_{N,r}(x_c, y_c) = \sum_{b=0}^{N-1} S_t(P_c, P_b), \tag{1}$$

where

P_c is the intensity value of the centre pixel;
P_b isthe intensity value of its N neighbourhood pixels;
N is the neighbourhood;

fridaya@uef.fi

R is the radius;
$\sum$ denotes concatenation operator of binary strings; and
t is a scale factor indicating the comparison range.

Because each comparison can result in one of the three values, SIcLTP encodes each pixel with two bits, and S_t is a piecewise function defined as follows:

$$S_t(P_c, P_b) = \begin{cases} 01, & \text{if} \quad P_b > (1+t)P_c, \\ 10, & \text{if} \quad P_b < (1-t)P_c, \\ 00, & \text{otherwise.} \end{cases} \tag{2}$$

The advantages of using this operator are as follows. The operator is computationally simple and efficient. Furthermore, the presence of a tolerant value makes it robust in the case of noisy images. Moreover, the scale invariance property makes it more robust to illumination changes.

3.2 Zero Mean Sum of Squared Differences (ZSSD)

The ZSSD, which is used to find matches and mismatches among images, is determined by subtracting the mean of the matched area from each intensity value [16]. Mathematically, the ZSSD can be calculated as follows:

$$ZSSD = \sum_{w} \left(I_{\text{img1}}(x,y) - M_{\text{img1}} - I_{\text{img2}}(x,y) - M_{\text{img2}} \right)^2, \tag{3}$$

where

$I_{\text{img1}}(x,y)$ is the first image intensity at point (x, y);
$I_{\text{img2}}(x,y)$ is the second image intensity at point (x, y);
M_{img1} is the mean first image intensity over window w; and
M_{img2} is the mean second image intensity over window w.

Figure 4 shows an example of the concatenated features extracted by SIcLTP for the face and iris and CV.

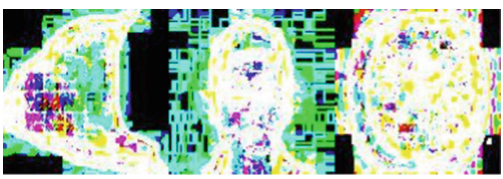


Fig. 4. Sample images of horizontal concatenated features of the three modalities.

To validate the result obtained after applying concatenation of the SIcLTP-extracted features of the face, iris, and CV, LBP is applied to the same data. The sample fused image of LBP-extracted features from the three modalities is shown in Fig. 5.

fridaya@uef.fi

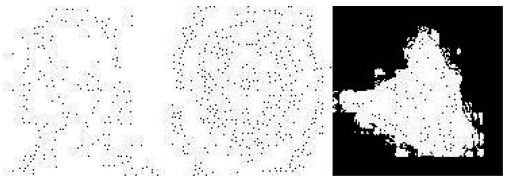


Fig. 5. Sample fused images of LBP extracted features from face, iris, and CV

The recognition accuracies obtained from the ROC analysis after applying the ZSSD for similarity match finding are listed below in Table 3; these results show that the combination/fusion of the face, iris and CV yields higher accuracy when SIcLTP as a feature extractor is applied. This implies that the combination/fusion of the three modalities performs better than the individual modalities taken separately. The ROC curve obtained has been shown in Fig. 6.

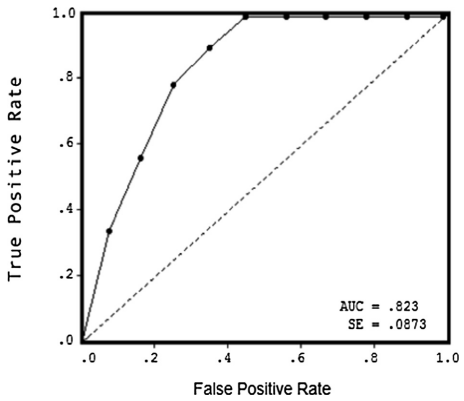


Fig. 6. ROC curve for the fused modalities

Table 3 compares the accuracies obtained when the modalities are independently tested after applying the SIcLTP operator with the accuracy obtained when the modalities are fused. In addition, it includes the recognition accuracy obtained after fusing LBP-extracted features from the face, iris, and CV. The results highlight the effectiveness of the fusion of modalities for multimodal biometrics. These tabulated results are shown graphically in Fig. 7.

Table 3. Comparison of individual modalities with the fused modality.

Modality	Recognition accuracy %	
	LBP	SIcLTP
Face	51.1	82.2
Iris	53.6	82.1
CV	21.43	67.86
Face, Iris, and CV	69.33	87.77

fridaya@uef.fi

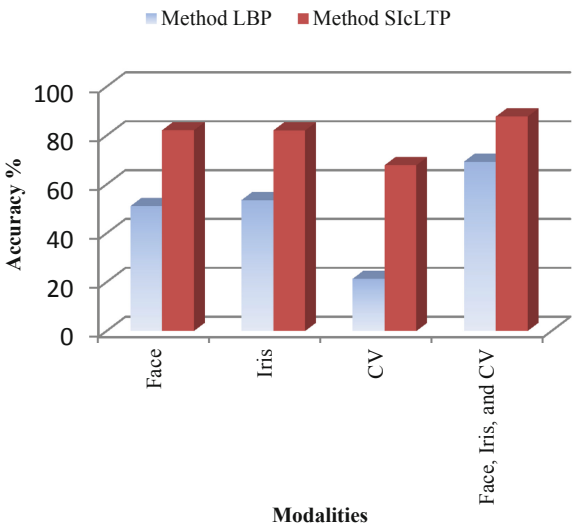


Fig. 7. Comparative chart of accuracy in unimodal and multimodal biometrics

It can be clearly observed that the accuracy of the fused modalities is higher than that of the unimodal results. Therefore, when three modalities are combined, the recognition accuracy is found to improve further; hence, multimodal biometrics can improve recognition accuracy.

Although the result obtained for multimodal traits of the face, iris, and CV is promising, the number of features obtained from all three modalities is significantly increased. Hence, the feature subset selection technique can be considered to reduce the dimensionality.

4 Dimensionality Reduction of the Fused-Feature Space Using Genetic Algorithm

The genetic algorithm (GA) tool in MATLAB 7.1b is used to reduce the large feature space that has been created after the modalities for the face, iris, and CV are fused. Typically, the SIcLTP-extracted features comprise the original feature set.

The following fitness function, which corresponds to the one in [19], has been applied:

$$\text{Fitness} = \frac{\alpha}{N_f} + \exp\left(-N_f^{-1}\right), \tag{4}$$

where

α is the ZSSD based matching error; and
N_f is the cardinality of the selected features.

fridaya@uef.fi

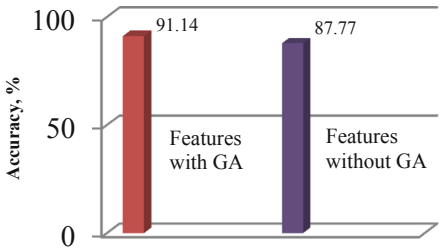


Fig. 8. Comparative accuracy with and without GA

The accuracy values presented in Sect. 4 are compared with the result obtained after applying GA; this is shown in Fig. 8. The result is promising; in particular, it shows an improvement of approximately 5% in accuracy, which is a significant increase. The comparative accuracies for unimodality, multimodality, and fused modality after applying GA are listed in Table 4.

Table 4. Comparative accuracies of all the modalities after GA application

Sl. no.	Modality	Accuracy %
1	Face	82.2
2	Iris	82.1
3	CV	67.86
4	Fused modalities: Face, Iris, and CV	87.77
5	Feature selection/dimensionality reduction using GA for fused modalities	91.14

5 Conclusions and Future Scope

Using the extracted features of individual traits, the fusion of multimodal (trimodal) traits was conducted by concatenation of the extracted features for the multimodal (all three modalities) traits of Face, iris, and CV.

The recognition accuracy obtained by concatenation of face, iris, and CV was better than those of the unimodal results; in particular, a recognition accuracy of 87.77% was achieved, as shown in Table 4. The results in Table 3 clearly demonstrate the accuracy of the SIcLTP operator when compared with the LBP-extracted results.

The concatenation of the fused features leads to a sufficiently large feature space; thus, a GA tool has been used to reduce the dimensionality. To evaluate the effectiveness of the reduced feature space, the recognition accuracy was further evaluated, and the accuracy of the fused modality of the face, iris, and CV after applying the GA wasapproximately 5% better, leading to an accuracy of 91.14% (Table 4).

In this study, we have only been implemented an improved feature extraction technique followed by fusion of the extracted features from the face, iris, and CV to achieve higher recognition accuracy. Factors such as age and major pose variations have not been considered. In particular, the present study was not intended to develop a

fridaya@uef.fi

complete multimodal biometric system; however, we aimed to widen the scope of real-time implementation of such systems in the areas of security, forensics, and medical research.

Moreover, the unavailability of standard data sets for various biometric traits, especially for CV, has also constrained the comparison of recognition accuracies to be universally accepted. Similar to most recent studies, the database in the present study has also been created randomly in an uncontrolled environment.

For wider acceptance of the multimodal biometric system, other modalities having highly discriminative, distinguished, and disjoint feature sets, such as DNA with fingerprints, may be used for further improvement of the recognition accuracy; real-time situations may require exploration of other methods of feature subset selection or dimensionality reduction that have higher computational speed and better resource utilisation.

References

1. Jain, A.K., Ross, A.: Multibiometrics systems. Commun. ACM **47**(1), 34–40 (2004)
2. Jain, A.K., Ross, A.: Multibiometricsystems. Commun. ACM **47**(1), 34–40 (2004). https://doi.org/10.1145/962081.962102
3. Kittler, J., Hatef, M., Duin, R.P.W., Matas, J.: On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell. **20**(3), 226–239 (1998). https://doi.org/10.1109/34.667881
4. Ross, A., Jain, A.: Information fusion in biometrics. Pattern Recognit. Lett. **24**(13), 2115–2125 (2003). https://doi.org/10.1016/S0167-8655(03)00079-5
5. Jain, A., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. Pattern Recogn. **38**(12), 2270–2285 (2005). https://doi.org/10.1016/j.patcog.2005.01.012
6. Sanderson, C., Paliwal, K.K.: Identity verification using speech and face information. Digit. Signal Process. **14**(5), 449–480 (2004). https://doi.org/10.1016/j.dsp.2004.05.001
7. Ross, A., Govindarajan, R.: Feature level fusion using hand and face biometrics. In: Proceeding of SPIE 5779, Biometric Technology for Human Identification II (2005). https://doi.org/10.1117/12.606093
8. Zhou, X., Bhanu, B.: Feature fusion of face and gait for human recognition at a distance in video. In: ICPR 2006, 18th International Conference on Pattern Recognition, 20–24 August 2006, Hong Kong, CHN, IEEE, New York (2006). https://doi.org/10.1109/ICPR.2006.556
9. Singh, S., Gyaourova, A., Bebis, G., Pavlidis, I.: Infrared and visible image fusion for face recognition. In: Proceeding SPIE 5404, Biometric Technology for Human Identification (2004). https://doi.org/10.1117/12.543549
10. Rattani, A., Tistarelli, M.: Robust Multi-Modal and Multi-Unit Feature Level Fusion of Face and Iris Biometrics, Computer Vision Laboratory, University of Sassari. Technical report no. CVL-2008-003, December 2008
11. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, 2nd edn. Wiley, New York (2001)
12. Khatoon, N., Ghose, M.K.: Steady Illumination colour Local Ternary Pattern as a feature extractor in iris authentication. In: Bi, Y., Kapoor, S., Bhatia, R. (eds.) Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016, vol. 2, pp. 966–973. Springer, Cham, CHE (2018). https://doi.org/10.1007/978-3-319-56991-8_72
13. Jain, V., Mukherjee, A.: The Indian face database, IIT Kanpur (2002)
14. Dobes, M., Machala, L.: http://www.inf.upol.cz/iris

fridaya@uef.fi

15. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recogn. **29**(1), 51–59 (1996). https://doi.org/10.1016/0031-3203(95)00067-4
16. Patil, S., Nadar, J.S., Gada, J., Motghare, S., Nair, S.S.: Comparison of various stereo vision cost aggregation methods. Int. J. Eng. Innov. Technol. **2**(8), 222–226 (2013)
17. Khatoon, N., Ghose, M.K.: Conjunctival Vasculature (CV) as a unique modality for authentication, using Steady Illumination colour Local Ternary Pattern (SIcLTP). Int. J. Recent Innov. Trends Comput. Commun. **3**(9), 5673–5677 (2015)
18. Khatoon, N., Ghose, M.K.: Use of illumination invariant feature descriptor for face recognition. Int. J. Comput. Sci. Inf. Secur. (IJCSIS) **16**(3), 118–122 (2018)
19. Babatunde, O., Armstrong, L., Leng, J., Diepeveen, D.: A genetic algorithm–based feature selection. Int. J. Electron. Commun. Comput. Eng. **5**(4), 899–905 (2014)

fridaya@uef.fi

Evaluating Machine Learning Models on the Ethereum Blockchain for Android Malware Detection

Md. Shohel Rana(✉), Charan Gudla, and Andrew H. Sung

School of Computing Sciences and Computer Engineering,
The University of Southern Mississippi, Hattiesburg, MS 39406, USA
{md.rana, charan.gudla, andrew.sung}@usm.edu

Abstract. Android, the most popular mobile operating system, with billions of active users and more than 2 million apps, has motivated advertisers, hackers, fraudsters and cyber-criminals to develop malware of all types for it. In recent years, extensive research has been conducted on malware analysis and detection for Android devices, even though Android has already implemented various security mechanisms to deal with the problem. In this paper, we developed a consortium blockchain network to evaluate various machine learning models for a given malware dataset. A reward is offered using smart contracts as an incentive to the competitors for their work by allowing them to submit solutions through training with selected machine learning models in a secure and trustworthy manner. The analysis of datasets by competitors helps various organizations in the network to enhance or boost their current malware detection or defense tools. The decentralized network provides transparency, enhances security and reduces the cost in managing all relevant data by eliminating third parties. We used DREBIN dataset in the developed framework for initial experiments and the encouraging results are presented.

Keywords: Machine learning · Blockchain · Smart contract · Google · Malware

1 Introduction

Android is a most popular mobile operating system developed by Google, built on a modified version of the Linux kernel in conjunction with additional open source software and designed primarily for touchscreen mobile devices and its architecture is divided into five components with two models of permissions: (i) A sandbox environment at the kernel level which prevent access to the file-system and other

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 446–461, 2019.
https://doi.org/10.1007/978-3-030-22868-2_34

fridaya@uef.fi

resources and (ii) API used to expose to the user during installation of an application [1]. Every application is assembled into an Android Application Package (APK) file contains application code (.dex files), resources, and the AndroidManifest.xml file that is considered as an important element that provides the information of the features and the security configuration of every application (e.g. permissions API, activities, services, content providers and the broadcast receivers). After Decompiling the APK file, the AndroidMenifest.xml (see Fig. 1) file has been experimented how the permissions are used and the Java files where the API functions are invoked in conjunction with the asset and resource files whether there is any dex executable (ELF) image file or any code hiding image script available or not [2–5]:

```
<uses-permission android:name="android.permission.INTERNET"/>
<uses-permission android:name="android.permission.ACCESS_WIFI_STATE"/>
<uses-permission android:name="android.permission.ACCESS_NETWORK_STATE"/>
<uses-permission android:name="android.permission.WAKE_LOCK" />
<uses-permission android:name="android.permission.READ_PHONE_STATE"/>
<uses-permissionandroid:name="android.permission.READ_EXTERNAL_STORAGE"/>
<uses-permission android:name="android.permission.WRITE_EXTERNAL_STORAGE"/>
<uses-permission android:name="android.permission.GET_ACCOUNTS"/>
<uses-permission android:name="com.google.android.gms.permission.ACTIVITY_RECOGNITION"/>
<uses-permission android:name="android.permission.RECEIVE_BOOT_COMPLETED"/>
<uses-permission android:name="android.permission.CHANGE_WIFI_STATE"/>
<uses-permission android:name="android.permission.GET_TASKS"/>
<uses-permission android:name="android.permission.ACCESS_BACKGROUND_SERVICE"/>
<uses-permission android:name="android.permission.RECEIVE_BOOT_COMPLETED"/>
<uses-permission android:name="android.permission.SYSTEM_ALERT_WINDOW"/>
<uses-permission android:name="com.android.vending.BILLING" />
```

Fig. 1. An AndroidMenifest.xml file with used permissions

In the above, red permissions are commonly used in malicious Android apps. It has also been found that some sensitive functions often related to malware, as follows:

fridaya@uef.fi

```
(TelephonyManager) context.getActivity().getSystemService("phone")).getDeviceId()
(TelephonyManager) context.getActivity().getSystemService("phone")).getLine1Number()
(ActivityManager) context.getActivity().getSystemService("activity")
List<RunningAppProcessInfo> procInfos = ((ActivityManager)
context.getActivity().getSystemService("activity")).getRunningAppProcesses();
for (int i = 0; i < procInfos.size(); i++) {
      if(((RunningAppProcessInfo)
      procInfos.get(i)).processName.equals(args[0].getAsString())){
         b = Boolean.valueOf(true);
     }
}
String[] strArr = new String[]{
   "android.permission.ACCESS_COARSE_LOCATION",
   "android.permission.ACCESS_FINE_LOCATION",
   "android.permission.ACCESS_WIFI_STATE",
   "android.permission.CHANGE_WIFI_STATE",
   "android.permission.VIBRATE",
   "android.permission.READ_CALENDAR",
   "android.permission.WRITE_CALENDAR",
   "com.google.android.gms.permission.ACTIVITY_RECOGNITION"
};
IntentFilter intentFilter = new IntentFilter("android.intent.action.PACKAGE_ADDED");
intentFilter.addAction("android.intent.action.PACKAGE_REMOVED");
intentFilter.addAction("android.intent.action.PACKAGE_REPLACED");
intentFilter.addDataScheme("package");
registerReceiver(this.f2955D, intentFilter);
intentFilter = new IntentFilter();
intentFilter.addAction("android.net.wifi.SCAN_RESULTS");
intentFilter.addAction("android.net.wifi.STATE_CHANGE");
intentFilter.addAction("android.net.wifi.WIFI_STATE_CHANGED");
intentFilter.addAction("android.net.wifi.supplicant.CONNECTION_CHANGE");
registerReceiver(this.an, intentFilter);

External links:

public static String f2920C = "http://www.shinhwa21.net/new/apps_judis_end.php?pkg=";
private static String al = "http://www.shinhwa21.net/new/apps_kakao_judis_7.php?pkg=";
```

To deal with the Android malware problem, much research has been done recently. For example, TaintDroid [6], DroidRanger [7] and DroidScope [8] are methods proposed to monitor the runtime behavior of applications. Also, static analysis methods to assess malicious behavior in source code, data or binary files without running the application have been proposed by Kirin [9], Stowaway [10] and Risk-Ranker [11].

In this paper, we evaluate machine learning solutions to the Android malware detection problem by building a consortium blockchain based on the DanKu [12] protocol that allows the initiator to post a dataset, the evaluation criteria, and a reward for the best machine learning model submitted. The proposed system is depicted in Fig. 2 below.

The paper is organized as follows: Sect. 2 describes related works, Sect. 3 describes the proposed malware detection system architecture, Sect. 4 describes the implementation, Sect. 5 describes threat analysis, Sect. 6 describes various machine learning algorithms which have been used in our experiment, Sect. 7 describes methodology includes dataset description and feature extraction, Sect. 8 describes results and analysis, including measurement metrices and comparison, and finally Sect. 9 describes conclusion and future work.

fridaya@uef.fi

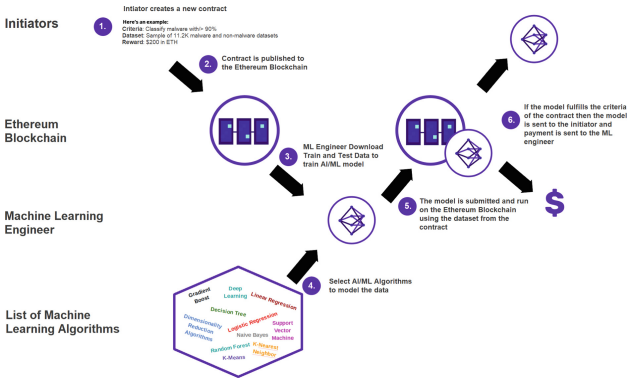


Fig. 2. Overview of proposed framework

2 Related Works

There are many studies in the literature of using machine learning algorithms for malware detection. Gu et al. [13] proposed a multi-feature detection method by constructing a framework named CB-MMDE for detection and classification of malware on Android devices through blockchain technology.

Raje et al. [14] proposed a decentralized firewall built using blockchain technology to classify Portable Executable (PE) files as malicious or benign using a deep belief neural network (DBN) as the detection engine. In order to classify grayscale images into two classes using DBN, the dataset of 10,000 files is used to train the network.

Ouaguid et al. [15] proposed a new framework ANDROSCANREG (Android Permissions Scan Registry) by extracting and analyzing the requested permissions in an Android application through a decentralized and distributed system based on the emerging technology called blockchain. ANDROSCANREG consists of two blockchains, (i) PERMBC, handles analysis, validation and preparation of the raw results so that they can persist in the second Blockchain, (ii) BTCBC, save the permissions history of each version of the applications being scanned via financial transactions. The presented approach can implement different analysis of an Android application allowing it to download and install malicious modules in the device file in order to upload, destroy or encrypt the victim's private data based on the acquisition of the administrator permissions.

Noyes [16] proposed design and implementation of a novel anti-malware environment called BitAV allowing decentralization to update and maintain the software. The feedforward scanning technique significantly enhanced the performance of malware matching system by decomposing the file matching process into efficient interrogations at constant ($O(1)$) time.

Firdaus et al. [17] proposed a bio-inspired method with machine learning to detect root exploit by examining three types of features including system command, directory path, and code-based features, in conjunction with the novel android debug bridge (ADB). This proposed method used four types of boosting algorithms including

fridaya@uef.fi

AdaBoost, real AdaBoost, logit boost, and multiboot for the classification purposes and developed a system called RODS.

Moubarak [18] proposed a new notion by utilizing the blockchain technology to explore the feasibility of new untraceable malwares. This proposed idea developed a 4-ary malware and tested it in real time where each chunk of the code interacts with the bitcoin network to validate and assure whether it belongs to malicious software.

3 Proposed System Architecture

3.1 Basic Structure

In order to ensure trust and integrity, the proposed system is consisted of three phases, see Fig. 3.

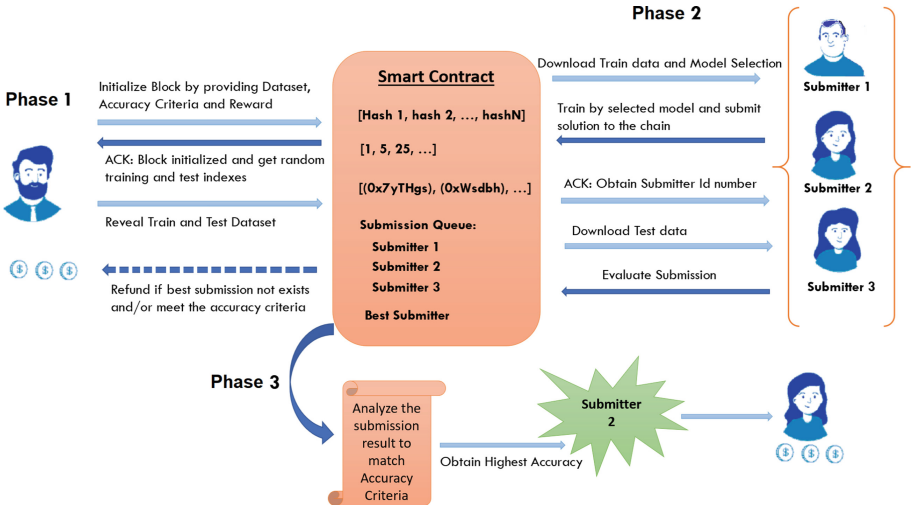


Fig. 3. Structure of the proposed system

- **Phase 1:** Initiator (Alice) initializes the genesis block by revealing a dataset defined with train and test ratio, an evaluation criterion that must be fulfilled by the miners or machine learning engineers, and a reward amount typically denominated in cryptocurrency for best model submitter to the Ethereum smart contract. The Ethereum smart contract provides a list to machine learning algorithms to model the data and evaluate the model for classification tasks.
- **Phase 2:** Miners or machine learning engineers acting as submitter download the training set provided by an initiator, and train with the selected machine learning algorithm to build a model. After successful evaluation of the model, submit their solutions to the blockchain. The smart contract ensures the submission by providing *Submitter_ID* to uniquely identify the submission results into the Blockchain.

fridaya@uef.fi

- **Phase 3:** After finalizing the competition, the blockchain analyzes the submission results and selects a winner of the competition who meets the accuracy criteria.

3.2 Definitions

- **user** is anyone who can interact with Ethereum contracts.
- **initiator** is a user who initializes the blockchain with genesis file and create smart contracts.
- **smart contract** is an Ethereum contract.
- **submitter (miner or machine learning engineer)** is a user who submits solutions to the blockchain for a reward.
- **period** is a timeframe needed to mine a block.
- **data point** is made up of input(s) and prediction(s).
- **data group** is a matrix made up of data points.
- **hashed data group** is the hash of a data group including random nonce.
- **contract wallet address** is an account that hold the reward amount until the finalization of the contract.

3.3 Features of Smart Contracts

- ***Initialization()***, using this function Bob will create a genesis block with some parameters including Dataset, Accuracy criteria and a reward amount.
- ***IsBlockInitialized()*** is responsible whether the contract is created, initialized and returns two arrays containing Training and Test indexes of the dataset.
- ***RevealData()*** splits the dataset into training and testing sets by creating two arrays containing the indexes generated randomly and reveals to the chain.
- ***DownloadTrainDataAndMLASelection()*** is used by miners/submitters to get the data revealed in previous step and returns two arrays (i) list of Machine Learning Algorithms, and (ii) list of Training Data Indexes.
- ***SubmitSolution()*** is invoked after modeling the data using training indexes with one of the machine learning algorithms and returns *Submitter_ID* as acknowledgement.
- ***EvaluateSubmission()*** takes submitter id as input parameter to evaluate the model from previous step and finally returns the evaluation result (Accuracy, Precision, Recall, F1-score, etc.) as an array.
- ***SubmissionQueue()*** is responsible to store the submission information into the blockchain along with *Submitter_ID*.
- ***BestSubmitter()*** retrieves submitted information from the blockchain to analyze accuracy criteria and returns *Submitter_Address* and *Submitter_ID*.
- ***ContractFinalized()*** To finalize the competition, this function is used with returning Boolean Value (True or False).
- ***CancelContract()*** is used to cancel the contract if any exception is occurred.
- ***RewardOrRefund()*** This function is used to send reward amount to the best submitter and/or to refund the reward amount to the initiator if no one fulfills the conditions.

fridaya@uef.fi

- ***GetTrainingIndex()*** is used to retrieve training indexes.
- ***GetTestingIndex()*** is used to retrieve testing indexes.
- ***GetSubmissionId()*** is used to retrieve submission id.

4 Implementation

4.1 Dataset Hashing

The initiator who creates the contract, splits the dataset into data groups. Nonce is a number generated randomly along with data groups and it is used to identify each data group. The initiator hashes these data groups using SHA256 algorithms.

4.2 Determine Train and Test Dataset

During Initialization in phase 1, the initiator calls the function *Initialization()* which uses previously mined block hash number as a seed for randomization. The dataset groups are randomly selected using nonce. The dataset is split into 80% for training and 20% for testing. We repeat the procedure until all data group indexes are selected. Algorithm used in DanKu protocol (in solidity) for randomly selecting hashes:

```
function randomly_select_index(uint[] array) private {
  uint t_index = 0;
  uint array_length = array.length;
  uint block_i = 0;
  // Randomly select training indexes
  while(t_index < training_partition.length) {
    uint random_index =
uint(sha256(block.blockhash(block.number-block_i))) %
array_length;
    training_partition[t_index] = array[random_index];
    array[random_index] = array[array_length-1];
    array_length--;
    block_i++;
    t_index++;
  }
  t_index = 0;
  while(t_index < testing_partition.length) {
    testing_partition[t_index] = array[array_length-1];
    array_length--;
    t_index++;
  }
}
```

fridaya@uef.fi

5 Competition Rules

The proposed system has been designed in such a way that no participants can cheat or get advantages over other users (e.g. the initiators, the submitters, and/or machine learning engineers of the Ethereum Blockchain).

5.1 Overfitting by Submitter

In order to prevent overfitting problem, the test dataset is kept secret until the submission period ends. Because, submitters or machine learning engineers can overfit their selected machine learning model if they have the access to test dataset. The smart contract evaluates the submission using test dataset.

5.2 Too Many Submissions

The competitor cannot submit solution more than once, because after first successful submission the smart contract save the information to the chain with his/her address along with submission id.

5.3 Block Hash Manipulation by Miners

A miner does not have the full control over the working procedure of training indexes selection. After observing which data groups are going to be selected, the initiator can take decision whether to mine the block or not which could result in disagreeable training indexes.

5.4 Distributed Reward System Abuse

To prevent malicious submitters from re-submitting the similar solution by making small changes to the original submitter solution and calling the evaluation function before the original submitter, a distributed reward system may de-incentivize submitters. Due to this reason, it ensures that the best submitter only be paid. The first submitter will be rewarded if there are two submissions with the same solution and evaluated.

6 Machine Learning Algorithms

This section describes basics of machine learning algorithms that have been used in this experiment as follows:

- **Decision Tree (DT):** Builds classification models in the form of a tree structure by breaking down a dataset (categorical and/or numerical data) into smaller subsets and the result is a tree with decision nodes having two or more branches and leaf nodes that represent a classification or decision. The topmost decision node also known as root node in a tree which corresponds to the best predictor [19].
- **Random Forest (RF):** A supervised learning algorithm trained with the *Bagging* method that builds multiple decision trees and merges them together to get a more accurate and stable prediction [20].

fridaya@uef.fi

- **Extremely Randomized Tree (ERT):** With respect to random forests, this method drops the idea of using bootstrap copies of the learning model without trying to find an optimum cut-point for each one of the *K* randomly chosen features at each node [21].
- **Gradient Boosted (GB):** An ensemble machine learning algorithm works for both regression and classification problems using boosting technique, combining several weak learners to form a strong learner [22].
- **Support Vector Machine (SVB):** A discriminative classifier defined by a separating hyperplane and the algorithm outputs an optimal hyperplane which categorizes new examples. In 2D space this hyperplane is a line dividing a plane in two parts where in each class lay in either side [23].
- **Neural Networks (NN-MLP):** The multilayer perceptron is a feedforward artificial neural network model that consists of multiple layers where each of the layer is fully connected to the next layer and maps sets of input data onto a set of appropriate outputs. The nodes of the layers are neurons using nonlinear activation functions, except for the nodes of the input layer. There can be one or more non-linear hidden layers between the input and the output layer [24].
- **Naïve Bayes (NB):** A supervised classification algorithm for solving binary or multi-class classification problems and for predictive modeling by calculating the probabilities for each factor and then selecting the outcome with highest probability using Bays theorem [25].
- **k-Nearest Neighbors (k-NN):** A simplest machine learning algorithms that can be used for both classification and regression predictive problems. In a classification task, the output is calculated as the class with the highest frequency from the *K*-most similar instances where each instance votes for their class and the class having most votes is taken as the prediction and class probabilities is calculated as the normalized frequency of samples that belong to each class in the set of K most similar instances for a new data instance [26].
- **Discriminant Analysis (DA):** Discovers a set of prediction comparisons based on independent variables that are used to classify entities into groups with having are two possible objectives, (i) finding a predictive equation for classifying new individuals, (ii) interpreting the predictive equation to better understand the relationships exists among the variables [27].
- **Logistic Regression (LR):** The appropriate regression analysis to conduct when the dependent variable (target) is categorical (e.g. to predict whether an email is spam ***1*** or ***0***) and used to describe data used when the dependent variable (target) is categorical and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables [28].

7 Methodology

In this experiment we have used 'DREBIN' [29] dataset in which 11,120 of 123,453 are real Android applications from 179 different malware families, 5,560 applications of real malware samples and 5,560 applications of real benign samples. The samples were collected in the period of August 2010 to October 2012. An overview of the top

fridaya@uef.fi

20 malware families in our dataset is provided in Table 1 including numerous feature families: api_call, feature, url, service_receiver, permission, call, intent, real_permission, activity, provider. Note that only top 20 families are exposed, and our dataset contains number of entries in each class of malware (values above 40).

Table 1. Top malware families of our dataset (values above 40)

Malware family	# Entries	Malware family	# Entries
FakeInstaller	925	Adrd	91
DroidKungFu	667	DroidDream	81
Plankton	625	ExploitLinuxL otoor	70
Opfake	613	Glodream	69
Ginmaster	339	MobileTx	69
BaseBridge	330	FakeRun	61
Iconosys	152	SendPay	59
Kwin	147	Gappusin	58
FakeDoc	132	Imlog	43
Geinimi	92	SMSreg	41

8 Results and Analysis

8.1 Measurement Metrices

Confusion Matrix: A confusion matrix is a matrix which contains information about actual and predicted classifications to measure the performance of algorithm using the matrix data [30, 31]. Table 2 represents confusion matrix.

Accuracy (AC) is the proportion of the total number of corrected predictions. Overall, how often is the classifier correct?

Table 2. Confusion matrix

		Actual class	
		Positive	Negative
Predictive class	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

$$accuracy\,(AC) = \frac{No.of\ correctly\ classified\ data\ (TP + TN)}{Total}$$

Precision (P) is the proportion of the correctly predicted positive cases determined by

$$precision = \frac{TP}{TP + FP}$$

fridaya@uef.fi

Recall or True Positive rate (TP) is the proportion of the correctly identified positive cases defined by

$$recall = \frac{TP}{TP + FN}$$

False Positive rate (FP) is the proportion of negatives cases that were incorrectly classified as positive, defined by

$$false\,positive\,rate = \frac{FP}{TP + FP}$$

f1- Score or F-Measure is a weighted average of the True Positive (TP) rate or recall and Precision (P) defined by

$$F = \frac{\left(\beta^2 + 1\right) * P * TP}{\beta^2 * P + TP}$$

Where β has a value from 0 to infinity and is used to control the weight assigned to TP and P.

ROC Curve is a graph to summarize the performance of the classifier over all probable thresholds generated by plotting the True Positive (TP) Rate in Y-axis against the False Positive (FP) Rate in X-axis.

8.2 Results

We compared the results of machine learning classifier submitted by the competitors or miners in a simulated environment. More than one competition is held, and their respective accuracy criteria are recorded (see Figs. 4 and 5). We conducted 4 competitions for our experimental purpose which is described in Table 3.

Table 3. Result of the performance

Competition	Submitter	Algorithms	Precision		Recall		f1-score		Accuracy
			0	1	0	1	0	1	
COMP1 With Accuracy Criteria 80%	1	k-NN	0.92	0.89	0.89	0.92	0.90	0.91	90.47
	2	DT	0.93	0.88	0.88	0.94	0.91	0.91	91.78
	3	LR	0.78	0.84	0.86	0.75	0.82	0.80	80.94
	4	NB	0.68	0.52	0.15	0.92	0.25	0.66	53.60
	5	NN (MLP)	0.85	0.91	0.92	0.83	0.88	0.87	87.54
COMP2 with Accuracy criteria 85%	1	ERT	0.93	0.93	0.93	0.93	0.93	0.93	93.66
	2	DA	0.76	0.88	0.91	0.70	0.83	0.78	80.71
	3	SVM	0.91	0.91	0.91	0.91	0.91	0.91	90.74
	4	GB	0.87	0.88	0.89	0.86	0.88	0.87	87.50
	5	k-NN	0.92	0.89	0.89	0.92	0.90	0.91	90.47

(*continued*)

fridaya@uef.fi

Table 3. (*continued*)

Competition	Submitter	Algorithms	Precision		Recall		f1-score		Accuracy
			0	1	0	1	0	1	
COMP3 with Accuracy criteria 90%	1	DA	0.76	0.88	0.91	0.70	0.83	0.78	80.71
	2	GB	0.87	0.88	0.89	0.86	0.88	0.87	87.50
	3	SVM	0.91	0.91	0.91	0.91	0.91	0.91	90.74
	4	LR	0.78	0.84	0.86	0.75	0.82	0.80	80.94
	5	RF	0.95	0.94	0.94	0.95	0.94	0.94	94.33
COMP4 with Accuracy criteria 95%	1	RF	0.95	0.94	0.94	0.95	0.94	0.94	94.33
	2	ERT	0.93	0.93	0.93	0.93	0.93	0.93	93.66
	3	DT	0.93	0.88	0.88	0.94	0.91	0.91	91.78
	4	NN (MLP)	0.85	0.91	0.92	0.83	0.88	0.87	87.54
	5	SVM	0.91	0.91	0.91	0.91	0.91	0.91	90.74

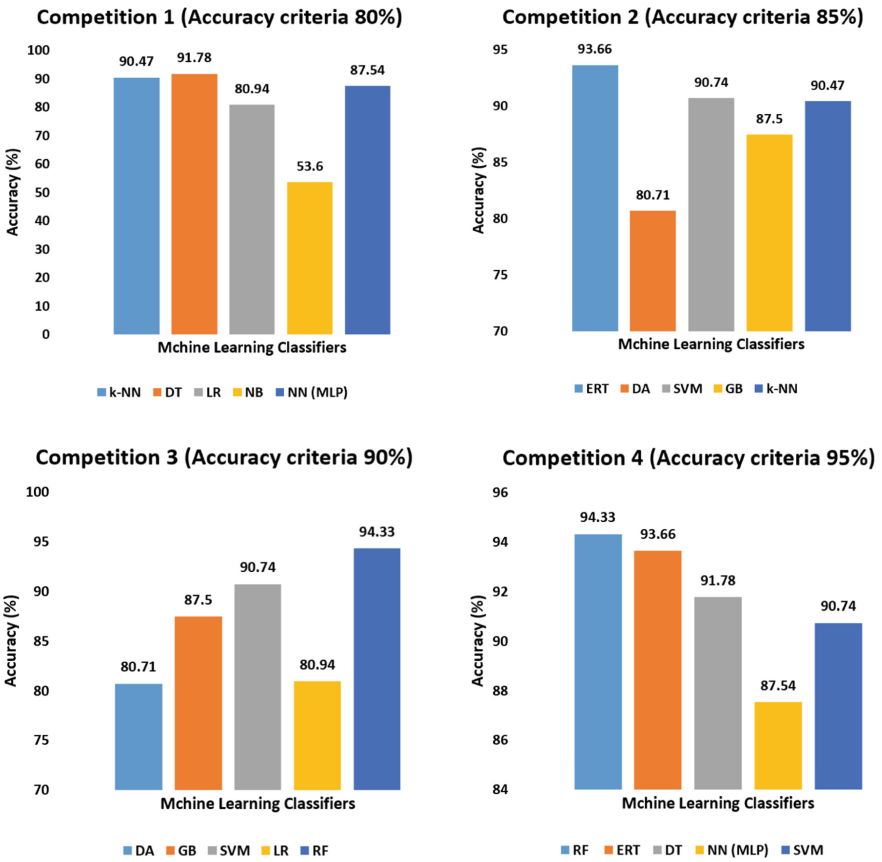


Fig. 4. Accuracy curves

fridaya@uef.fi

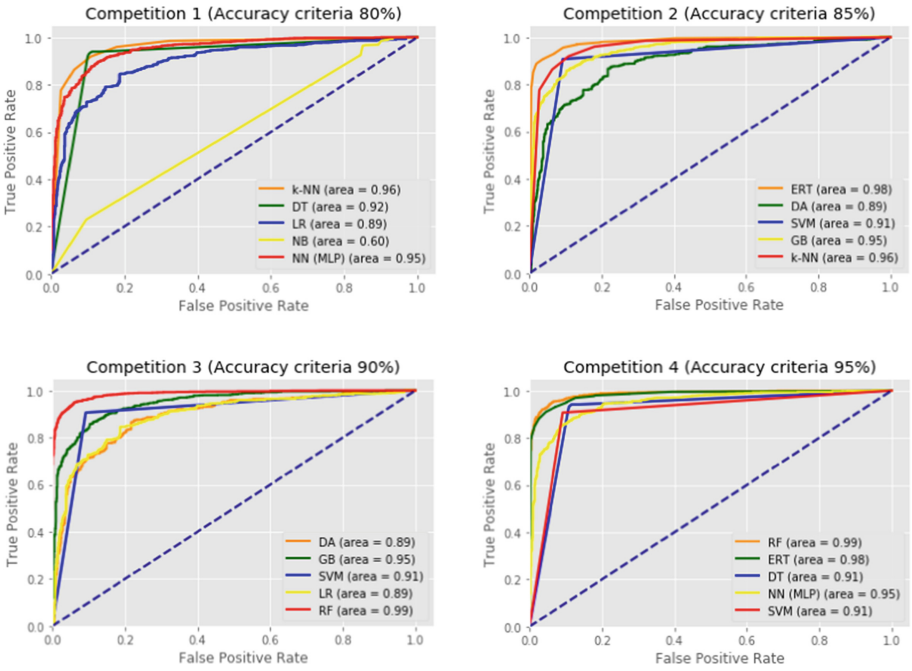


Fig. 5. ROC curves

Finally, we obtained our best outcome by running several trials, using 80% for training and 20% for testing, done under the top 100 most frequent API calls and system permissions applied with information gain. During the experiment we conducted a competition where 5 competitors can take place to submit their trained model to the Blockchain. We obtain various results if we see each competition, for example, in the first competition that assign the accuracy criteria (80%), so anyone can achieve the reward who will obtain more than or equal to accuracy criteria, here *Submitter_2* achieves the reward by obtaining the highest accuracy 91.78% among the others by using Decision Tree classifier. Similarly, in the second competition *Submitter_1* is rewarded who obtained accuracy of 93.66% using Extremely Randomized Tree with fulfilling the accuracy criteria (85%). In the third competition, *Submitter_5* is rewarded by obtaining 94.33% using Random Forest classifier with fulfilling the accuracy criteria (90%). But, in terms of fourth competition, No one is rewarded due to failing to fulfill the accuracy criteria defined 95%, so the reward amount will be back to the initiator who created or established competition. We can define the "BEST" outcome as the experiment that maximizes the difference between the TPR and the FPR since we have achieved higher TPR and lower FPR shown in Table 3.

fridaya@uef.fi

9 Conclusion and Future Work

In this paper, for a given dataset we evaluated various machine learning models in a consortium blockchain network. An incentive is awarded to the best machine learning model submitter. The analysis of dataset by competitors helps various organizations in the network to enhance or boost their current malware detection tools. The decentralized network provides transparency, enhances security and reduces the cost in managing all relevant data by eliminating third parties.

For future work, we develop a real-time malware detection system with blockchain integration to detect and prevent unknown malicious attacks on a network. Also, we will conduct more experiments with additional datasets and other machine learning algorithms to improve malware detection.

Acknowledgment. The authors wish to acknowledge the valuable help received from Besir Kurtulmus, Algorithmia Inc., for his guidance on technology and domain knowledge pertaining to applying machine learning within blockchain.

References

1. Drake, J.J., Lanier, Z., Mulliner, C., Fora, P.O., Ridley, S.A., Wicherski, G.: Android Hacker's Handbook. Wiley, Indianapolis (2014)
2. Rana, M.S., Sung, A.H.: Malware analysis on android using supervised machine learning techniques. Int. J. Comput. Commun. Eng. **7**(4), 178–188 (2018)
3. Rana, M.S., Rahman, S.S.M.M., Sung, A.H.: Evaluation of tree based machine learning classifiers for android malware detection. In: Nguyen, N., Pimenidis, E., Khan, Z., Trawiński, B. (eds.) Computational Collective Intelligence. ICCCI 2018. Lecture Notes in Computer Science, vol. 11056. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98446-9_35
4. Rana, M.S., Gudla, C., Sung, A.H.: Android malware detection using stacked generalization. In: Proceeding of 27th International Conference on Software Engineering and Data Engineering, pp. 15–19 (2018)
5. Rana, M.S., Gudla, C., Sung, A.H.: Evaluating machine learning models for android malware detection – a comparison study. In: Proceeding of International Conference on Network, Communication, and Computing, Taipei, Taiwan (2018)
6. Enck, W., Gilbert, P., Chun, B., Cox, L.P., Jung, J., McDaniel, P., Sheth, A.: Taintdroid: an information-flow tracking system for real-time privacy monitoring on smartphones. In: Proceeding of USENIX Symposium on Operating Systems Design and Implementation (OSDI), pp. 393–407 (2010)
7. Zhou, Y., Wang, Z., Zhou, W., Jiang, X.: Hey, you, get off my market: detecting malicious apps in official and alternative android markets. In: Proceeding of Network and Distributed System Security Symposium (NDSS) (2012)
8. Yan, L.K., Yin, H.: Droidscope: seamlessly reconstructing OS and dalvik semantic views for dynamic android malware analysis. In: Proceeding of USENIX Security Symposium (2012)
9. Enck, W., Ongtang, M., McDaniel, P.D.: On lightweight mobile phone application certification. In: Proceeding of ACM Conference on Computer and Communications Security (CCS), pp. 235–245 (2009)

fridaya@uef.fi

10. Felt, A.P., Chin, E., Hanna, S., Song, D., and Wagner, D.: Android permissions demystified. In: Proceeding of ACM Conference on Computer and Communications Security (CCS), pp. 627–638 (2011)
11. Grace, M., Zhou, Y., Zhang, Q., Zou, S., Jiang, X.: Risk-ranker: scalable and accurate zero-day android malware detection. In: Proceeding of International Conference on Mobile Systems, Applications, and Services (MOBISYS), pp. 281–294 (2012)
12. Kurtulmus, A.B., Daniel, K.: Trustless Machine Learning Contracts; Evaluating and Exchanging Machine Learning Models on the Ethereum Blockchain, Algorithmia Research (2018). https://algorithmia.com/static/documents/d3a4c04/Machine-Learning-Models-on-the-Ethereum-Blockchain.pdf. Accessed 18 Sept 2018
13. Gu, J., Sun, B., Du, X., Wang, J., Zhuang, Y., Wang, Z.: Consortium blockchain-based malware detection in mobile devices. In: IEEE Access, vol. 6, pp. 12118–12128 (2018). https://doi.org/10.1109/access.2018.2805783
14. Raje, S., Vaderia, S., Wilson, N., Panigrahi, R.: Decentralised firewall for malware detection. In: 2017 International Conference on Advances in Computing, Communication and Control (ICAC3), pp. 1–5 (2017)
15. Ouaguid, A., Abghour, N., Ouzzif, M.: A novel security framework for managing android permissions using blockchain technology. Int. J. Cloud Appl. Comput. (IJCAC) **8**(1), 55–79 (2018)
16. Noyes, C.: BitAV: Fast Anti-Malware by Distributed Blockchain Consensus and Feedforward Scanning, CoRR, abs/1601.01405 (2016)
17. Firdaus, A., Anuar, N.B., Razak, M.F., Hashem, I.A., Bachok, S., Sangaiah, A.K.: Root exploit detection and features optimization: mobile device and blockchain based medical data management. J. Med. Syst. **42**, 1–23 (2018)
18. Moubarak, J., Filiol, E., Chamoun, M.: Developing a K-ary malware using Blockchain. https://arxiv.org/abs/1804.01488. Accessed 20 Oct 2018
19. Decision Tree – Classification. https://www.saedsayad.com/decision_tree.htm. Accessed 20 Oct 2018
20. Towards Data Science | The Random Forest Algorithm. https://towards-datascience.com/the-random-forest-algorithm-d457d499ffcd. Accessed 20 Oct 2018
21. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. Mach. Learn. **63**(1), 3–42 (2006)
22. A Comprehensive Guide to Ensemble Learning. https://www.analyticsvidhya.com/-blog/2018/06/comprehensive-guide-for-ensemble-models/. Accessed 20 Oct 2018
23. Towards Data Science | Support Vector Machine - Introduction to Machine Learning Algorithms. https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47. Accessed 20 Oct 2018
24. Neural Networks with Scikit. https://www.python-course.eu/neural-networks-with-scikit.php. Accessed 20 Oct 2018
25. Naive Bayes for Machine Learning. https://machinelearningmastery.com/naive-bayes-for-machine-learning/. Accessed 20 Oct 2018
26. K-Nearest Neighbors for Machine Learning. https://machinelearningmastery.com/k-nearest-neighbors-for-machine-learning/. Accessed 20 Oct 2018
27. Discriminant Analysis. https://ncss-wpengine.netdna-ssl.com/wp-content/themes/nc-ss/pdf/Procedures/NCSS/Discriminant_Analysis.pdf. Accessed 20 Oct 2018
28. Towards Data Science | Logistic Regression - Detailed Overview. https://towards-datascience.com/logistic-regression-detailed-overview-46c4da4303bc. Accessed 20 Oct 2018

fridaya@uef.fi

29. Arp, D., Spreitzenbarth, M., Hubner, M., Gascon, H., Rieck, K.: DREBIN: effective and explainable detection of android malware in your pocket. In: NDSS, vol. 14, pp. 23–26, USA (2014)
30. Confusion Matrix. http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion-matrix/confusion-matrix.html. Accessed 20 Oct 2018
31. Simple guide to confusion matrix terminology. http://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/. Accessed 20 Oct 2018

fridaya@uef.fi

Blockchain-Based Authentication for Network Infrastructure Security

Chi Ho Lau, Kai-Hau Yeung(✉), and Kai Man Kwong

Department of Electronic Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong SAR, China
p420lch2005@gmail.com, eeayeung@cityu.edu.hk, kmkkwong2-c@my.cityu.edu.hk

Abstract. Network infrastructure devices form the backbone of a network. Any computer connecting to the network can communicate with these devices by running an appropriate software. Because of this, hackers could easily attack the network infrastructure. The reason why such attacks is possible is that there is no authentication mechanism in most of today's networks. To solve this, this paper proposes a solution to the problem by introducing an authentication scheme. Blockchain technology is used to store authentication information because malicious tempering is difficult. Authenticated Devices Configuration Protocol (ADCP) is proposed to configure and form a secure overlay network infrastructure automatically. A working implementation has been built to validate the authentication design. Experimental results have also been presented that show the validation.

Keywords: Blockchain · Network infrastructure · Network security

1 Introduction

Security problem is always a concern for computer networks. Many security solutions have been proposed to a secure network infrastructure. One of the widely used solution is Public-Key infrastructure (PKI) [1]. However, deployment and management cost for PKI environment in LAN is high. Besides, this solution is based on asymmetric cryptography which is computational expensive [2]. Modern personal computers might possess such computational power. However, such level of computational power is usually absent in network infrastructure devices such as routers and switches. This problem is discouraging network administrator from using it. Without authentication on network infrastructure devices, any malicious device can act as a network infrastructure device. Attacks like packet dropping or malicious routing against these devices can jeopardies confidentiality, integrity and availability [3, 4].

To protect a network, isolation is the most straightforward solution. It is to isolate different types of devices and limit their connectivity. An example is NI-Switches [5]. However, network admin needs to scrutinize all devices and store their identities. The management and configuration overhead are frustrated. Alternative solution for authentication without using PKI are often rely on central server such as kerberos [6].

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 462–473, 2019.
https://doi.org/10.1007/978-3-030-22868-2_35

fridaya@uef.fi

This kind of solution might suffer from single point of failure. If the authentication server down, devices can no longer confirm the identity of each other.

Blockchain technology [7] has been adapted to store all authentication information across authentication devices. The basic unit of blockchain is block. Each block can store part of the digital asset. The integrity of the blockchain is guarantee by the digital fingerprints. Each block contained two digital fingerprints, its own fingerprint and previous block's fingerprint. Blocks are linked up chronologically using these digital fingerprints to form a chain. Since every block has the digital fingerprint of the previous block, malicious tempering is very difficult. Blockchain will be distributed across devices which can be validate by cross reference check.

In this paper, we therefore present a solution to construct a secure overlay network infrastructure. All network devices can be protected by the solution. Since the digital assets in blockchain is difficult to temper with, it is suitable to store device identities. The blockchain will only be distributed among the overlay network infrastructure. All traffics within the overlay network may be encrypted by symmetric cryptography based on the identity information. This enhances confidentially without requiring expensive computational power. The configuration of the overlay network infrastructure will also be handled by the system itself automatically to reduce the deployment cost.

2 System Overview

In this paper we propose a solution to solve the problem of authenticating devices among network infrastructure devices. Network infrastructure devices such as routers and switches have no effective authentication scheme to prevent malicious access. When these devices cannot verify each other's identity, this enables intruder to sniff information or gain access to the network infrastructure. In small scale networks, the network administrator can keep tracking of the physical addresses of all connected devices. A white listed network can be used to prevent such intrusion. However, keep tracking all network devices in a large-scale network is costly. Also, many low-layer network service protocols do not adopt authentication scheme, it could be forged easily and jeopardizing the network. To address these issues, we proposed an authentication scheme working on network infrastructure devices. The scheme utilizes Blockchain technology to store authentication information across the network.

The authentication scheme relies on an offline hardware authenticator (HA) which is physically kept by the network administrator. HA is a standalone device which accepts wired and wireless connections. The program running in the HA would accept authentication requests. Network administrator can monitor connected devices and authenticate them by triggering the API as provided by the HA. During the authentication process, HA will generate authentication record which include the identity of the authenticated device. Such identity can also be used to examine the traffics. For example, BPDU frames should only come from a device as identified as a switch.

All authentication records will be stored in every authenticated device. To protect authentication records against forgery, we use Blockchain technology to store records. Confirmed authentication records will be sealed within blocks. Each block contains the hash value of previous block. This structure creates a chronological chain of blocks.

fridaya@uef.fi

Since all authenticated devices has a copy of the blockchain database, authentication record is traceable and difficult to forge.

This solution aims to build a secured overlay network infrastructure which connects only authenticated devices (Fig. 1). To achieve this, all authenticated infrastructure devices need to identify themselves during communication. Any authenticated infrastructure device can confirm the identity of the sender with its own blockchain database. The incoming traffic will be discarded if this identity declaration is missing or invalid. All communication may also be encrypted based on the identity information. Such design can protect all network infrastructure devices against cyber-attacks and prevent intruder sniffing the network.

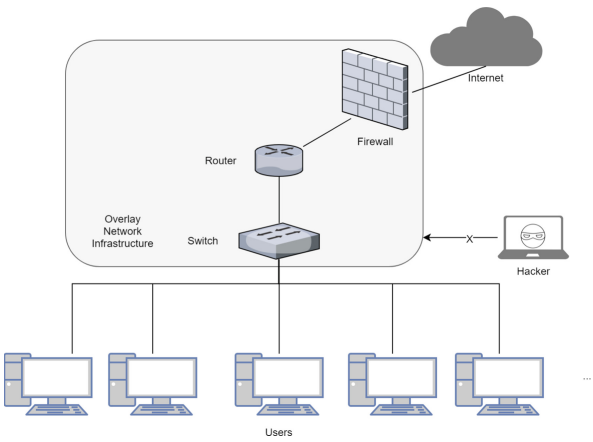


Fig. 1. Overview of the overlay network infrastructure

3 System Design and Mechanism

3.1 Authentication Scheme

Hardware authenticator (HA) is the most important component in the authentication scheme. To get authenticated, network devices need to communicate with the HA first. HA accepts wireless and wired connection. Once connected, an authentication request will be sent to HA. Network administrator can monitor these requests and configure the device identity. Device identity can be used to determine which network protocols should be trusted when received from an authenticated device. Being authenticated by the HA is the only way to join the secured overlay network infrastructure.

The authentication scheme is being supported by a coin-based blockchain system. Every network device adapts this scheme have a pair of public and private keys. These keys are generated using Elliptic Curve Digital Signature Algorithm (ECDSA) [8]. The public key is the account number to accept coins. Devices that own coins is considered authenticated. Determine the authentication status of a network device is through examine the total number of coins associate with the public key it owns. This public

fridaya@uef.fi

key is also act as device identifier (Id). The number of coins will also be used to determine the device identity. Device identity is important to determine which protocol should be accepted from a network device.

Some rules were made to the coin based blockchain system to make sure HA is the authentication authority. The only device that can transfer coins is HA. Authenticated devices cannot transfer its coins to another device. This rule prohibits authentication status transfer. No coins offering is allowed except the Genesis block which is the first block of the blockchain database. The genesis block is generated by the HA when the system starts. It contains a coin offering record to HA. All blocks behind genesis block should not contain any other coin offering record.

3.2 Authentication Process

Once a network device connected to the HA, it sends an authentication request immediately. The blockchain database will be delivered to the target network device when HA received confirmation from the network administrator. The blockchain database included all previous authentication records. Genesis block is the first block of the blockchain database. The only transaction record in genesis block is a coin offering record to HA. This transaction declares the authority of HA because the blockchain system does not generate new coins afterwards. Another important element in the genesis block is the Blockchain public key (Kc) (see Fig. 2). This key is important for overlay network formation and communication.

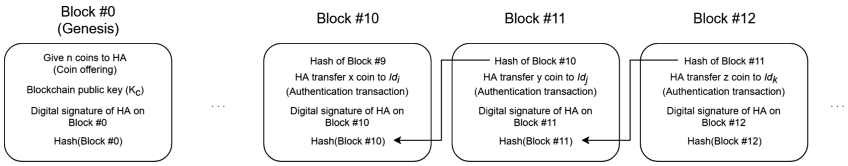


Fig. 2. Structure of blockchain

Next, the HA will sign an authentication transaction to the target device. The number of coins being transferred in the transaction determined the device identity of the target device. This transaction called Authentication transaction (AT) which is a declaration of coin transfer which is protected by the private key of HA. To confirm the transaction, HA need to generate a new block. HA is the only device that can generate new block. A digital certificate is included in every block to make sure that HA generates and confirms the authentication. All other network devices should accept blocks when the issuer of digital certificate is HA (see Fig. 2).

3.3 Communication Between Authenticated Devices

Unauthenticated devices cannot communicate with authenticated devices because any communication attempts will be discarded (see Fig. 3). The communication between authenticated devices require identity clarification. The identity information is included

fridaya@uef.fi

in the communication. Symmetric encryption based on the device identifier (Id) could be adopted to protection the communication from third party sniffing. In that case, all communication from an authenticated device will be encrypted, therefore unauthenticated device is difficult to intercept their communication. Traffics from authenticated device will be examined before passing to user application. A software firewall is included in our software suite to perform filtering, encrypting and decrypting traffics. If the incoming traffic type mismatched the device identity that associate with the device identifier (Id), those will also be discarded by the software firewall.

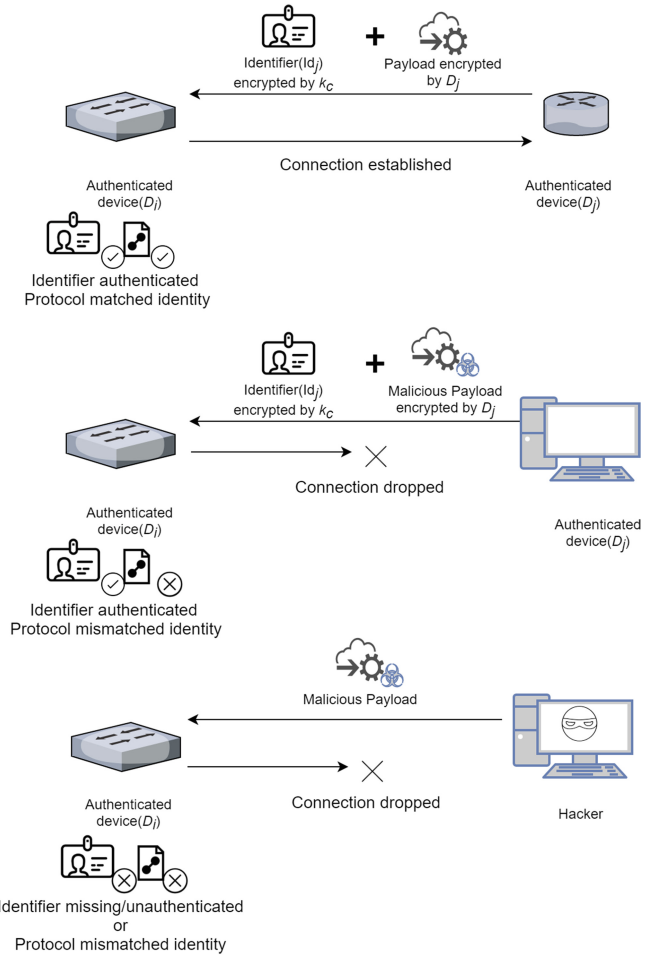


Fig. 3. Communication scenarios for authenticated and unauthenticated network devices

fridaya@uef.fi

3.4 Overlay Network Infrastructure Formation

After authentication process, devices will disconnect from Hardware Authenticator (HA) and enter defensive mode. Under defensive mode, authenticated device accepts the configuration protocol that we will be discussed later. Any other traffics will be discard under this mode. The latest block contains the authentication information of the freshly authenticated device. Authenticated device must communicate with at least one existing authenticated device and broadcast the latest block among the overlay network infrastructure to get accepted as a member. Authenticated Devices Configuration Protocol (ADCP) is proposed to configure authenticated device to form a secure overlay network automatically.

Figure 4 is the timing diagram of ADCP when a freshly authenticated device (D_i) joining the overlay network infrastructure through authenticated device (D_j). As mentioned, fresh authenticated device entered defensive mode which blocks all communication ports. D_i broadcasted the hashed Blockchain public key (Kc) among the entire LAN. Both authenticated and unauthenticated devices that connected to this LAN can receive this message. In this case, D_j received this broadcast. D_j computes the hash value of its own Kc and compared with the hash value from the received broadcast. If they matched, a random key (Kt) was generated. Kt and its device identifier (Id_j) was encrypted with Kc as the unicasted respond to D_i. Then, D_i will try to decrypt the respond using Kc resulting Id'_j and $k't$. D_i will verify $Id'j$ with its own blockchain database to check whether the device identifier is authenticated or not. If $Id'j$ is authenticated, TCP port 6001 will be opened and establish TCP connection with D_j. Once established, D_i will send its own IP address to D_j which encrypted by Kt. D_j received and decrypted the message to get the IP address of D_i. If the decrypted IP address matched with source IP address from the previous TCP connection, D_j will accept the connection and receive new blocks from D_i. Kt will be expired after used or after a period without using. After ADCP communication with another authenticated device, D_i exit defensive mode and accept encrypted communication from any other ports.

Figure 5 shows the timing diagram for unsuccessful ADCP communication. Unauthenticated device (Dx) received the ethernet broadcast from fresh authenticated device (Di). Dx tries to generate unauthenticated device identifier (Idx) to fool the authenticated device. However, Dx does not have genuine Kt to encrypt the message back to Di. The decryption on Di will be failed and no respond will be given to Dx resulting waiting until timeout. The design of ADCP allow authenticated devices connect each other automatically in a secure way. Afterwards, the secured overlay network was formed which accepts communication from authenticated devices only. The system can protect the machine from cyber-attack. Even there is preexisting exploit; attacker cannot use it through internet because attacker is unauthenticated.

fridaya@uef.fi

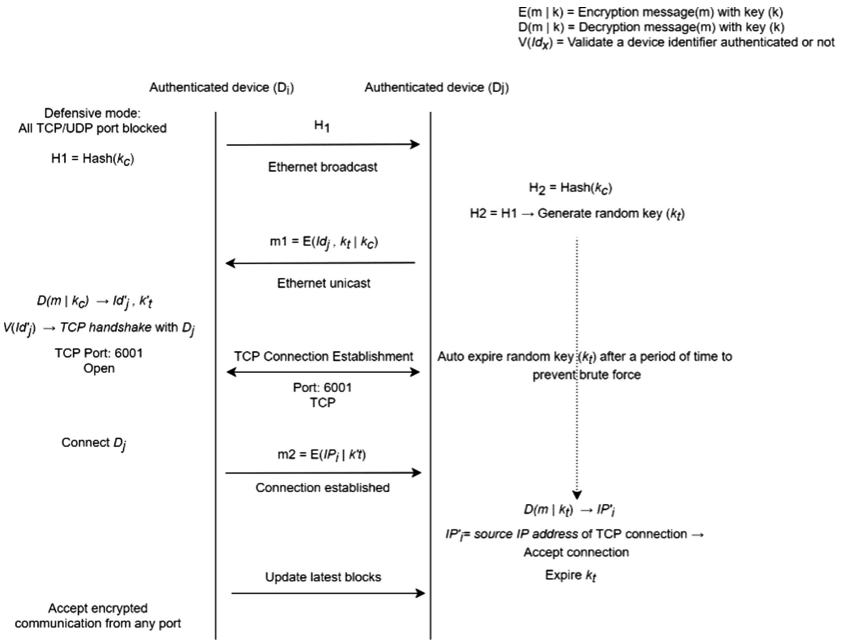


Fig. 4. Communication between authenticated network devices using ADCP

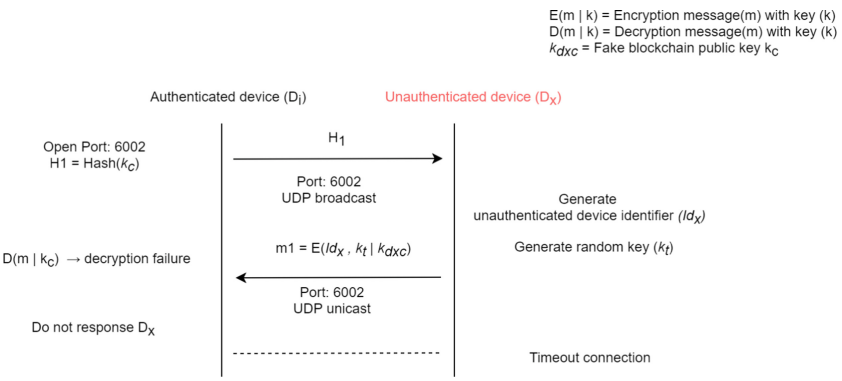


Fig. 5. Communication between authenticated and unauthenticated network

4 Implementation

A security software system has been built to validate our design. We used node.js [9] and python [10] to build the software. The software system can be installed in network devices to handle the blockchain system, ADCP communication, and perform protocol checking. There are two versions of software: one is for the Hardware authenticator (HA) and the other is for other network devices.

fridaya@uef.fi

Figure 6 shows the system diagram for HA. HA accepts authentication requests from unauthenticated device. An HTTP interface is provided to the network administrator to monitor those requests and authenticate them one by one. Once HA receives authentication confirmation from the network administrator, it signs the authentication transaction (AT) and generate new blocks. Then, the new block will be sent to the target device. Figure 7 shows the system diagram for network devices. The software firewall is the major difference from that showing in Fig. 6. The software firewall will examine all traffics to see if the protocol matched with the device identity or not.

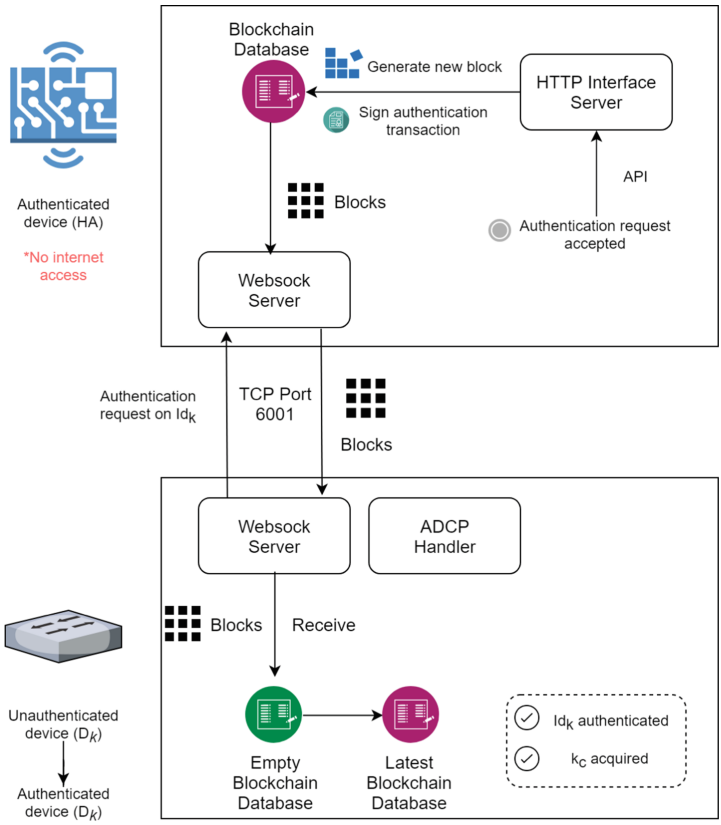


Fig. 6. System diagram of a hardware authenticator (HA)

fridaya@uef.fi

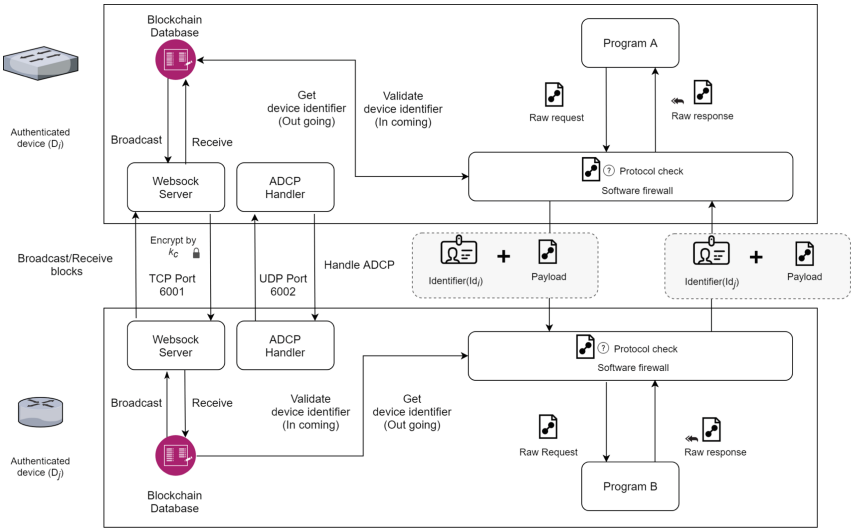


Fig. 7. System diagram of network infrastructure device

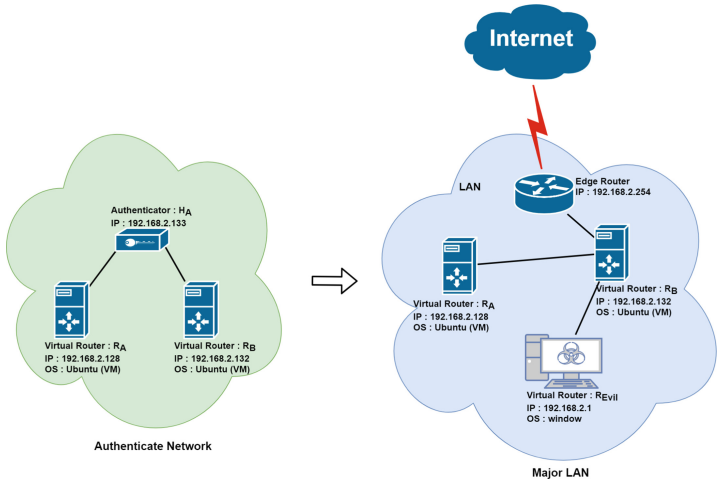


Fig. 8. VM connection diagram

5 Experiment and Results

Experiments were conducted to demonstrate our design. As shown in Fig. 8 shown, two Virtual Machines [11] were used to emulate the network infrastructure devices. Both are Virtual Routers (RA and RB) [12]. Another virtual machine was used to emulate the Hardware Authenticator (HA). First, RA and RB join the authentication network and being authenticated by the HA. Then, both were disconnected from the authentication

fridaya@uef.fi

network and join the major local area network (LAN). After the handshake of Authenticated Devices Configuration Protocol (ADCP), an overlay network infrastructure was formed within the Major LAN. An unauthenticated Virtual Machine (REvil) was added to the network to test the security enhancement bought by our solution.

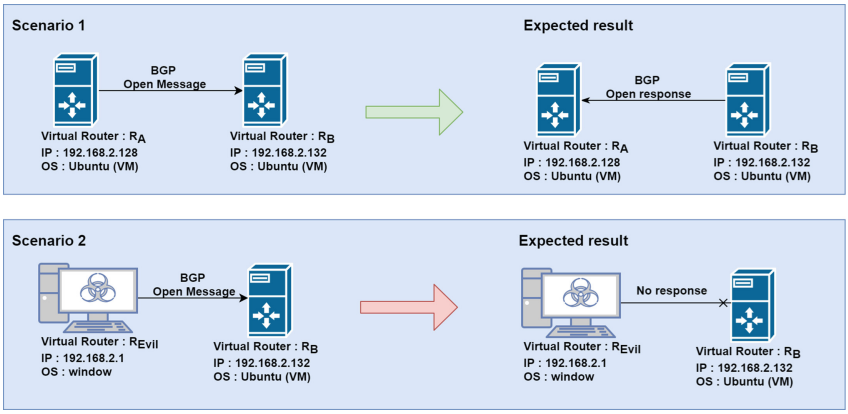


Fig. 9. Two scenarios and expected result for the experiments

Figure 9 shows two typical scenarios which were tested in the experiments. Border Gateway Protocol (BGP) [13] was chosen to demonstrate the security enhancement when adapting our solution. Although BGP does not provide authentication by itself, our software firewall is expected to discard any BGP communications from unauthenticated infrastructure devices. TCPdump [14] was installed on both RA and RB to capture traffics.

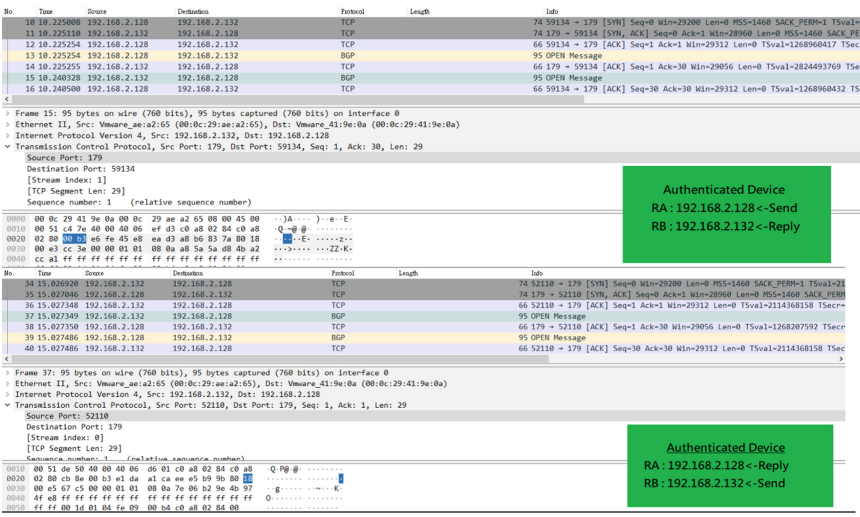


Fig. 10. Experimental result for scenario 1

fridaya@uef.fi

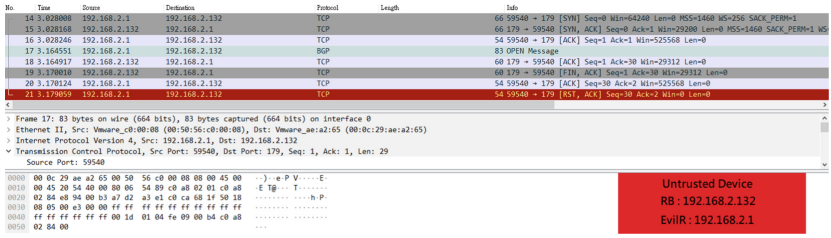


Fig. 11. Experimental result for scenario 2

For scenario 1, RA sends a BGP Open message to RB. Since both are authenticated devices, RB should accept the BGP message from RA and reply to RA. For scenario 2, REvil also sends an Open message to RB. Since the request comes from the unauthenticated device, RB should discard the message and terminate the connection.

Figure 10 shows the result of TCPdump in scenario 1. All captured traffic was transferred and opened by Wireshark [15]. The upper screenshot shows the traffic when RA sent the Open message to RB, and the lower screenshot shows the traffic when RB sent the Open message to RA. According to the screenshot, the BGP Open message communication only successes between RA and RB. The experimental results matched the expected result.

Figure 11 shows the result of TCPdump in scenario 2. According to the screenshot, a BGP Open message has been sent from REvil to RB. However, no data was returned from RB because the message is from an unauthenticated device (REvil). Afterwards, the TCP connection was closed. The experimental matched the expected result.

6 Conclusion

To prevent the security issue causing by authentication problem, an authentication scheme supported by blockchain technology is proposed. The detail system design and experimental results has shown the feasibility of the solution. Such design can be implemented in current networks without changing any standard protocol in use. Also, building an encrypted overlay network infrastructure is possible based on this solution.

References

1. Adams, C., Lloyd, S.: Understanding PKI: Concepts, Standards, and Deployment Considerations. Addison-Wesley Professional, Boston (2003)
2. Kumar, Y., Munjal, R., Sharma, H.: Comparison of symmetric and asymmetric cryptography with existing vulnerabilities and countermeasures. Int. J. Comput. Sci. Manag. Stud. **11**(03), 60–63 (2011)
3. Zhang, X., et al.: Malicious packet dropping: how it might impact the TCP performance and how we can detect it. In: Proceedings 2000 International Conference on Network Protocols. IEEE (2000)

fridaya@uef.fi

4. Padmanabhan, V.N., Simon, D.R.: Secure traceroute to detect faulty or malicious routing. ACM SIGCOMM Comput. Commun. Rev. **33**(1), 77–82 (2003)
5. Yeung, K.H., Leung, T.C.: Building secure network infrastructure for LANs. IPSI Trans. Adv. Res. **2**(2), 32–37 (2006)
6. Neuman, C., et al.: The Kerberos network authentication service (V5). No. RFC 4120 (2005)
7. Zyskind, G., Nathan, O.: Decentralizing privacy: using blockchain to protect personal data. In: 2015 IEEE Security and Privacy Workshops (SPW). IEEE (2015)
8. Lopez, J., Dahab, R.: An overview of elliptic curve cryptography. Tech. Rep. (2000)
9. Node.js: nodejs.org (2018). https://nodejs.org/
10. Python: Python Software Foundation (2018). https://www.python.org/
11. Ali, I., Meghanathan, N.: Virtual machines and networks - installation, performance, study, advantages and virtualization options. Int. J. Netw. Secur. Appl. **3**(1), 1–15 (2011). https://doi.org/10.5121/ijnsa.2011.3101
12. Anwar, K.: The Linux Router, 31 July 2002. https://www.linuxjournal.com/article/5826. Accessed 27 Oct 2018
13. Rekhter, Y., Li, T., Hares, S.: A Border Gateway Protocol 4 (BGP-4), January 2006. https://tools.ietf.org/html/rfc4271. Accessed 28 Oct 2018
14. TCPdump: Tcpdump/libpcap public repository (2018). http://www.tcpdump.org/
15. Wireshark: Wiresharkt • go deep (2018). https://www.wireshark.org

fridaya@uef.fi

Blockchain: Analysis of the New Technological Components as Opportunity to Solve the Trust Issues in Supply Chain Management

Adnan Imeri[1,2(✉)], Nazim Agoulmine[2], Christophe Feltus[1], and Djamel Khadraoui[1]

[1] Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg
adnan.imeri@list.lu
[2] Universitè of Èvry Val d'Essonne, Evry, France

Abstract. The issues of trust in the area of supply chain management are an immense concern among the stakeholders cooperating in the supply chain. For a sustainable process of logistics and transportation, an efficient information sharing is considered crucial. The current systems that serve as a base for the operations in supply chain have several drawbacks in terms of data security and trust among stakeholders, who share information as part of their cooperation. Information is shared in a paper-based or semi-digitalized way due to the lack of trust or risk of competitive disadvantages in the current systems. This paper aims to analyze the trust issues in supply chain management and propose new ways of improving trust by considering these issues at the design level. Technological components are analyzed for designing new models and to support the implementation of new architectures that improve trust in a systems which manage supply chain processes.

Keywords: Trust · Blockchain · Supply chain management · Design science research

1 Introduction

The dynamic growth of markets and trading has greatly increased the need for extending the cooperation of stakeholders (members) in the Supply Chain (SC) [1,2]. The globalization era turned the SC into a dynamic and distributed environment, with strong dependencies on communication for work synchronization [1,3,4]. In the ecosystem of SC stakeholders (shippers/suppliers, manufacturers/recipients, retailers/customers and logistics services [5,6]), a significant step for organizational work is the sharing of information. The process of communication or exchange of information enables better collaboration among stakeholders as members of the SC. This process is strongly dependent on the technological selection for managing the context of communication between stakeholders of the

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 474–493, 2019.
https://doi.org/10.1007/978-3-030-22868-2_36

fridaya@uef.fi

SC [3,7]. The performance of the SC is dependent on the common agreement of SC stakeholders, which is closely related to the timely sharing of information [3,4]. This process will face many difficulties if an SC stakeholder hesitates in sharing information, due to a lack of trust of the other parties in the SC. This would considerably damage the process and may cause many issues with the effectiveness of processes in the SC [4]. The globalization era and high competitiveness oblige the stakeholders of SC to outsource some of their sub-services by delegating them to third-party service providers, e.g., third-party logistics or IT firm [3]. From the different perspectives of stakeholders in the SC, sharing information between themselves may be a challenge, since this information may be outsourced to third parties [3,4].

Trust is an essential component for ensuring long-term cooperation between stakeholders in the SC and leads to competitive advantages [3,4,8]. Trust is the key mechanism for a long relationship and the commitment of the SC stakeholders, and it is consider as an essential component for preforming inter-organizational business processes [4,8]. Although the term "trust" is a classic case study from scholars and professionals, there is some divergence on a common definition of it. Trust is defined in different ways based on the needs and purposes of the authors. The most generic definition for the term trust is proposed by research from [3], and is most likely related to the context of our study: "trust means that a company is willing to take the risk, or expose itself about another company" [3]. The concern of the SC stakeholders in their daily SC operations is the issues of trust with different stakeholders. Sharing information with third parties exposes high risks, because of the possibility of disclosing information to other, unauthorized stakeholders [4]. The reliability of such systems decreases since in particular the system may be considered untrusted. This research aims to analyze the issue of trust in supply chain management (SCM) with the purpose of supporting the design of new models and identifying the technological component for improving the trust and reliability in the systems that supports SCM operations.

This research is synthesized as follows. In Sect. 1, we present the definition of the problem and the research questions raised. Section 2 shows the analysis of trust issues based on a use case. In Sect. 3, we present state of the art for promoting new technologies, i.e., Blockchain technology and related trust issues. Further, in Sect. 4 we analyze the new opportunities and perspectives for improving trust by design in SCM. Our conclusion and future trends are detailed in Sect. 5.

1.1 Research Method

As explained in the introduction, the objective of this paper is to analyze the trust issues in supply chain management and propose a new approach for solving them. Hence, the scope of the paper is not the design of a new artefact (model, method, software, tool, etc.) but rather an effort to understand the context related to trust in the supply chain and identify the appropriate technology

fridaya@uef.fi

to investigate in order to propose a solution to the trust problem (research questions elaboration). According to Wieringa [9] the design of a new artefact in the field of design science research needs to be achieved following an iterative regulative cycle composed of the following four steps: *(1) Problem investigation, (2) Solution design, (3) Design validation, (4) Solution implementation.* After the first iteration, *Problem investigation* turns into *Implementation evaluation.* Each of these four steps corresponds to practical problems or knowledge questions, i.e., (1) and (2) are knowledge questions (respectively: what is the problem? Is the design valid?) and (3) and (4) are practical problems (respectively: design the solution and implement the design). Acknowledging this design science approach, the paper proposes first to answer a knowledge question related to the problem investigation and second to set up the basis for the design of the solution, which is a practical problem. The main knowledge question is **"what are the trust issues related to the supply chain?"** The answer to this question aims to provide scientific information about the concrete trust problems arising when supply chains are deployed. The semantic of that question is enhanced in Sect. 1.3 by three sub-questions (Q1 to Q3). To answer this question, we first propose a use case of the supply chain in the field of transport of dangerous goods, and explain, regarding to that case, what the issues of trust are. Afterward, we performed a systematic review of the scientific literature in order to illustrate to what extent solutions are already available (Sect. 3), and we reviewed traditional trust in SCM technical solutions to analyze how existing architectures allow existing trust frameworks to be deployed (Sect. 3.1). Finally, we highlighted a new trend aiming at strengthening the management of trust based on blockchain. The answers to the knowledge questions about the detection of the trust issues in SCM has provided four trust requirements to be fulfilled by trust solutions. Based on the four requirements, the second part of the paper investigates how an upcoming solution could fulfill these requirements. Therefore, the second main research question (Q4), answered in Sect. 4, corresponds to a practical problem: **"what are the possible options for fulfilling the trust requirements?"**. With the advantage of the blockchain technology, this question can be answered thorough analysis of the way this new blockchain technology allows the fulfillment of trust requirements.

1.2 Problem Definition

One of the main research challenges in recent years for the scientific community and retail industry enterprises is supply chain management (SCM). The supply chain (SC) is a network composed of entities such as "Suppliers", "Manufacturers", "Retailers" and "Logistics", which have different activities in the SC [6,10,11], as presented in Fig. 1. The "Suppliers" are responsible for providing raw materials for the "Manufacturers". The "Manufacturers" use these materials to produce their products (goods). The goods are exposed to the markets through "Retailers". In the SC network, "Logistics Services" cover almost any interaction processes between stakeholders. Therefore, "Logistics Services" is the

fridaya@uef.fi

middle part of all these entities, covering the transportation and management of goods, until the customers receive the desired goods.

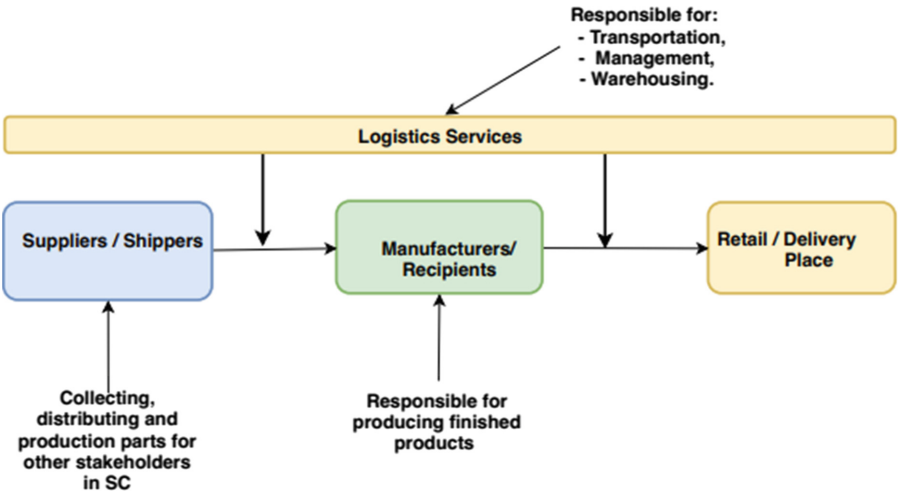


Fig. 1. The schema of the interactions of entities in SC [6]

The flow of goods, information and financial transactions needs to be appropriately managed to comply with business contracts, legal policy, and national and international standards [14]. For managing activities in the SC, the general term introduced is "Supply Chain Management", which covers any workflow management in the SC. The term supply chain management (SCM) is defined by research in [12] as "the systemic, strategic coordination of the traditional business functions and the tactics across these business functions within a particular company and across businesses within the supply chain, for the purposes of improving the long-term performance of the individual companies and the supply chain as a whole". The research from [13], defines SCM as "the integration of key business processes from end-user through original suppliers, that provides products, services, and information that add value for customers and other stakeholders". Due to the distributed nature of the SCM and the fact that the marketplace is growing daily, the logistics and transportation are subject to disturbances while close cooperation is required between stakeholders involved in the supply chain (SC) [6]. A large number of communications, the centralized storage of information, lack of process auditing, lack of long-life management of goods, traffic congestion during the transportation of goods, and other possible obstacles, are current SC challenges. Furthermore, an SCM concern considered is data security, trust among suppliers, and the asymmetry of information (on the trading markets). The number of participants in the SC is increasing, meaning that each partner relies on another, and trust among these partners is an essential factor for daily business and the optimization of processes in SCM [14] [1].

fridaya@uef.fi

Existing systems, such as ERP[1], as are usually used for SCM, are centralized services. The concern is on "trust", since the information is stored in a centralized database, and "owned" by the host of the service, e.g., third-party IT firm or stakeholder. This way of storing information, as shown in Fig. 2, allows the administrators of the database to modify, delete or insert other related information. This model and its technological components enable the sharing of information among stakeholders, and these types of architectures are usually implemented in the current systems, ERP systems [16,17]. Updating and maintaining the centralized database is a crucial problem since parties are dependent on each other in the SC, and the database requires a high level of work organization [15,18]. For example, for upgrading systems, some of the processes will remain on hold for the other parties in SC, which may cause delays and inconsistencies in SC processes, e.g., in the transport of goods.

The current challenges in the SCM come from the management of the workflows in the SC, which are related to the organization of processes by current systems [4]. Given that these systems are mainly centralized, several issues become evident:

- The lack of trust among stakeholders who cooperate in the SC [19].
- Sharing sensitive information, e.g., the substance (goods) for transportation, contractual business terms, capacity, source, destination, current warehousing, the timestamp of the movement of goods and final destination [4,20].
- Interoperability issues: The new stakeholders should adapt their systems to exchange information with other parties in the SC [4,19].
- Competitive disadvantages: The possibility to share the information with other business competitors [20,39].

The challenges increases with the number of participants in the SC, increasing the level of complexity in terms of communication, fulfilling the tasks in the demand-response environment, and the possible disclosure of information for unauthorized parties or exposing the risk for competitive business disadvantages. These may include a company's attempts to keep information about types of products, capacity, and destinations confidential. Furthermore, trust tends to decrease due to the many interactions and the shared ecosystem of information, which provides the possibility for third parties to access this information. Following, for the model presented in Fig. 2, the trust is expressed as follows:

ρ - number of supply chain stakeholders;

$\alpha = 1$ - maximal value of trust;

ε - trust value over system;

In a third party trusted (centralized) approach, the trust decrease, if the number of participants increases:

$$\varepsilon = \frac{\alpha}{\rho}$$

[1] Enterprise Resource Planning. The current software systems used in SCM, such as BI, APS, CM, SCE, FDM, etc., [37,38].

fridaya@uef.fi

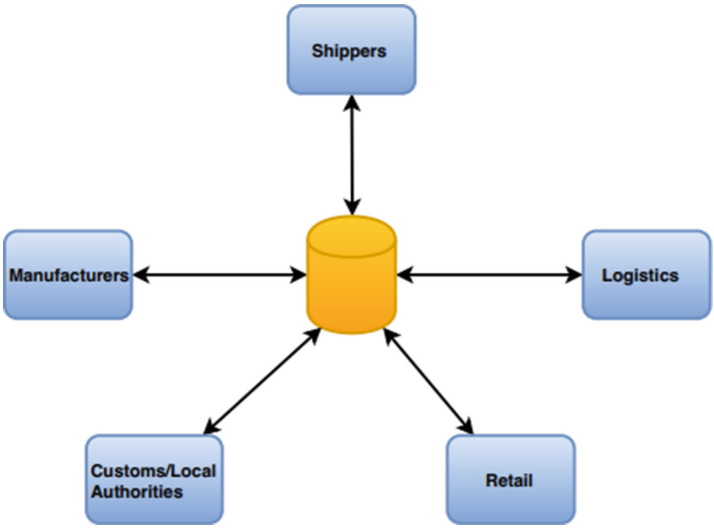


Fig. 2. The traditional (client-server) systems in SC (adapted from [4] and [1])

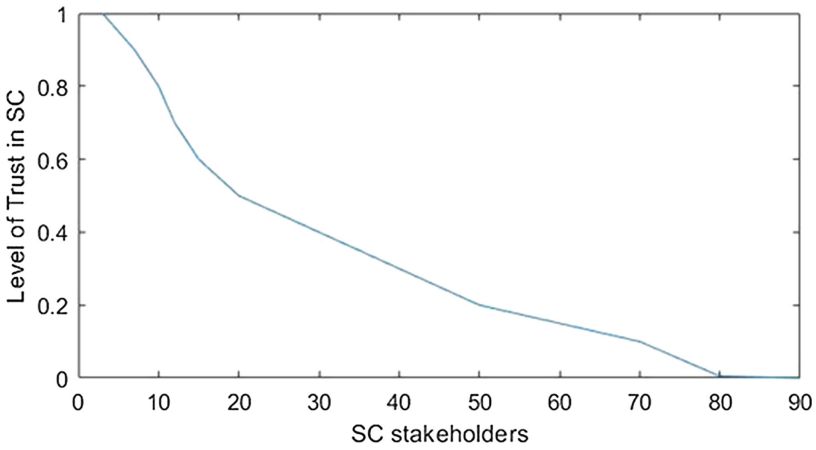


Fig. 3. The level of trust in centralized approaches concerning the number of SC stakeholders

The results from the formula above are showed in Fig. 3[2]. The level of trust (y-axis) in the report with the number of participants (x-axis) in SC.

[2] The results for Figs. 3, 5 and 6, respectively, are obtained after an implementation of given formulas by using Matlab. A github solution is provided in the following link: https://github.com/adnanimeri/MCodeSAIPapers.

fridaya@uef.fi

1.3 The Research Question

The SCM is a complex and dynamic environment to manage, and remains challenging in the different levels of organization of its processes. We extract some research questions from the challenges mentioned in the previous section and address them based on information discovered during the exploration of literature and potential real use cases. This resulted in the following research questions:

Q1: What are the issues in the current supply chain management?
Q1: What is the state of the art of the technology that can respond to the current issues?
Q3: What is the gap between requirements and solutions?
Q4: What are the possible options for fulfilling the trust requirements?

For answering these questions, we first propose a use case of the supply chain in the field of transport of dangerous goods, and explain, regarding that case we explore what the trust issues are. Afterward, we review the scientific literature to illustrate, what is the extent the current blockchain-based solutions regarding trust issues (Sect. 3.1). Finally, we highlighted a new trend aiming at strengthening the management of trust based on blockchain (Sect. 4).

2 Analyzing SCM Trust Based on Use Case

The dangerous goods are considered any material (gases, liquid, or solids) which shows risks potential for harming humans, animals or environment [40]. The activities of businesses and humans are requesting on transportation of dangerous goods [41]. Several accidents with dangerous goods occurred and the number is increasing yearly. The consequences from these accidents implies on risks for life preventive of humans, living organisms, property and environment.

Use Case: In the context of the transportation of dangerous goods (TDG), many stakeholders cooperate to fulfill the process of transport. This process is very delicate because of the risks for the environment and human life. The process reaches a certain level of complication due to the many standards and strict local and international policies. For the TDG from the point of origin to the point of destination, many processes must be followed, which generate sensitive information to be shared among the stakeholders. This information is crucial for a sustainable, correct, and risk-free for malicious parties to access them as a material, e.g., nuclear waste. The core part of this information shared between stakeholders are contractual issues, types of goods for transport, the timestamp of movement of goods, itinerary and journey described by these goods, warehousing, and the mode of transportation [1,20].

fridaya@uef.fi

2.1 The Issues of Trust Related to the Use Case

The cooperation of the stakeholders in the TDG process is more stringently organized than for the usual goods. They exchange information in a paper-based, semi digitalized way, e.g., email or phone calls and digitalized by using IoT devices and the centralized storage of information. In this context of organizing work for TDG, the main concern related to this information is whether the shared information can be trusted [1]. This comes from the fact that the information in a centralized database is editable. Since the single database stores this information, in the event of any failure, the whole system goes offline as it is the single point of access. Further, unauthorized access from third parties to this information remains a challenge for the entire ecosystem of the SC [20,21]. From a business perspective, a huge concern remains on the possible disclosure of information for competitors from the same domain of operation. This means that when it comes to sharing information, the TDG process becomes mostly paper-base and semi-digitalized [21]. The current systems do not offer a guarantee that the information shared between stakeholders is completely secured, and that unauthorized parties do not have any access to this information. This obviously decreases trust among stakeholders while the technological aspects do not offer any mechanisms to avoid these problems. Thus, stakeholders cooperate by relying on and trusting their working partners, instead of trusting the system. For overcoming such challenges and improving SCM, we researched new possibilities for designing new models and identifying the technological features, to reassure the stakeholders that the system is trusted. The technological aspects of such system are presented in Sect. 3, by portraying the main characteristics as advantageous for using it to develop the SCM systems.

3 Blockchain: The Rise of New Technologies

The new technology we have selected is the Distributed Ledger Technology or Blockchain. This technology is the current most disruptive technology in many cross-industries. The blockchain technology is a distributed-decentralized database, which enables digital information to be stored from different nodes in the network [22,23]. The blockchain network is composed of several nodes, which operate in a peer-to-peer mode of communication. These nodes rely on each other, instead of relying on a central authority [22,23]. For the nodes in the network, the consensus algorithms validate the set of transactions gathered in the data containers, known as the "block". The block is then added to the blockchain, and all nodes receive this "block" [22–26]. The blocks of data in the blockchain technology are linked together by following an ordering rule, where each new block contains the hash address of previous blocks, appending it in the chain of blocks or blockchain [22–26].

Blockchain technology has shown some potential technological features. The term "block" presents a data structure which is composed of the following components: *timestamp*, which indicates the time the block was created (appended into the blockchain). The previous hash is the hash value from the previous

fridaya@uef.fi

block, e.g. the block *n* contains the hash from the header of block *n-1*. Nonce is the value generated by the consensus algorithm, e.g., Proof of Work [22–26]. Transaction Root is the root of all transactions received from the nodes in the network for a determined timestamp. These transactions are organized in the tree by using the Merkle Tree. The main properties of the blockchain technology are:

- *Decentralized*: Blockchain technology relies on a peer-to-peer mode of communication. It does not have any central authority for storing and retrieving data [1 - 5].
- *Consensus Algorithm*: In the blockchain technology, the consensus refers to the agreement of nodes in shared content. Different consensus algorithms are used, such as Proof of Work (PoW), Proof of Stake (PoS), Practical Byzantine fault tolerance (PBFT), etc. [22–26].
- *Data Security*: The data stored in the blockchain are cryptographically checked. The blockchain technology uses digital signatures (public key cryptography) for signing and verifying transactions [22–26].
- *Data Immutability*: The data recorded on blockchain are cryptographically checked and distributed over all nodes in the network. For changing a transaction in the blockchain, the user should change all transactions at the same time in all nodes in the network. Furthermore, this is impossible when the consensus algorithm compares the hash root of the transaction and denies these changes [22–26]. Therefore, the transaction recorded in blockchain cannot be altered or deleted [22–26].
- *Auditability*: The timestamp of the validation transaction enables any user to trace the previous transaction executed by a specific user. This is possible by having access in the blockchain to any node in the network.
- *Smart Contracts*: A computer code deployed on the blockchain for performing specific tasks after some predefined conditions are fulfilled. A suitable mechanism for data processing and the management of processes in the supply chain is business contractual implementation.
- *Process automation*: The usability of the *smart contracts* and their ability for self-execution when a specific condition is fulfilled, present one of the main features for the automation of processes in SCM.
- *Interoperability:* The core component of blockchain technology enables many parties to access blockchain under pre-defined conditions.
- *Low-cost maintenance*: Blockchain technology does not use any central authority for the exchanging of messages and validation of transactions. This enables low-cost operations when using blockchain since there is no need to develop server infrastructures for the validation of transactions. This is in contrast to traditional systems, which use central servers for messages exchange and validation, and which usually have high database maintenance costs (upgrade, backup) [22–26].
- *Sustainability*: If several nodes fail or are disconnected, the blockchain is still available and works properly on the remaining nodes. When the "offline" nodes come back into "online" mode, they receive the latest state of the ledger [22–26].

fridaya@uef.fi

- *Public blockchain:* In the public blockchain, access is without permission, and it is considered fully decentralized. Any end user can join the network, execute the transaction or explore the block of the transaction conducted by other end users [22,24].
- *Consortium blockchain:* This type of blockchain is mainly used by organizations, and is considered partially decentralized. For example, if there are 15 organizations that host this type of blockchain, a consensus may be achieved if 10 organizations sign the transaction. Further, access rights are granted in specific cases.
- *Private blockchain or permissioned blockchain:* In private mode, different levels of access and read and write permissions are presented. To access the private blockchain network, permission must be granted [22,24].
- *Data management:* On-Chain vs Off-Chain. In the context of data management, the common practice in blockchain is to sort the raw data off-chain, and store the meta-data related to transactions, e.g., hashes of the transactions, on-chain [23,25].

Since its invention, blockchain technology has influenced many industries. The technical specifications of the blockchain are the considered the backbone for solving a specific problem in several domains of application including Supply Chain Management.

3.1 Blockchain and Trust in SCM

Blockchain technology and its possible implementations is a subject of study in the different level of SCM and other systems where trust is needed as a component. We present some state-of-the-art approaches that use blockchain as a technology component for developing their systems, e.g., for health care, trading, SCM, etc. and show specific research related to blockchain and trusted supply chain management. These studies are mainly explorative for the potential of blockchain in the supply chain domain, although some approach a specific use case. It is worth noting that this research paper is an extensive study of our previous research showed in [42].

The research presented in [27] proposes a new way of managing health information (patient data) with the creation of a new platform based on blockchain technology. This platform will bring together all medical organizations, and will manage patient data, minimizing the risk of exposing health information. The potential of blockchain properties as decentralization, distribution, security, interoperability, non-reputation, are evaluated as a core solution for developing such platforms, and for achieving trusted and sustainable healthcare platform. As opposed to the current solutions, while the data are centralized, and disclosed to patients and other health organization, this platform will propose a decentralized and secure platform for patients and health organizations to manage health information [46].

A real case for the trading market has been developed as a proof of concept using blockchain. The research by [28], presents the implementation of concepts

fridaya@uef.fi

in the market of trading old cars. The primary intention behind this research was to provide a single trust-free platform based on the blockchain that manages information related to the history of cars on the market. This platform helps buyers and sellers avoid the asymmetry of information and governmental organization to manage the registration of cars [46].

The research by [29] shows how the trust-based solution can be replaced by a trust-free solution based on the blockchain. The case study shown in this research is the concept of a trust-free coffee shop payment solution using digitalized punched cards. The advantages of using blockchain for designing and developing a trust-free solution also highlighted the current obstacles, e.g., scalability for using such an application in the real world [46].

This research [30] explores the potential of the blockchain as a validation tool. The trust, validation, distribution mechanism, and compliance issues and the blockchain impacts on them are explored. The standardization of blockchain technology is considered a crucial issue for achieving the massive usability of the blockchains, and the main points, which need to be standardized, are storage algorithms, signature algorithms and web-based access protocols. Moreover, this research shows a comparison of blockchains and their characteristics related to a given pattern. Another explorative study, which shows the potential of the blockchain in supply chain management, is presented by [31]. The main benefits of blockchain technology, such as transparency, trust, security and access to new markets are highlighted in this research. Several research projects were reviewed in this study, such as "Maersk and IBM" and "Walmart and IBM". The main idea behind these projects is to ensure the security of information during transport, which increases trust among suppliers. This research also shows some difficulties in adapting and using the blockchain technology [46].

The research from [32] explores the potential of blockchain for more transparency in supply chain management and logistics. The potential improvement in transparency between suppliers and consumers was investigated. To evaluate blockchain as an innovation tool, a well-known method, comprising the following five main points was used by [32]: relative advantages, compatibility, complexity, trialability, observability. The conclusion confirms that blockchain has the potential to generate transparency and ensure the fulfillment of contractual terms between suppliers in the supply chain [46].

The research from [33] study the challenges for data integration for the business-to-business case in the digital supply chain (DSC). The way the current supply chain system works, for now, stakeholders use third parties to share and execute specific work. This research investigates the requirements and functionalities of supply chain integration. It intends to present a new approach based on the blockchain, respond to questions on the acceleration of DSC integration and show how this technology supports that integration [1].

The study from [34] provides a framework for businesses to evaluate the possible applicability of blockchain in supply chain management to improve traceability. The aim of this framework is to discover the necessary inputs and evaluation tools needed to apply blockchain technology into supply chain management.

fridaya@uef.fi

This framework is based on the theory presented for supply chain traceability and blockchain technology, and comprises three main steps: (1) identification of drivers for traceability (or areas for improvement), (2) creating the main principles of blockchain and (3) evaluating sustainability applicability and technical limitation. The process methods could follow these steps in a disordered manner, and as a result, it outputs the effectiveness from the implementation of the blockchain technology in supply chain traceability (or any area targeted for evaluation) [1,46].

The research showed in [35] describes the whole traceability process, and gives a step-by-step example. Blockchain technology has been adapted for smartphones, for following production from the manufacturer, then through warehouses, to the store (retailer). The research highlights that, for the possible implementation of this technology and its continuous usability, cooperation between suppliers is required, then for having a footprint the usability of the smartphone is mandatory, and finally an integrated system should exist between the suppliers involved in the process [46].

The results conducted by these research papers show the potential of blockchain in SCM. The exploration of this research permits the evaluation of current conceptual models and implementations. We intend to provide an analysis of possibilities to provide trust in SCM, by considering trust from the design level expressed in a model, and using technological components to define a trusted SCM model.

4 Analysis of New Technological Opportunities and Perspectives

This analysis aims to study the potential technological aspects for improvement of trust mechanism in SCM. For performing such an analysis, we use the defined use case, as presented in Sect. 2, and measure the possible improvement of trust by using technological components of blockchain. Initially, we define the basic requirements from the scenario of TDG:

- R1: The shared information between stakeholders in the TDG process is required to be confidential and secure (immutable and with non-repudiation properties).
- R2: The communication and monitoring activities are enabled from a decentralized platform for the parties involved in the TDG process.
- R3: The platform and data loading needs to be unique, i.e., standardized, and interoperability issues need to be resolved.
- R4: The process should be auditable, from the beginning of contractual business terms to the delivery of goods.

This set of requirements concerns mainly the issues of trust, with data security, which is a common subject of sharing among stakeholders, and the data considered as the process determination for the TDG [1].

fridaya@uef.fi

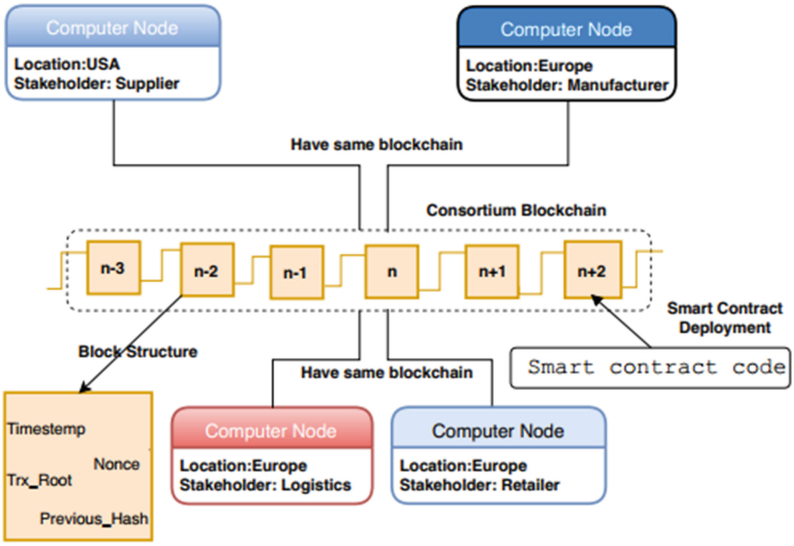


Fig. 4. A blockchain-based conceptual solution for TDG. Inspired from [1]

For a trusted process of TDG, the data sharing and data security is responsibility of stakeholders involved in this process. For fulfilling these requirements, we investigate and analyze the improvement of trust in SCM by using recent blockchain technology. The application of this technology in SCM has several benefits regarding trust enhancement, such as peer-to-peer data exchange, data security, data integrity, non-repudiation properties, an interoperable solution, and auditing. The blockchain-based applications are considered trusted due to the two main components they contain: public key cryptography and consensus protocol. The information exchange and process management, presented in sketch from Fig. 4, are performed using a peer-to-peer mode of communication. This avoids the need to share the information with stakeholders that "own" this information in their databases, or with third parties such as IT firms. This method of communication, based on the principles of the blockchain, continuously enables immutable record-keeping in a distributed-decentralized shared database [1]. That gives exceptional features for a transparent and auditable process in TDG between the stakeholders involved. For example, for the transportation of goods from the origin point to the destination point, for the stakeholders using the blockchain-based solution, subsequent events are recorded in the shared ledger for the whole TDG process [1].

The workflow management in the SC is enhanced using this technology due to its secure data management characteristics. Furthermore, this advances the collaboration of stakeholders and improves confidence in TDG process. The set of requirements extracted from our use case are examined in Table 1, and the possible design solutions are addressed to them based on the blockchain technology characteristics. The purpose of the design decision (solution) is to resolve

fridaya@uef.fi

Table 1. The design solution for use case requirements based on blockchain technology

Requirement	Blockchain solution designs
R1	The information stored in blockchain is cryptographically checked. The cryptographic tools guarantee the integrity of the data. The transactions, which are stored in the blockchain, are considered immutable with the non-repudiation properties. For data exchange, the blockchain technology mechanism uses the combination of public key cryptography, and the nodes in the blockchain validate these transactions. The communication provided by blockchain, continuously enable record-keeping in a distributed-decentralized shared database [1,3]. For privacy reasons, the consortium or private blockchain allows better management of the access rights and read-write permissions [3]
R2	The blockchain-based solution provides a shared database for stakeholders. For any SC event, the data inserted in the blockchain are distributed between all participants in the network, as shown in Fig. 4. These blockchain features and the combination of smart contracts will facilitate the monitoring of processes in real time. In such a use case, a monitoring subject is identified as the digital movement of goods, administrative documents, and financial information. For monitoring reasons, a distributed web interface is enabled by blockchain technology[a] [1]
R3	For standardized communication, and a collaborative data set, a set of predefined standards for data sharing, such as the "EPICS Standard" for data sharing from GS1 or any other internal standard agreed between stakeholders can be applied for communication issues [19,36]. The current blockchains can process any data, enabling an interoperable platform for the TDG process since the equal access and communication methods are granted by blockchain-based solutions [19,24,25]
R4	The immutable record-keeping properties of the blockchain are the backbone for ensuring an auditable process in SCM. The "blocks" of the transaction are exportable into the blockchain technology, and for any process, an ID process will be introduced, and all related information will be stored for the documentation of the process. Furthermore, the continues recording of immutable transactions by all nodes involved in the system, ensures a sustainable maintenance of entire system [1,19]

[a]PFS is the Distributed Web https://ipfs.io/.

fridaya@uef.fi

trust issues and enhance the collocation between stakeholders. Beyond the use case for TDG, these solution designs could be used for any other use case in SCM.

Considering these solution designs based on the blockchain, we consider that this architecture enhances trust in the TDG and SCM in general. The scientific community and business enterprises are interested in using several characteristics of the blockchain, for approaching issues in SCM. Distribution, decentralization and data security show the higher likelihood of the selection of blockchain technology for designing and developing new SCM systems.

4.1 Measurements and Predictions for Blockchain-Based Solutions for SCM

The plot in Fig. 5 shows a comparison of the centralized architectures in the report, with the number of stakeholders in the SC, versus the decentralized architecture, as shown in the following formula:

ρ - number of supply chain stakeholders
β - trust value over each stakeholder (decentralized trust value)
ε - trust value over system

Based on the distributed nature of blockchain technology, the reliability of such a system increases as the number of stakeholders increases:

$$\varepsilon = \sum_{i=1}^{\rho} \beta_i$$

A blockchain-based solution is considered sustainable because it does not have a single point of failure. The measurements from Fig. 6 are also a result of the blockchain technology including the above-mentioned technological features, as well as its potential for SCM. The confidential level for the suggested models based on blockchain technology shows significant results, as presented in Fig. 6. The measurement intends to show that the blockchain-based models are more likely to be trusted for SCM. The measurements for the confidentiality level show that 90% of the R-squared RMSI = 0.115 [43,44]. The standard deviation for the linear distribution of stakeholders is *"31.398"*. This indicates that the potential real cases have approximately a 90% possibility of being solved by blockchain-based technology for the SCM. Based on the dataset, our model is behaving linearly, but for prospect higher problem, i.e., with higher dimension data sets, we might apply a nonlinear prediction model, e.g., polynomial regression [45].

4.2 Solving Trust and Enforcement Reliability in the Context of Our Use Case

In the context of our use case for TDG, the blockchain-based solution will enhance trust among the stakeholders involved in this process. Sharing information in an ecosystem in a secure way, allowing stakeholders to rely on the system supporting

fridaya@uef.fi

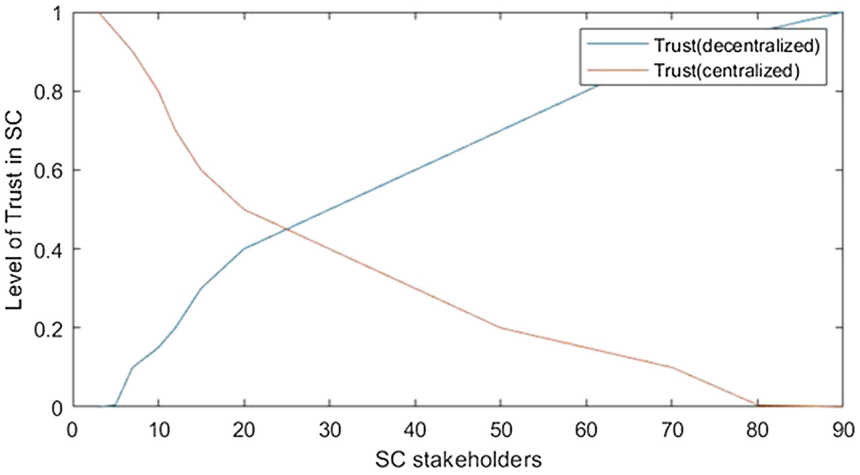


Fig. 5. The comparison of trust and reliability components in decentralized versus centralized mode

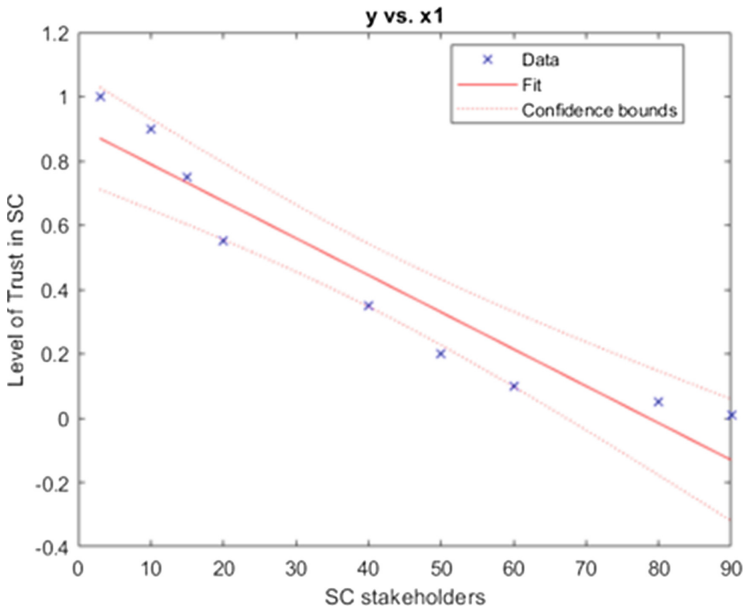


Fig. 6. Confidence level measurements for blockchain-based models

their operation, aims to enhance the TDG process. The current form of cooperation by TDG stakeholders is mainly based on paperwork or a semi-digital solution such as a fax. Stakeholders prefer exchanging information in this way to avoid using digital devices to share information. Because of possible data disclosure to competitors in the same domain of business, while the current technology does

fridaya@uef.fi

not guarantee data security and potential access from third parties entirely, using unauthorized channels would make this a possibility [1].

Among several blockchain solutions proposed already, the perspective of our solution proposes a *consortium blockchain*, which allows better management of stakeholders, consensus between them, and improve security and privacy aspects. Considering the distributed nature of the blockchain and the technological features, especially in data security, provides a core mechanism for the improvement of trust in TDG. Stakeholders who need to transport particular goods share the necessary information directly and securely with another party in the business contract, thus avoiding sharing this specific data with third parties, e.g., either on the stakeholder's side or IT firms for data management. Furthermore, a blockchain-based system allows stakeholders to work together in a standardized, interoperable framework and does not reveal any confidential information that might be harmful to their business activities. In such systems, data are not "owned" by any stakeholder, and the possibility of data manipulation tends to impossible, due to blockchain technological characteristics. Since the data are considered secured and immutable, the stakeholders rely on the system instead of trusting another party in business. The information shared among stakeholders remains accessible and auditable only by them. This increases trust and improves collaboration in the process of TDG by the participants involved in this process [1]. The design principles of blockchain (Table 1) and confidence level measurements support this conclusion regarding trust enforcement between stakeholders in the process of TDG.

5 Conclusion and Future Trends

We have analyzed current SCM trends, and discussed workflow management challenges and trust issues as one of the most highlighted problems in the SC. Therefore, we propose this research as an evaluation of possibilities for solving or improving the issues of trust in SCM. These analyses and measurements establish a stable base for treating the trust issues in SCM by considering the blockchain as a technological tool for resolving them. A fundamental way of designing new models should take into account the trust mechanism at a design level. The technological features considered should support new models, i.e., the introduction of the blockchain technological feature, to increase trust and provide a better operations system for such processes. A model and prototype of the solution will be presented in our future work. We will use the current use case, and in cooperation with business partners, we will implement our model in a real case. The results and related works will be published as an extension of this paper. This paper treats one of the most sensitive issues in SCM, such as trust. With several levels of trust and different interpretations based on the context of the trust-related challenges, still, this challenge remains debatable. We approached the trust related to data security in the context of data exchange among stakeholders in SCM. The analysis extracted from this study establishes components for resolving trust issues in SCM, by placing blockchain technology in pole position for addressing these issues.

fridaya@uef.fi

References

1. Imeri, A., Khadraoui, D.: The security and traceability of shared information in the process of transportation of dangerous goods. In: 2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS), pp. 1–5. IEEE (2018)
2. World Payments Report. https://www.worldpaymentsreport.com/. Accessed 11 Jan 2019
3. Sahay, B.S.: Understanding trust in supply chain relationships. Ind. Manag. Data Syst. **103**(8), 553–563 (2003)
4. Márquez, A.C.: Dynamic Modelling for Supply Chain Management: Dealing with Front-end Back-end and Integration Issues. Springer, Heidelberg (2010)
5. Mehmann, J., Teuteberg, F.: Process reengineering by using the 4PL approach: a case study on transportation processing in the agricultural bulk logistics sector. Bus. Process Manag. J. **22**(4), 879–902 (2016)
6. Witthaut, M., Deeken, H., Sprenger, P., Gadzhanov, P., David, M.: Smart objects and smart finance for supply chain management. Logist. J.: referierte Veröffentlichungen **2017** (2017). http://publica.fraunhofer.de/dokumente/N-470262.html
7. Chandra, C., Grabis, J.: Supply Chain Configuration: Concepts, Solutions, and Applications. Springer, Heidelberg (2007)
8. Weber, I., Xu, X., Riveret, R., Governatori, G., Ponomarev, A., Mendling, J.: Untrusted business process monitoring and execution using blockchain. In: International Conference on Business Process Management, pp. 329–347. Springer, Cham (2016)
9. Wieringa, R.: Design science as nested problem solving. In: Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology, p. 8. ACM (2009)
10. Lin, F.R., Sung, Y.W., Lo, Y.P.: Effects of trust mechanisms on supply-chain performance: a multi-agent simulation study. Int. J. Electron. Commer. **9**(4), 9–112 (2005)
11. Werner, H.: Supply Chain Management. Gabler, Wiesbaden (2000)
12. Mentzer, J.T., DeWitt, W., Keebler, J.S., Min, S., Nix, N.W., Smith, C.D., Zacharia, Z.G.: Defining supply chain management. J. Bus. Logist. **22**(2), 1–25 (2001)
13. Lambert, D.M.: Supply Chain Management: Processes, Partnerships Performance. Supply Chain Management Institute, Sarasota (2008)
14. Daniel, D.: What is supply chain management (SCM)? https://searcherp.techtarget.com/definition/supply-chain-management-SCM. Accessed 11 Jan 2019
15. Sreenivasaiah, P.K., Kim, D.H.: Current trends and new challenges of databases and web applications for systems driven biological research. Front. Physiol. **1**, 147 (2010)
16. Verwijmeren, M.: Software component architecture in supply chain management. Comput. Ind. **53**(2), 165–178 (2004)
17. Somers, T.M., Nelson, K.X.: The impact of critical success factors across the stages of enterprise resource planning implementations. In: Proceedings of the 34th Annual Hawaii International Conference on System Sciences, p. 10. IEEE (2001)
18. Coy, S.P.: Security implications of the choice of distributed database management system model: relational vs. object-oriented. In: NISSC, vol. 96, pp. 428–437 (2008)

fridaya@uef.fi

19. Staples, M., Chen, S., Falamaki, S., Ponomarev, A., Rimba, P., Tran, A.B., Weber, I., Xu, X., Zhu, J.: Risks and opportunities for systems using blockchain and smart contracts. Data61 (2017)
20. Imeri, A., Khadraoui, A., Khadraoui, D.: A conceptual and technical approach for transportation of dangerous goods in compliance with regulatory framework. J. Softw. **12**(9), 708–722 (2017)
21. Portofantwerp.com.: Antwerp start-up T-Mining develops Blockchain solution for safe, efficient container release. https://www.portofantwerp.com/en/news/antwerp-start-t-mining-develops-blockchain-solution-safe-efficient-container-release. Accessed 11 Jan 2019
22. Zheng, Z., Xie, S., Dai, H.N., Wang, H.: Blockchain challenges and opportunities: a survey. Work Pap. (2016)
23. Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System (2008). https://scholar.google.com/scholar?as_sdt=0%2C5&hl=en&inst=15483811489544987344&oq=Peer-to-&q=Bitcoin%3A%20A%20peer-to-peer%20electronic%20cash%20system
24. Xu, X., Weber, I., Staples, M., Zhu, L., Bosch, J., Bass, L., Pautasso, C., Rimba, P.: A taxonomy of blockchain-based systems for architecture design. In: IEEE International Conference on Software Architecture (ICSA), pp. 243-252 (2017)
25. Xu, X., Pautasso, C., Zhu, L., Gramoli, V., Ponomarev, A., Tran, A.B., Chen, S.: The blockchain as a software connector. In: 13th Working IEEE/IFIP Conference on Software Architecture (WICSA), pp. 182-191 (2016)
26. Walport, M.: Distributed ledger technology: beyond blockchain. Technical report, UK Government Office for Science, p. 19 (2016)
27. Nichol, P.B., Brandt, J.: Co-creation of trust for healthcare: the cryptocitizen framework for interoperability with blockchain (2016). https://doi.org/10.13140/RG.2.1.1545.4963
28. Cholewa, J.B., Shanmugam, A.P.: Trading real-world assets on blockchain-an application of trust-free transaction systems in the market for lemons. Bus. Inform. Syst. Eng. **59**(6), 425–440 (2017)
29. Beck, R., Czepluch, J.S., Lollike, N., Malone, S.: Blockchain-the gateway to trust-free cryptographic transactions. In: ECIS (2016)
30. Anjum, A., Sporny, M., Sill, A.: Blockchain standards for compliance and trust. IEEE Cloud Comput. **4**(4), 84–90 (2017)
31. Biggs, J., Hinish, S.R., Natale, M.A., Patronick, M.: Blockchain: revolutionizing the global supply chain by building trust and transparency (2018)
32. Badzar, A.: Blockchain for securing sustainable transport contracts and supply chain transparency-an explorative study of blockchain technology in logistics (2016)
33. Korpela, K., Hallikas, J., Dahlberg, T.: Digital supply chain transformation toward blockchain integration. In: Proceedings of the 50th Hawaii International Conference on System Sciences (2017)
34. Petersen, O., Jansson, F.: Blockchain technology in supply chain traceability systems (2017)
35. Jeppsson, A., Olsson, O.: Blockchains as a solution for traceability and transparency (2017)
36. GS1.org. EPC Information Services (EPCIS). https://www.gs1.org/sites/default/files/docs/epc/EPCIS-Standard-1.2-r-2016-09-29.pdf. Accessed 11 Jan 2019
37. Jennings, N.R., Wooldridge, M.: Applications of intelligent agents. In: Agent Technology, pp. 3–28. Springer, Heidelberg (1998)
38. Mofarrahi, G., Mohammadi, A.: The state of the art information sharing technologies for supply chain management. J. Theor. Appl. Inform. Technol. **59**(1) (2014)

fridaya@uef.fi

39. European Union Note. http://ec.europa.eu/competition/international/multilateral/2010_information_exchange.pdf. Accessed 11 Jan 2019
40. The European Agreement concerning the International Carriage of Dangerous Goods by Road (ADR) (2018). Unece.Org. Accessed 27 Sept 2018. https://www.unece.org/fileadmin/DAM/trans/danger/publi/adr/adr2017/ADR2017e_web.pdf. Accessed 11 Jan 2019
41. Carriage Of Dangerous Goods Manual - Classification (2018). Hse.Gov.Uk. http://www.hse.gov.uk/cdg/manual/classification.htm. Accessed 11 Jan 2019
42. Imeri, A., Feltus, C., Khadraoui, D., Agoulmine, N., Nicolas, D.: Solving the trust issues in the process of transportation of dangerous goods by using blockchain technology. In: Proceedings of the 11th International Conference on Security of Information and Networks, p. 25. ACM, September 2018
43. Coefficient Standard Errors and Confidence Intervals. https://nl.mathworks.com/help/stats/coefficient-standard-errors-and-confidence-intervals.html. Accessed 11 Jan 2019
44. Root Mean Square Error of Approximation. https://www.rasch.org/rmt/rmt254d.htm. Accessed 11 Jan 2019
45. Polynomial Regression. https://onlinecourses.science.psu.edu/stat501/node/325/. Accessed 11 Jan 2019
46. Imeri, A., Khadraoui, D., Agoulmine, N.: Blockchain for solving the security and trust issues in supply chain management and logistics: a survey. In: Engineering Digital Transformation. Springer (2018)

fridaya@uef.fi

Password Manager Combining Hashing Functions and Ternary PUFs

Bertrand Cambou(✉)

Northern Arizona University, Flagstaff, USA
Bertrand.cambou@nau.edu

Abstract. Hashing functions protect passwords against various hacking techniques because message digests can replace the passwords when stored in the network for future authentication. However, the message digests remain exposed to password guessing attacks, most hashing functions are known, and public. The objective of the protocols presented in this paper is to offer additional lines of defense using physical unclonable functions to convert the message digests into challenge-response pairs. The use of ternary physical unclonable functions reduces false rejection rates, and lowers the latencies during the processing of the authentications. Without having access to the PUFs, the look up tables storing challenge-response pairs are more difficult to attack than those storing message digests: they are unclonable, contain high levels of randomness, and quasi unique. The modeling efforts, and algorithms developed in this paper to validate the schemes, use commercially available components, and SRAM based ternary PUFs.

Keywords: Password management · Physical unclonable function · Hash functions · Ternary states

1 Introduction

The essence of access control is to match reference patterns such as passwords, secret keys, and biometric prints, against the same reference patterns of the users, which are stored in look up tables of the secure servers managing the networks [1–3]. One of the known exposures in the cyber space, is the potential loss to malicious entities of the look up table storing UserID-password pairs. Hackers are often able to damage institutions, and individual users by getting un-authorized access to databases of millions of passwords at once. Examples include attacks on internet providers, US government agencies, the IRS, health institutions, political parties, banks, and much more [4–7]. These types of cyber-attack, creates huge financial damages, and can be the result of insiders that are silent for years. A false sense of security can exposes millions of users overnight.

Ways to protect passwords and look up tables, from both external and internal attackers, have been proposed [8–14]. Several layers of protections for the access control are proposed in this paper; the first layer is the replacement of the userID-password pair by message digests of hashing functions [15–17]; the second layer is the replacement of the message digests by challenge response pairs (CRP) of physical

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 494–513, 2019.
https://doi.org/10.1007/978-3-030-22868-2_37

fridaya@uef.fi

unclonable functions (PUFs) [18–26]. In both cases, the exposure to stolen databases is lower, because the content of the tables is not directly exploitable to gain access to the original UserID-password pairs. The protection offered by PUFs is much higher, and mitigates attacks based on the knowledge of look up tables using message digests, and hashing functions [27, 28]. However, the need to use an error correcting code can be extremely complex [29–35]. Therefore, novel methods based on ternary PUFs, that are error correction free, are proposed. The architectures developed in this paper, reuses known technology modules, such as memory based PUFs, true random number generators, functions, and public key infrastructure. Several prototypes, containing both hardware and software components, were designed to evaluate, and validate the proposed schemes.

2 Background Information

2.1 Hash Function for Password Managers

Hash functions [36–38] are one-way cryptographic methods to generate message digests of fixed lengths, from incoming data streams of variable lengths. Hash functions have the following properties:

- A single bit change in the incoming data stream results in a totally different message digest.
- The incoming data streams cannot be extracted from the message digests.
- The hash functions are collusion resistant: the probability that two different streams result in the same message digest is extremely low.

Hash functions are pivotal in multiple cryptographic applications such as multi-factor authentications, and blockchains [39–44]. Hash functions replace, the look up tables containing User ID-password pairs [45], as shown in Fig. 1.

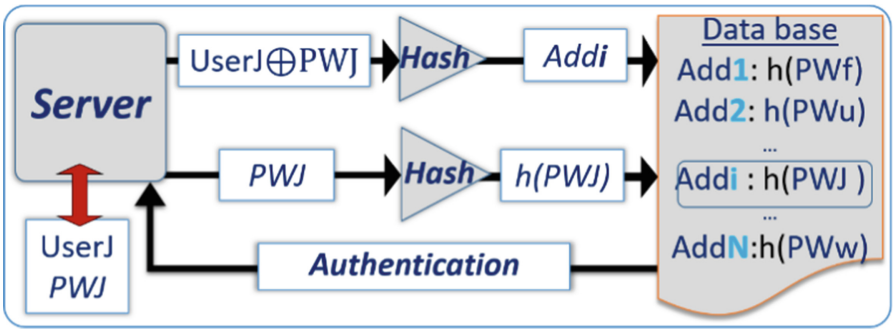


Fig. 1. Block diagram describing the data flow. Passwords and UserIDs are converted into message digests then into addresses (h(ID⊕pass)). On the right, the look up tables store the message digests.

fridaya@uef.fi

The hashing of the password ***PWj*** results in the first message digest ***h(PWj)***; the user ID ***Userj*** and ***PWj*** are XORed; the hashing of ***Userj*** ⊕ ***PWj*** results in the second message digest ***h(Userj*** ⊕ ***PWj)***. A portion of the second message digest is then used to generate the coordinate **XY** of the look up table storing the first message digest. The XORing of the user ID, and the password, enhances the protection, and this because the hashing of the User ID alone is weak, considering that most hashing functions are known and public.

To explain this method we developed a small example shown in Tables 1 and 2. Here we are using SHA-1 for its simplicity; however, for the prototypes proposed in the remaining sections of this paper, we used SH-3, which is safer, even in cases of attacks from quantum computers. In this example, user#1 has a user ID "a6c26", a password "12ae5", and a first message digest "a639d" generated with SHA-1; only the first five characters of the message digest are kept. The XORing the first three hexadecimal characters of the user ID and the password have a value of "b86". The hashing of "b86" with SHA-1 generates a message digest that has "3e" as first two hexadecimal characters. We then place "a639" in the address "3e" of Table 2. Five more UserID-password pairs are stored in a similar way in the look up Table 2.

Table 1. Example showing the conversion of user IDs, and passwords into message digests, and an addresses, with hash functions.

User	ID	PW	h(PW)	ID⊕PW	h(ID⊕PW)
1	a6c26	12ae5	a639d	b86	3e
2	a311f	221ab	44d46	81f	90
3	ed011	22131	a3171	dfe	8c
4	15cab	320f1	7d5ff	27c	af
5	87a1b	775ed	0454b	f0f	93
6	1acf6	523c0	9b4fc	48f	ba
…	…	…	…	…	…

Table 2. Look up table storing the hashing of the passwords.

YX	0	1	2	3	4	5	6	7	8	9	A	B	c	d	e	f
0																
1																
2																
3															**a639d**	
4																
5																
6																
7																
8													**a3171**			
9	**44d46**			**0454b**												

(*continued*)

fridaya@uef.fi

Table 2. (*continued*)

YX	0	1	2	3	4	5	6	7	8	9	A	B	c	d	e	f
A																**7d5ff**
B											**9b4fc**					
C																
D																
E																
F																

Such a method is applicable to store very large quantities of UserID-Password pairs. During authentication cycles, the users provide their UserID/Password pair; the information extracted from the table at the corresponding address is then compared to the message digest provided by the password. When SHA-1 is replaced by more powerful hashing functions (ex: SHA-3), such look-up tables are much more secure than the tables directly storing UserID-Password pairs. The method can handle low levels of collisions, i.e. multiple messages digests located in the same address, or multiple Users sharing the same message digest.

Vulnerability of the Method. When look-up tables storing message digests are lost to the enemy, information can also be lost overtime with password guessing methods, big data analysis, and brute force hashing of commonly used passwords. For example, if the attacker knows that a particular user trends to use a set of passwords, the attacker can hash these passwords, and see if the message digests exist in the table. When millions of users are involved, the loss of can have a considerable impact, with legal exposure for the operator. Other issue of this password manager is the potential loss of the password by the user. The password manager cannot erase the message digest because the address in the look up table is also lost. One remedy is to implement a counter, which is incremented every time a user forgets the password. When the user login after the loss, its password can be concatenated with an index that is incremented at each loss. If the previously lost password is used again, the concatenated version with the index will not be recognizable.

2.2 Physical Unclonable Functions

PUFs generate patterns from hardware components that act as human fingerprints, which are unclonable, random, and make each device authenticable from each other [18–25]. The design of the PUF exploits the manufacturing variations created during their fabrication. In this paper, we are calling challenges, or initial responses, the patterns generated upfront by the PUF, while the responses are generated during each access control cycle. The authentications are positive when the matching challenge-response-pairs (CRPs) are high, with a low rate of error, i.e. mismatch. Methods to design PUFs include the use of logic components with gate delays, arbiters, and ring oscillators [18–20]. A PUF designed with ring oscillators is shown in Fig. 2.

fridaya@uef.fi

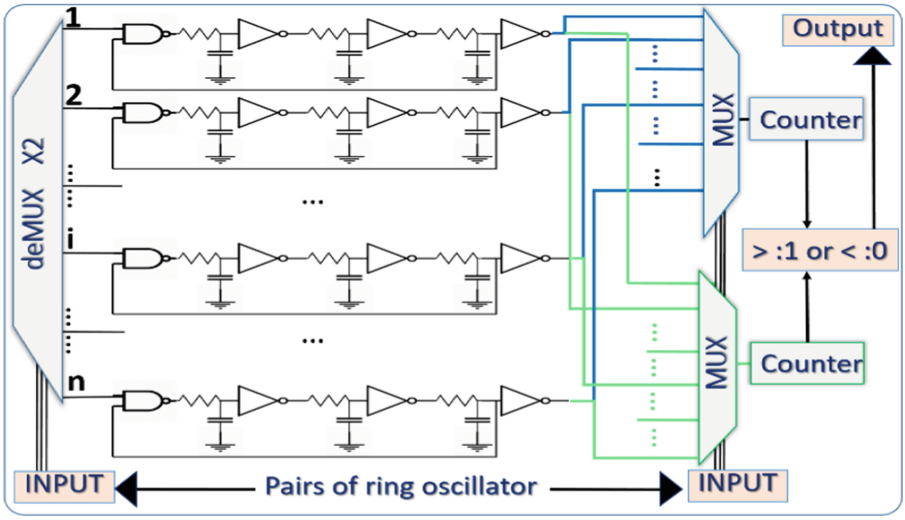


Fig. 2. Block diagram of a PUF designed with ring oscillators.

Pairs of ring oscillators selected among the n possible pairs, are analyzed using two counters. If the number of oscillations of the ring going to the top MUX is greater than the number of oscillations of the ring going through the bottom MUX, the challenge (during initial set up, or response during authentication) is a "1", if not, the challenge (or response) is a "0". If for example, the address of 128 successive pairs is transmitted to the PUF, 128-bit challenges (or responses) are generated.

Memory structures [46–55], SRAM [46, 47], DRAM [48], Flash [49, 50], ReRAM [51–53], and MRAM [54, 55], are also excellent elements to generate strong PUFs. The method to generate PUFs from SRAM arrays is to subject the device to power-off-power-on cycles. A significant proportion of the cells always restart as either a 0, or a 1, thereby generating CRPs. One of the generic methods to generate CRPs from memory devices is to characterize a particular parameter $\mathcal{P}$ of the cells of the array. The values of parameter $\mathcal{P}$ vary cell to cell, and follow a distribution with a median value **T**. In order to generate challenge and response pairs, all cells with $\mathcal{P} < \mathbf{T}$ generate "0" states, and the others generate "1" states. The resulting streams of data become PUF CRPs to authenticate the array.

Various PUFs have different figures of merit, and strength against attackers. A hostile party cannot simply read the entire PUF array of a strong PUF, and use the reading to communicate with a network of client devices. The cloning of the entire array would be a security threat; however, the level of randomness of such a component is so high that this type of attack is highly unlikely. The use of ternary PUFs reduces the challenge-response-pair (CRP) error rates [56–60].

fridaya@uef.fi

3 Password Managers with PUFs

3.1 Password Generators with PUFs

One time password generators can be designed with addressable PUF generators (APG) [45, 61]. The block diagram of the APG is shown in Fig. 3.

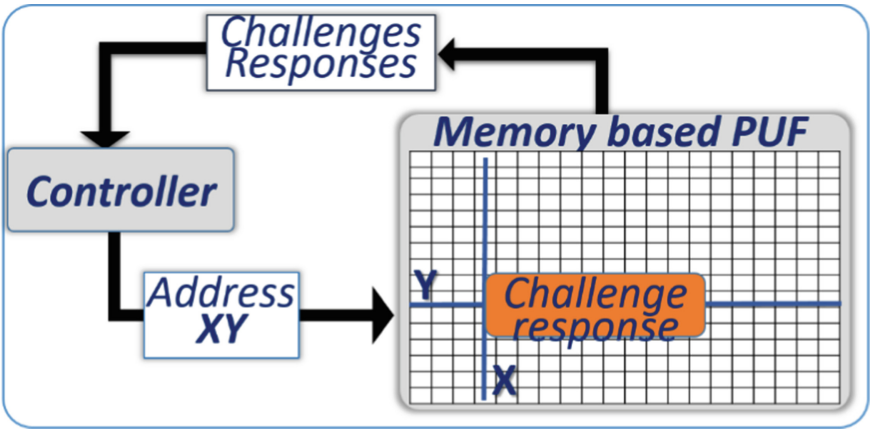


Fig. 3. Block diagram of the APG

In this architecture, the controller sends an address XY to the memory array to read a block of cells from which challenges and responses are generated. For a SRAM PUF, 0 or 1 are read in the block of cells located at address XY after power-off-power-on cycles. For other memory PUFs, parameter $\mathcal{P}$ of the block of cells located as XY is characterized. The ring oscillator of section II-B is also used as APG. In this case series of pairs are characterized at each address. The CRP generation varies with the relative value of the parameters within the multiple cells that are selected at a particular address. Therefore, as a particular cell could be a "0" when part of one group of cells, and a "1" when part of a different group, or when read with different instructions.

As described in [45, 61], APGs can be used to generate temporary passwords, and authenticate users. A block diagram of a temporary password generator with APG is shown in Fig. 4:

- For *password generation*, a random number ***TRN*** is generated and concatenated with the UserID, to feed the hash function. An address XY is extracted from the message digest to point on a particular block of cells in the APG. The reading of the PUF generates challenges, i.e. the initial responses. The resulting temporary password is the combination of ***TRN***, and the challenges.
- For *authentication*, the user transmits the password and its user ID. The same ***TRN*** and ***UserID*** are generating the same message digest, the same address ***XY***; the same block of cells is read to generate responses. These responses are matched with the challenges provided as part of the temporary password. If the CRP error rate is low, the authentication is positive.

fridaya@uef.fi

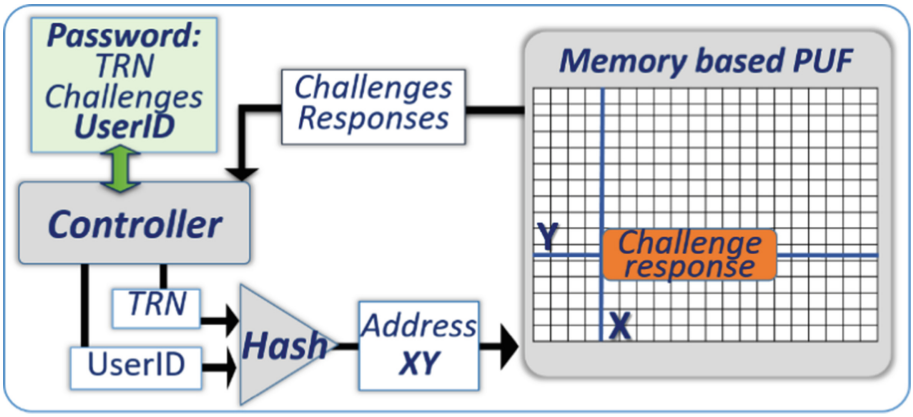


Fig. 4. Block diagram of a temporary password generator with APG

Multi-factor authentication and encryption was included in our implementation to protect the password. In the application the PUFs reside in the server, therefore the device is not subject to significant thermal and other environmental variations, this results in low CRP error rates. The method uses new random numbers ***TRN*** at each operation for one-time use of the passwords and concatenate ***TRN,*** with an index to eliminate the relevance of previous used passwords.

3.2 Password Manager Architecture with PUFs

Unlike the temporary password generator shown in section III-A, a password manager allows the user to pick its own password *PWj*, and user ID *Userj*. The design of password managers with PUFs is more complicated than designing password generators with PUFs. In the architecture presented in this paper, addressable PUF generators (APG) are combined with the hash function based password manager described in section II-A. A block diagram describing the scheme is shown in Fig. 5.

- During *enrollment*, the hashing of ***Userj*** $\oplus$ ***PWj*** results in the message digest ***h (Userj*** $\oplus$ ***PWj)*** from which the address ***Addi*** is extracted. The hash of password ***PWj*** generates the address ***XY*** in the APG. In our implementation, the memory space of the APG is an array of 4096 $\times$ 4096 cells, the first 12 bits of the message digest were used to find the ***X*** in the array, the next 12 bits to find the ***Y***. Starting from this address ***XY***, challenges are generated from the n-cells following this address. The parameter ***P*** is then measured, and challenges ***Chj***, the stream of bits *{Ch1, Ch2, ...Chn}*, are generated. The challenges ***Chj*** are finally stored in the look up table at the address ***Addi***.
- The *authentication* cycles follow the same step by step scheme, resulting in the same address ***Addi***, and a stream of responses ***Rej***. These responses are matched with the challenges ***Chj***. Assuming that the CRP error rates are low enough, the authentication is positive.

fridaya@uef.fi

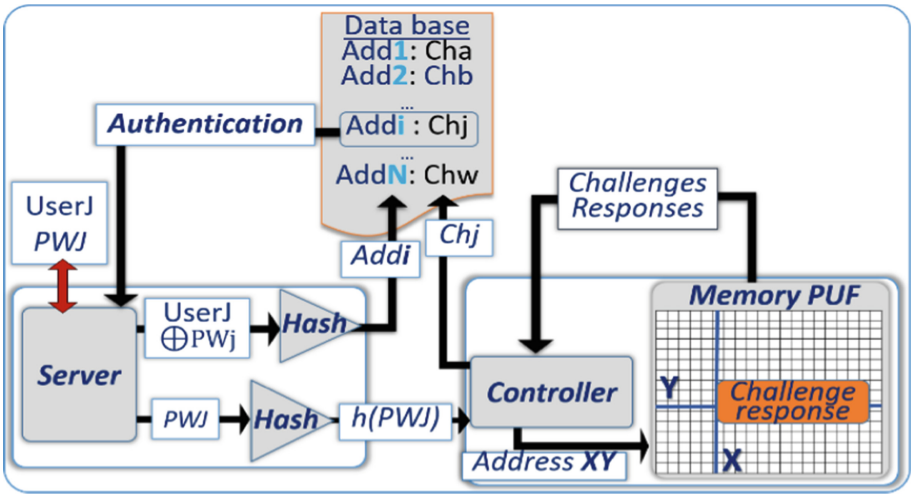


Fig. 5. Block diagram showing a password manager with APG.

Since hash functions are one-way functions, it is impossible to deduce the input of the hash function by looking at the address of a PUF array. Unlike traditional data storage units, the memory arrays used in the APG do not store information; CRPs cannot be generated from these PUFs without knowing the necessary instructions. If an attacker is able to have access to the database, the knowledge of the challenges do not disclose the message digest of the password. This offers additional layers of security, and makes password guessing more difficult.

3.3 Algorithm of Password Managers with PUFs

A summary of the algorithm developed to store the user ID and password in the look-up table with APGs is presented in Table 3:

Table 3. Algorithm to generate, and store challenges, replacing the passwords.

Step	Description of the instructions	Data stream/information	Where
1.0	*Hashing of the password* ***PWj***	***h(PWj)***	***Server***
2.0	***h(PWj)*** *converted into the address* ***XY***	***h(PWj)*** → ***X Y***	***APG***
3.0	*Select n-cells in the array located after* ***XY***	***{n-cells}***	***APG***
3.1	*Measure parameter* $\mathcal{P}$ *for the n-cells*	***{P1, P2, …, Pn}***	***APG***
3.2	*Generate the challenges* ***Chj*** ϵ *{0, 1}*	***Chj = {C1, C2, …, Cn}***	***APG***
4.0	*XORing the user ID* ***Userj*** *with* ***PW***	***Userj*** ⊕ ***PW***	***Server***
5.0	*Hashing of* ***Userj*** ⊕ ***PW***	***h(Userj*** ⊕ ***PW)***	***Server***
6.0	***h(Userj*** ⊕ ***PW)*** *converted into address* ***Addi***	***h(Userj*** ⊕ ***PW)*** → ***Addi***	***Server***
7.0	*Store* ***Chj*** *in the database at address* ***Addi***	***Addi; Chj***	***Database***

fridaya@uef.fi

- *Step-1:* The password ***PWj*** feeds the hash function to generate a message digest ***h (PWj)***. Additional security features are not described in this paper, which include the use of cryptography to protect the passwords, and multi-factor authentication to protect the user. The hash function can be based on some standard hash algorithms (SHA).
- *Step-2:* The message digest of step one is used to generate the coordinate ***XY*** of the memory array. As suggested above, this could be the first 24-bits of the message digest, if the size of the memory array is 4,096 × 4,096. Different possible schemes are presented in Sect. 4.
- *Step-3:* The generation of a stream of challenges ***Chj = {Ch1,….Chn}*** from the ***n-cells*** located in the memory array which passes the address ***XY,*** can be done differently based on the type of memory technology, and various generation protocols. Some methods, as suggested Sect. 2.1, can be based on the measurement of a particular parameter Ƥ.
- *Step*-4 and 5: The user ID ***Userj***, and the password ***PW*** are Xored to generate the data stream ***Userj*** ⊕ ***PW***. The Boolean function XOR can be replaced by other methods to create a data stream from the user ID.
- *Step*-6: After the hashing of ***Userj*** ⊕ ***PW***, an address ***Addi*** is generated from the message digest. If, for example, the look-up table has 4,096 × 4,096 positions, the first 24 bits of the message digest can be used to generate ***Addi***.
- *Step*-7: The challenges ***Chj = {Ch1,….Chn}*** is stored at the address ***Addi***. In case of a collusion, i.e. multiple users at the same address, multiple challenges are stored at the same location, which does not make the authentication more difficult.

A summary of the algorithm developed to authenticate a user with password is summarized in Table 4.

Table 4. Algorithm to authenticate users with CRPs.

Step	Description of the instructions	Data stream/information	Where
1.0	*Hashing of the password* ***PWj***	***h(PWj)***	***Server***
2.0	***h(PWj)*** *converted into the address* ***XY***	***h(PWj)*** → ***X Y***	***APG***
3.0	*Select n-cells in the array located after* ***XY***	***{n-cells}***	***APG***
3.1	*Measure parameter* ***Ƥ*** *for the n-cells*	***{P1, P2, …, Pn}***	***APG***
3.2	*Generate responses* ***Rej*** ϵ *{0, 1}*	***Rej = {C1, C2, …, Cn}***	***APG***
4.0	*XORing the user ID* ***Userj*** *with* ***PW***	***Userj*** ⊕ ***PW***	***Server***
5.0	*Hashing of* ***Userj*** ⊕ ***PW***	***h(Userj*** ⊕ ***PW)***	***Server***
6.0	***h(Userj*** ⊕ ***PW)*** *converted into address* ***Addi***	***h(Userj*** ⊕ ***PW)*** → ***Addi***	***Server***
7.0	*Read* ***Chj*** *in the database at address* ***Addi***	***Addi; Chj***	***Database***
8.0	*Find Hamming distance* ***Hj Chj*** *to* ***Rej***	***Hj***	***Server***
9.0	*Authentication if* ***Hj*** *below the threshold*	***Yes*** *or* ***No***	***Server***

- *Step*-1 to 5: These steps are similar to the one described above, to program the look-up tables. Rather than generating challenges, responses are generated by the APG at the same addresses.
- *Step*-6 to 9: The challenges previously stored in the look-up table are compared with the freshly generated responses. When the Hamming distance ***Hj*** between challenges and responses is bellow a set threshold, CRP error rates are low enough, the authentications are positives.

fridaya@uef.fi

4 Password Manager with Ternary PUFs

4.1 Ternary APGs

Typically, PUFs experience 3 to 10% CRP error rates. This is usually acceptable for authentication when the length of challenge-response-pairs is high enough. The ternary PUFs are based on three states, (0, 1, X), X being the fuzzy state that is used to characterize the cells that are not predictable. The CRP error rate measured on the non-fuzzy states is significantly lower. The results shown in Fig. 6 are based on the characterization of commercially available SRAM memories. Due to micro-manufacturing differences, the flip-flop of each SRAM cell randomly emerges from power-off-power-on as a 0, or a 1. The vast majority of the cells are responding in a predictable way, therefore acting as a "finger print" of the device. We characterized 32 KB SRAMs, and measured 3 to 5% CRP error rate at each power-off power-on cycles. We then subjected the memory cells with successive power-off power-on cycles, and masked with a ternary state "X" the cells having CRP errors. As a function of the number of queries, the percentage of the cells that are not stable (the one marked "X"), are plotted at the bottom portion of Fig. 6. The corresponding CRP error rates of the non-fuzzy states are plotted at the top portion of Fig. 6. Below 10 cycles, the CRP error rates are in the 0.3 to 5% range, above 80 cycles the rates are in the 0.01% to 0.03% range. In our software implementation, the cells were cycled 100 times during enrollment, 12% of the cells are marked X, 44% are "0"s, and 44% are "1"s. The average CRP error rates of the non-fuzzy cells were at 0.01%. Two possible modifications of the protocol described in Sect. 3 are suggested to process the ternary PUFs.

First method: To store the ternary states in the look up table. In this method, during the enrollment of the password ***PWj***, the server generates streams of trits (0, 1, X) ***Chj*** from the PUF array. ***Chj*** is stored at the address ***Addi*** in the look-up table shown in the block diagram of Fig. 4. In our implementation, the "0" are stored in the look up table as (01), the "1" as (10), and the fuzzy "X" as (11). During authentication, the response generation is binary, only "0"s and "1"s are extracted. Only the non-fuzzy states stored in the look up table are counted for CRP error rates estimate. This protocol is implementable with SRAM PUFs, and with memory PUFs that are based on a parameter ***P***. Within the n-cells that are used for challenge generation, at a particular address ***XY*** of the array, the bottom 1/3 are "0"s, the middle 1/3 are "X"s, and the top 1/3 are "1"s. During response generation, the same n-cells are sorted in the bottom half for the "0"s, and into the top half for the "1"s. In this method, the error rates are extremely low; no error correcting methods are needed. The latency of the enrollment of each new password is here higher, while the authentication cycles are fast.

Second method: Upfront characterization of the entire array. In this method, the entire PUF array is characterized upfront to identify all fuzzy cells of the array, and mark them with an "X". During the enrollment of the passwords, and the corresponding challenge generation, the fuzzy cells located pass the address ***XY*** are ignored, and the n-bits kept are non-fuzzy, i.e. binary. The resulting challenges consist of streams of bits that are stored in the look-up table. During response generation, the same fuzzy cells are ignored, resulting in binary streams that are compared with the challenges stored in

fridaya@uef.fi

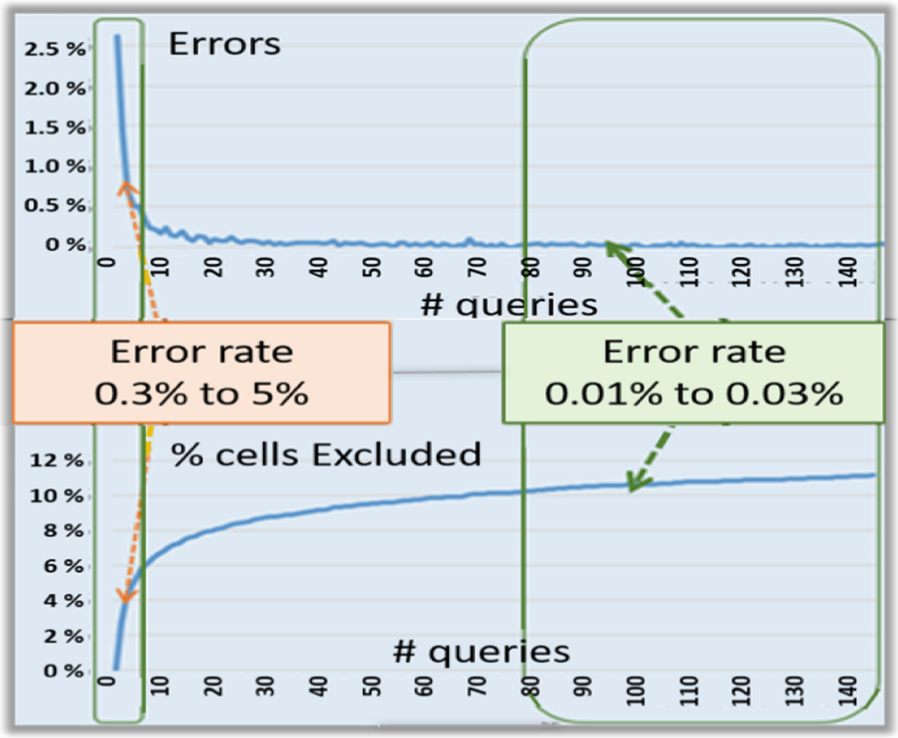


Fig. 6. CRP error rate reduction by masking the fuzzy cells of a SRAM.

the lock up table. The lengthy upfront characterization is done only once, both enrollment and authentication cycles are relatively fast.

The first method is more desirable than the second one, on a security standpoint. The entropy of the stored challenges ***Chj*** is higher. The number of possible configurations of a stream of n trits is 3^n versus 2^n for a stream of bits. Attacks such as password guessing are therefore more complex.

4.2 Double Addressing with Ternary APGs

In the scheme presented above in section IV-A, only the content of what is stored in the address Addi incorporate the patterns generated by the PUFs. Small error rates, below 7%, in CRPs do not add latencies in the matching algorithm. The use of ternary PUF reduces the error rates, and latencies. The question addressed in this section is the potential use of PUF CRPs to replace both password *PWj* by *Chj*, and address *Addi* by the address *Cddi*, also extracted from the PUFs. An error in the address would send the search engine to the wrong part of the look up table, which could be difficult to handle. The block diagram of a ternary PUF password manager with double addressing is shown in Fig. 7.

fridaya@uef.fi

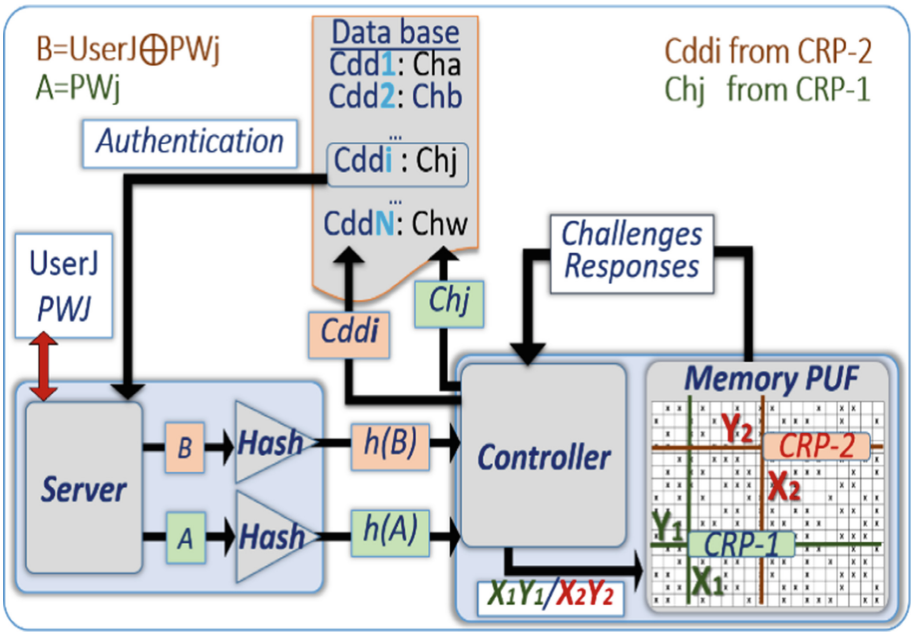


Fig. 7. Block diagram of a password manager with two addressing schemes.

- The *enrollment* of a new UserID-Password pair has two steps. The first one, as described previously, where the APG generated at the address $\boldsymbol{X_1Y_1}$, the challenges Chj from the password ***PWj***, to store it in the look up table. In the second step, the message digest $\boldsymbol{h(Userj \oplus PWj)}$ is used to find a second address $\boldsymbol{X_2Y_2}$ in the APG. The challenges extracted at this second address are then used to find the address ***Cddi*** in the look up table where ***Chj*** is stored. For example, if the size of the look-up table is 4096 $\times$ 4096, the first stream of the 12 bits of the challenges is used as the first coordinate, and the second stream of 12 bits for the second coordinate.

As a result, not only the content of the look up table is generated from the PUF, but the addressing system is also extracted from the PUF.

- The *authentication* of the user follows a similar method. The password generates, at the address $\boldsymbol{X_1Y_1}$, the responses ***Rej***, and the message digest $\boldsymbol{h(Userj \oplus PWj)}$ is used to find the address ***Cddi*** from the responses generated at the address $\boldsymbol{X_1Y_1}$. If ***Cddi/Rej*** is matching the reference ***Addi/Chj***, the authentication is positive.

The use of ternary PUFs, as described below in section IV-C, is important to mitigate potential mismatches between ***Addi*** and ***Cddi*** in case of CRP errors, which could result in pointing to the wrong stream in the look up table.

fridaya@uef.fi

4.3 Error Matching Algorithms

The two figures of merit of the password generator that are analyzed in this section are the false rejection rate (FRR), and latency of the matching algorithm. The objective of the matching algorithm is to minimize FRR to an acceptable level, while keeping the latency low enough. To keep the user experience enjoyable, we target an FRR below 10^{-4}, and latencies under 3 s.

With PUFs having CRP error rates of 3%, the rate of bad addressing is estimated with statistical models such as Poisson.

- If the CRP error rate is 3%, with 24-bit address, the probability to get at least 1bit error in the address is 51%, and the probability to get at least 2 is 22%. This error rate is prohibitive.
- With ternary PUFs, if the CRP error rate is 0.03%, the probability to get at least 1 bit error is 0.7%, and the probability to get at least 2 is $2.6 \cdot 10^{-5}$. The probability to get at least 3 errors is only $6 \cdot 10^{-8}$.

To find the matching address we designed our codes based on the algorithm shown in Fig. 8. The challenge stored in the look up table, at the address ***Cddi*** is ***Chi***. The authentication is positive when the hamming distance between ***Rej***, the response generated by the APG from the password ***PWj***, and ***Chi*** is below an acceptable threshold ***Hmax***.

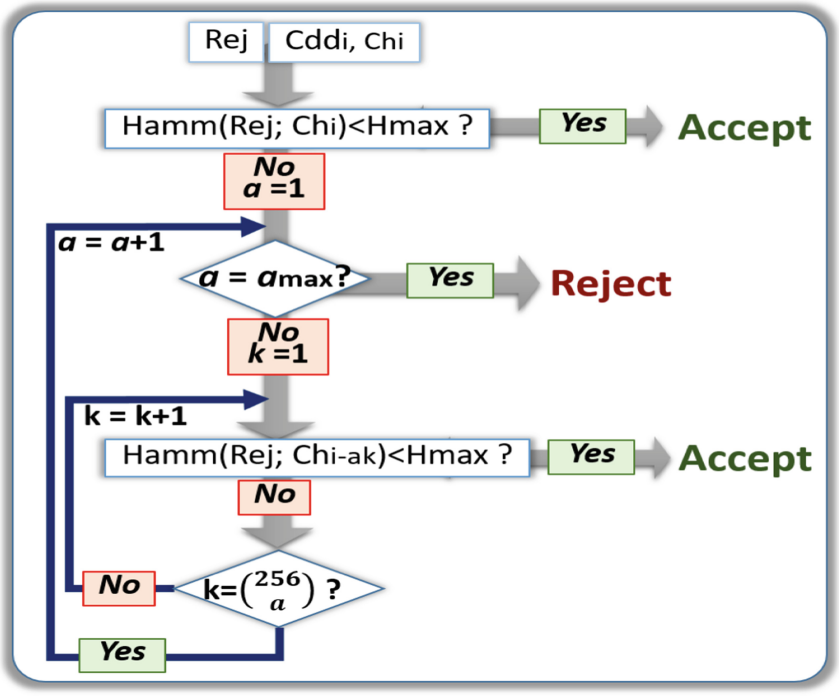


Fig. 8. Algorithm to find the matching address in the look up table.

fridaya@uef.fi

If negative, all ***k*** challenges ***Chi-1k*** located in the look-up table at the addresses ***Cddi-1k*** are considered. The addresses ***Cddi-ak*** have a Hamming distance of one with the address ***Cddi***. The authentication is positive when the Hamming distance between one of the k challenges ***Chi-1k***, and ***Rej*** is below the threshold ***Hmax***.

If negative, the process iterates to consider all ***k*** challenges; ***Chi-ak*** located in the look up table at the addresses ***Cddi-ak*** having a Hamming distance ***a*** with ***Cddi***. The iteration stops when the Hamming distance reaches the threshold ***amax***, therefore rejecting the authentication.

With 24-bit addresses, there are only 24 addresses with a Hamming distance of 1 within a given response, so step-2 is quick. The likelihood to have to go to the next step, ***a = 2***, is small (If CRP error rate is 0.03%, the probability is $2.6 \cdot 10^{-5}$). The number of addresses with a Hamming distance of 2 with ***Cddi*** is $\binom{24}{2} = 273$.

Based on the ternary SRAM PUF that we characterized, if $a_{max} = 2$, the resulting false rejection rate (FRR) of the protocol is $2 \cdot 10^{-5}$, which is the probability to have the hamming distance at 2 or higher. When $a_{max} = 3$, the resulting FRR is $6 \cdot 10^{-8}$, with the probability to have the hamming distance at 3 or higher. Such levels of FRRs are very small, and acceptable to users. The latencies of the methods are also acceptable.

To find the Hamming distance between two 256-bit long streams usually takes less than 10 clock cycles, or 10 ns for a 1 GHz processor. With the ternary PUFs that we characterized, and ***a*** = 2, the sever needs to find the Hamming distance between ***Rej*** and ***Chi***, then the 24 Hamming distances between ***Chi-1 k*** and ***Chi***, a total of 24 searches, or 250 ns for a 1 GHz processor. With ***a*** = 3, the maximum number of searches needed to iterate is 297, and 2.97 μs latency.

This protocol is also scaling well when the size of the lock-up table increases. For example, look-up tables with 36-bit addresses can store 68G addresses, which is 65,000 larger than a look-up table bit 24-bit addresses. There are only 36 addresses with a Hamming distance of 1 within a given response, and $\binom{36}{2} = 630$ addresses with a Hamming distance of 2.

4.4 Double Addressing with Two Ternary APGs

The same method as the one described above can be implemented with two different ternary APGs, as shown in Fig. 9. One APG, in green in the figure, generates the challenges *Chj* to be stored in the look-up table based on the passwords *PWj*. The second APG, in red in the figure, generates the addresses *Cddi* in the look-up table based on *Userj* ⊕ *PWj*. To enhance security the two APGs can use different PUF technologies. The first APG for *Chj* generation is less sensitive to CRP error rates, and does not have to be based on Ternary PUFs. The control circuitry shown in Fig. 8 can also be separated into two different circuits to increase the protection of the system.

fridaya@uef.fi

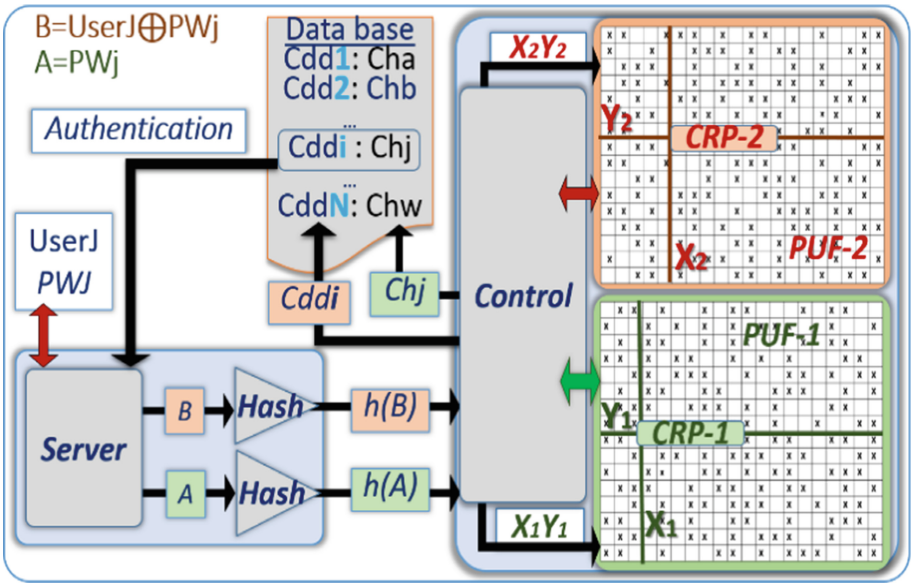


Fig. 9. Block diagram of a password manager with two ternary APG, driving two addressing schemes.

4.5 Double Addressing with PUF Based on Ring Oscillators

PUF designed with component different than SRAM were studied, such as the ones designed with gate delays and ring oscillators. The diagram of such password manager is shown in Fig. 10. Pairs are formed to group the stream of bits, which are generated from the message digest. Each pair generates a challenge (or response), in such a way that a message digest of 512 bits can generate 256 CRPs. Ternary states can be generated with ring oscillator PUFs. During challenge generation, a particular pair of ring oscillators generates the state "0" or "1" when the difference between the numbers of oscillation per milli-second of the two rings is large enough. When this difference is below a given threshold, the pair generates a fuzzy state "X". During the response generation, all pairs are generating only "0", or "1".The CRP matching estimate is based on the non-fuzzy pairs that were generated during challenge generation, while the pairs with fuzzy state are ignored. This results in a reduction of the CRP error rates.

4.6 Entropy Enhancement

In the software implementation presented above, the randomness, and entropy was increased with longer message digests. The protocols presented in Sects. 3 and 4 are based on the conversion of message digest into addresses. The example given is the one of an APG based on an array of 4,096 × 4096 cells. The first 12 bits are used to find the ***X*** coordinate, and the next 12 bits are used to find the ***Y*** coordinate. The n-cells located after that address are used to generate PUF challenges (or responses) that consist of n-bits. Message digests contain long streams of bits, typically 512, so only 24

fridaya@uef.fi

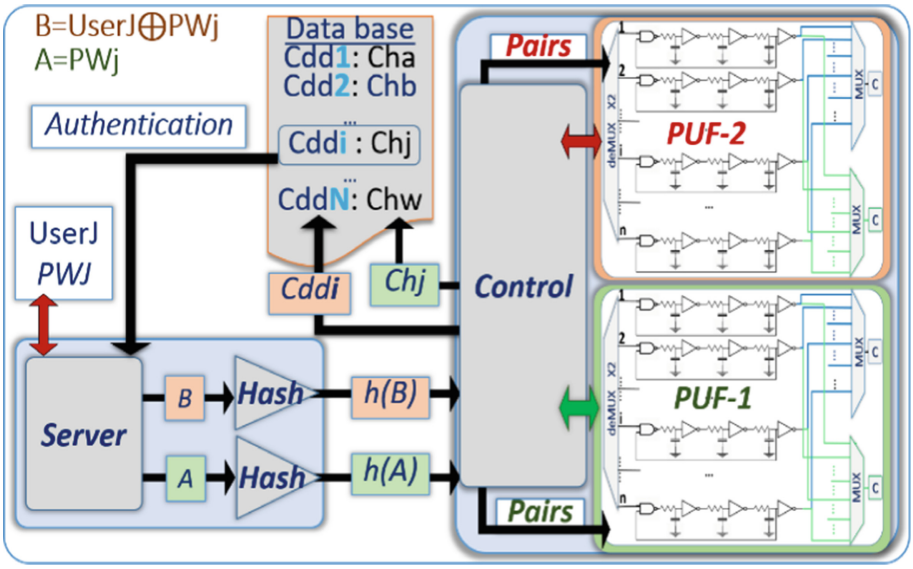


Fig. 10. Block diagram of a password manager with two PUFs based on ring oscillator driving two addressing schemes.

bits out of these streams are used in the protocol. The opportunity to increase randomness is to find multiple addresses in the APG from the message digests, and to reduce the number of cells involved at each address. Assuming that the size of the streams is kept at **n**, a total of f addresses in the APG are selected from each message digest, and m cells are selected by address to generate the n-bit challenges (or responses), with $n = fm$. For example if $n = 512$, and $f = 16$ (addresses selected from the message digests), then $m = 32$ bits are generated at each address. The number of bits of the message digest used for addressing in the protocol is now $16 \times 24 = 384$. There are 2^{24} different 24-bit possible addresses, and entropy of 24 ($\text{Log}_2(2^{24}) = 24$). The number of possible ways to pick 16 addresses having 24 bits is $(2^{24})^{16}$, an entropy of 384.

5 Conclusion and Future Work

As part of the design of password managers, hash functions are effective to replace the storage of the password into message digests. Hash functions can also protect the address in look up tables where these message digests are stored. Such methods offer reasonable protections against the exposure of the look up tables to malicious parties. However, certain attacks, such as password guessing, are still effective against such schemes, because the hash functions are usually known, and public. The use of encrypted hashing functions is stronger, but can create a false sense of security when the cryptographic keys are also lost to the hackers. The combination of hashing functions with addressable PUF generators allows the replacement of the message

fridaya@uef.fi

digests of the passwords by challenge-response-pairs, which provide an additional layer of security. Even if the hash functions are known, the PUFs, which can be unclonable, random, and unique, mitigate attacks against lost look up tables. The task of malicious entities is thereby much more challenging; in order to break the system they have to steal the hardware containing the PUFs, or find a way to download a representation of the PUFs. The suggested ternary PUF architecture handles the fuzzy states, which reduces the challenge-response pair error rates, making the method more reliable, and reduces latencies. The proposed matching algorithms are critical to protect the addresses in look up tables using ternary PUFs, with low false rejection rates (FRR) at low computing latencies. The implementation using ternary SRAM PUF is showing FRR lower than the part per million, and latencies in the microsecond range. This suggested additional protection can use a single PUF, two PUFs, or PUFs based on different technologies. The hardware-software implementation that was developed, and described in this paper, uses existing components, and is extremely low cost. This method has the objective to complement other cryptographic protections, not to replace them. For example, the use of multi-factor authentication, and the encryption of the passwords with public key infrastructure (PKI) are recommended.

Moving forward the suggested algorithms should be further improved, and use stronger PUFs beside SRAMs. The functionality of commercial password managers, including password expiration and replacement, verification of their strength, firewall protection, and others, will be incorporated, and tested in future prototypes. SRAM PUFs are relatively easy to break, and the proposed password managers should be designed with tamper resistant PUFs. Examples of stronger candidates include memristor and ReRAM based low power PUFs [51–53, 56, 57]. Finally, the development boards used in this work should be replaced with application specific integrated circuits (ASIC), which incorporate hardware protection against differential power analysis, and other side channel attacks, and using a cryptoprocessor, to provide encryption of the passwords with PKI, and hardware implementation of SHA-3.

Acknowledgments. The author is thanking the contribution of several graduate students at Northern Arizona University, in particular Sareh Assiri, Christopher Philabaum, Duane Booher, Vince Rodriguez, Ian Burke, and Mohammad Mohammadi.

References

1. Jeong, Y.S., Park, J.S., Park, J.H.: An efficient authentication system of smart device using multi factors in mobile cloud service architecture. Int. J. Commun. Syst. **28**(4), 629–674 (2014)
2. Saxena, N., Choi, B.J.: State of the art authentication, access control, and secure integration in smart grid. Energies **8**(10), 11883–11915 (2015)
3. Zhang, M., Zhang, J., Zhang, Y.: Remote three-factor authentication scheme based on Fuzzy extractors. Secur. Commun. Netw. **8**(4), 682–693 (2014)
4. Keane, J.: Security researcher dumps 427 million hacked Myspace passwords, July 2016. https://www.digitaltrends.com/social-media/myspace-hack-password-dump/
5. Morgan, S.: 2017 Cybercrime Report, Cybercrime damages will cost the world $6 trillion annually by 2021. Cybersecurity Ventures, Herjavec Group (2017)

fridaya@uef.fi

6. Target: Data stolen from up to 70 million customers: USA Today. https://www.usatoday.com/story/money/business/2014/01/10/target-customers-data-breach/4404467/
7. Blocki, J., Harsha, B., Zhou, S.: On the economics of offline password cracking. In: IEEE Symposium on Security and Privacy (SP) (2018)
8. Lee, H.W., Noh, M.J., Chol, H.M., Feng, X.: Password system, method of generating password, and method of checking password. Patent application US2009/0228977A1 (2009)
9. Fitzgerald, J.: Systems and methods for providing a covert password manager. US patent 9,571,487 B2 (2017)
10. Harper, R.: STARTS password manager. Patent publication, US2005/0125699A1 (2005)
11. Mimlitsch, J.: User-administrated single sign-on with automatic password management for WEB server authentication. Patent publication US2007/0226783A1 (2007)
12. Safriel, M.: Portable password manager. Patent publication US2004/0193925A1
13. Tsai, C.-S., Lee, C.-C., Hwang, M.-S.: Password authentication schemes: current status and key issues. IJ Network Security (2006)
14. Zhang, Z., Yang, K., Hu, X., Wang, Y.: Practical anonymous password authentication and TLS with anonymous client authentication. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 1179–1191. ACM (2016)
15. Forler, C., List, E., Lucks, S., Wenzel, J.: Overview of the candidates for the password hashing competition. In: Conference on Passwords, Springer (2014)
16. Tsai, J.-L.: Efficient multi-server authentication scheme based on one-way hash function without verification table. Comput. Secur. **27**(3–4), 115–121 (2008)
17. Janzen, W.S.: Iterated password hash systems and methods for preserving password entropy. US patent 8,769,637 B2 (2014)
18. Pappu, R., Recht, B., Taylor, J., Gershenfield, N.: Physical one-way functions. Science **297** (5589), 2026–2030 (2002)
19. Jin, Y.: Introduction to hardware security. Electronics **4**, 763–784 (2015). https://doi.org/10.3390/electronics4040763
20. Gassend, B., et al.: Silicon physical randomness. In: Proceedings of the 9th ACM Conference on Computer and Communications Security, CCS'2002, pp. 148–160 (2002)
21. Naccache, D., Frémanteau, P.: Unforgeable identification device, identification device reader and method of identification. Patent US5434917, August 1992
22. Gao, Y., et al.: Emerging physical unclonable functions with nanotechnologies. IEEE. https://doi.org/10.1109/access.2015.2503432
23. Delavor, M., et al.: PUF based solution for secure communication in advanced metering infrastructure. ACR Publication (2014)
24. Herder, C., Yu, M.-D., Koushanfar, F., Devadas, S.: Physical unclonable functions and applications: a tutorial. Proc. IEEE **102**(8), 1126–1141 (2014)
25. Maes, R., Verbauwhede, I.: Physically unclonable functions: a study on the state of the art and future research directions. In: Towards Hardware-Intrinsic Security (2010)
26. Wang, D., Zhang, Z., Wang, P., Yan, J., Huang, X.: Targeted online password guessing: an underestimated threat. In: Proceedings of the ACM CCS, pp. 1242–1254 (2016)
27. Pass the Hash attack: Microsoft research as on 12 August 2015. http://www.microsoft.com/PTH
28. Bonneau, J., Herley, C., van Oorschot, P., Stajano, F.: Passwords and the evolution of imperfect authentication. Commun. ACM **58**(7), 78–87 (2015)
29. Taniguchi, M., Shiozaki, M., Kubo, H., Fujino, T.: A stable key generation from PUF responses with a fuzzy extractor for cryptographic authentications. In: 2013 IEEE 2nd Global Conference on Consumer Electronics (GCCE), Tokyo (2013)
30. Price, N.E., Sherman, A.T.: How to generate repeatable keys using physical unclonable functions, correcting PUF errors with iteratively broadening and prioritized search

fridaya@uef.fi

31. Boehm, H.M.: Error correction coding for physical unclonable functions. In: Austrochip, Workshop in Microelectronics (2010)
32. Yu, M., Devadas, S.: Secure and robust error correction for physical unclonable functions. IEEE Design Test Comput. Verifying Phys. Trustworthiness ICs Syst. **27**, 48–65 (2010)
33. Kang, H., Hori, Y., Katashita, T., Hagiwara, M., Iwamura, K.: Cryptographie key generation from PUF data using efficient fuzzy extractors. In: 16th International Conference on Advanced Communication Technology, Pyeongchang (2014)
34. Becker, G.T., Wild, A., Güneysu, T.: Security analysis of index-based syndrome coding for PUF-based key generation. In: 2015 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), Washington, DC (2015)
35. Hiller, M., Weiner, M., Rodrigues, L., Birkner, M., Sigl, G.: Breaking through fixed PUF block limitations with differential sequence coding and convolutional codes. In: TrustED'13 (2013)
36. Paar, C., Pezl, J.: Understanding Cryptography - A Text Book for Students and Practitioners. Springer, Berlin (2011)
37. Mel, H.X., Baker, D.: Cryptography Decrypted. Addison-Wesley, Boston (2001)
38. Pfleeger, C.P., et al.: Security in Computing, 5th edn. Prentice Hall, Upper Saddle River (2015)
39. Croman, K., Decker, C., Eyal, I., Gencer, A.E., Juels, A., Kosba, A., Miller, A.: On scaling decentralized blockchains. In: Springer International Conference on Financial Cryptography and Data Security, Berlin, Heidelberg (2016)
40. Luu, L., Narayanan, V., Zheng, C., Baweja, K., Gilbert, S., Saxena, P.: A secure sharing protocol for open blockchains. In: ACM SIGSAC Conference on Computer and Communication Security (2016)
41. Eyal, I., Gencer, A.E., Sirer, E.G., Renesse, R.V.: Bitcoin-NG: a scalable blockchain protocol. In: NSDI (2016)
42. Dorri, A., Kanhere, S.S., Jurdak, R.: Blockchain in internet of things: challenges and solutions. arXiv preprint arXiv: 1608.05187 (2016)
43. Gervais, A., Karame, G.O., Wüst, K., Glykantzis, V., Ritzdorf, H., Capkun, S.: On the security and performance of proof of work blockchains. In: ACM SIGSAC Conference on Computer and Communications Security (2016)
44. Zheng, Z., Xie, S., Dai, H.-N., Wang, H.: Blockchain challenges and opportunities: a survey. Int. J. Web Grid Serv. 1–25 (2016)
45. Cambou, B.: Addressable PUF generators for database-free password management system. In: CryptArchi (2018)
46. Holcomb, D.E., Burleson, W.P., Fu, K.: Power-up SRAM state as an identifying fingerprint and source of TRN. IEEE Trans. Comput. **57**(11), 1198–1210 (2008)
47. Maes, R., Tuyls, P., Verbauwhede, I.: A soft decision helper data algorithm for SRAM PUFs. In: 2009 IEEE International Symposium on Information Theory (2009)
48. Christensen, T.A., Sheets II, J.E.: Implementing PUF utilizing EDRAM memory cell capacitance variation. Patent No.: US 8,300,450 B2, 30 October 2012
49. Prabhu, P., Akel, A., Grupp, L.M., Yu, W.-K.S., Suh, G.E., Kan, E., Swanson, S.: Extracting device fingerprints from flash memory by exploiting physical variations. In: 4th International Conference on Trust and Trustworthy Computing, June 2011
50. Plusquellic, J., et al.: Systems and methods for generating PUF's from non-volatile cells. WO20151056887A1 (2015)
51. Chen, A.: Comprehensive assessment of RRAM-based PUF for hardware security applications. IEEE (2015). 978-1-4673-9894-7/15/IEDM

fridaya@uef.fi

52. Cambou, B., Afghah, F., Sonderegger, D., Taggart, J., Barnaby, H., Kozicki, M.: Ag conductive bridge RAMs for physical unclonable functions. In: 2017 IEEE International Symposium on Hardware Oriented Security and Trust (HOST), McLean (2017)
53. Korenda, A., Afghah, F., Cambou, B., A secret key generation scheme for internet of things using ternary-states ReRAM-based physical unclonable functions. In: Submitted to International Wireless Communications and Mobile Computing Conference (IWCMC 2018)
54. Zhu, X., Millendorf, S., Guo, X., Jacobson, D.M., Lee, K., Kang, S.H., Nowak, M.M., Fazla, D.: PUFs based on resistivity of MRAM magnetic tunnel junctions. Patents US 2015/0071432 A1, March 2015
55. Vatajelu, E.I., Di Natale, G., Barbareschi, M., Torres, L., Indaco, M., Prinetto, P.: STT-MRAM-based PUF architecture exploiting magnetic tunnel junction fabrication-induced variability. ACM Trans. **13**(1), 5 (2015)
56. Cambou, B., Orlowski, M.: Design of PUFs with ReRAM and ternary states. CISR 2016, April 2016
57. Cambou, B., Afghah, F.: Physically unclonable functions with multi-states and machine learning. In: 14th International Workshop on Cryptographic Architectures Embedded in Logic Devices (CryptArchi), France (2016)
58. Cambou, B., Telesca, D.: Ternary computing to strengthen information assurance, development of ternary state based public key exchange. In: Computing Conference, IEEE, London, July 2018
59. Cambou, B., Flikkema, P., Palmer, J., Telesca, D., Philabaum, C.: Can ternary computing improve information assurance? Cryptography, MDPI, February 2018
60. Cambou, B.: Physically unlonable function generating systems and related methods. US patent 9,985,791 (2018)
61. Cambou, B.: Encoding ternary data for PUF environment. US patent 10,050,796 (2018)

fridaya@uef.fi

Capabilities of Email Forensic Tools

Ahmad Ghafarian(✉)

Department of Computer Science and Information Systems,
University of North Georgia, Dahlonega, GA 30005, USA
Ahmad.Ghafarian@UNG.edu

Abstract. In many digital forensic investigations, the analysis of email data may be necessary. However, this may violate the privacy of the person whose emails are being analyzed. To avoid this violation, investigators generally run keyword searches to reveal only those emails that contain at least some of the keywords. Forensics investigators need to choose an email forensics tool that supports comprehensive keyword search and reporting. Selection of an appropriate tool requires knowledge of the capabilities of the tool. In this research, our aim is to help the investigators in the tool selection process by experimentally examining four different email forensics tools and evaluating their keyword search and report generation capabilities as well as some other features. The other features include but not limited to email format they support, size of the file they accept, format of the reports, etc. We use Enron email dataset as input to the tools.

Keywords: Email forensic · Report generation · Keyword search · Tools · Enron · Search filtering · Capabilities

1 Introduction

E-mail forensics refers to the study of email details including sender, receiver, content, date/time of transmission, detailed record of e-mail transaction as well as the intent of the sender [10]. However, by looking at the suspects' entire email, the investigator may violate privacy of the individuals whose emails are being analyzed. Generally, forensics investigators use email forensics tools to extract details of email communication between senders and the recipients. To avoid privacy violation, the investigator run keyword searches to reveal only those emails that contain the keywords. Therefore, the investigator needs to have knowledge of the keyword search capabilities of the email forensics tools. In an effort to assist forensics experts in selecting an appropriate tool, we experimentally evaluate the search and report generation capabilities of several email forensics tools, namely, *Add4Mail, Paraben Email Examiner (EMX)*, *MailXaminer*, and *eMailTrackerPro.* We use the Enron email dataset file that is maintained by MIT [8]. We import the email files to these tools, run the tools, and evaluate their keyword search filtering, report generation and other capabilities. The other features include but not limited to email format they support, size of the file the accept, format of the reports, etc.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 514–528, 2019.
https://doi.org/10.1007/978-3-030-22868-2_38

fridaya@uef.fi

The rest of this paper is organized as follows. Section 2 provides literature review. Section 3 describes tools, technologies and the email forensics process. Section 4 covers analysis of the results. Conclusion appears in Sect. 5. Section 6 shed light on possible extension of this work.

2 Literature Review

Email forensics is an integral part of many digital forensics investigations. Several researchers and practitioners have studied various aspects of email forensics such as tool development, technologies, and methodologies to assist forensics investigators. In this section, we review some of the most recent researches in this field.

The paper by Banday [3] illustrates e-mail architecture from forensics perspective. The author describes the roles and responsibilities of different e-mail actors and components, itemizes meta-data contained in e-mail headers, and lists protocols and ports used in it. In addition, the author provides the details of various open source and proprietary email forensics tools and their capabilities. The article by Meghanathan et al. [13], provides a comprehensive survey of email and network forensics tools as well as their specific features. The paper by Paglierani et al. [14] defines a general methodology for email forensics and show proof of concept with evidence results. Other study of the main characteristics of email forensics tools and techniques can be found in [4, 10]. Cohen [5] presents an automated technique for bulk-email analysis and presentation to aid in evidence interpretation. Devendran et al. [6] theoretically have examined a set of common features of some popular email forensic tools.

The paper by Hajidj et al. [9] presents an e-mail forensic analysis software tool, which is developed by integrating existing statistical and machine learning techniques complemented with social networking techniques. The authors proposed a framework that offers different functionalities ranging from e-mail storing, editing, searching, and querying. However, we found no report on the widespread use of their framework.

Stolfo and Hershkp [16] have developed a data-mining tool that visualizes a very wide range of detailed analysis of email and email flows derived from an archive of email in a variety of formats. The authors claim that their tool may be leveraged in many applications. In a similar research, Khan et al. [11] proposed a framework that employs data mining techniques to perform statistical analysis, email classification & clustering, author identification and social network analysis. The researchers evaluated their implementation by using the Enron e-mail corpus made available by MIT [8]. Their proposed framework is implemented in Java.

Another study of privacy of email investigation can be found in [2]. The authors introduced a cryptographic scheme that allows encrypting entire email boxes before handing them over for investigation. The key feature is that the investigator can non-interactively run keyword searches on the encrypted data and decrypt those emails for which a configurable number of matches occurred. They implemented a standard forensics framework as a plug-in.

fridaya@uef.fi

3 Tools, Technologies and Methodologies

This section describes the hardware and software tools that we have used in this research as well as our experiment process.

A. *Software Tools*

- *Aid4Mail v3.8* [1]
- *Paraben's EMX 8.6.5277* [15]
- *MailXamine 4* [12]
- *eMailTrackerPro 10* [7]
- Microsoft windows 10 Ultimate 64 bits

B. *Hardware Specifications*

A laptop with Intel core 2Duo CPU, 2.5 GHz, 4 GB RAM, and 320 GB HDD (2 partitions). First, we uninstalled the operating systems from the laptop and installed Windows 10 fresh. This process ensures the sound forensics process. Then we installed one email forensics tool at a time, imported the email files, performed analysis and then proceeded with the next tool.

C. *The Experimental Process*

As we mentioned above, we used Enron bulk email dataset file that is maintained by MIT. The dataset contains hundreds of thousands of emails. Processing all those emails are very time consuming and is beyond the scope and purpose of this study. In addition, the email forensics tools under this study impose limitation on the number of emails they can process per input folder. Therefore, we selected three folders from the email dataset. The selected folders contain all types of emails including sent emails, received emails, deleted emails, etc. The selected email folders is sufficient to demonstrate the behavior of the email forensics tools. The format of email dataset is called *maildir* (see Fig. 1). In some circumstances, we had to convert this format to other formats, so that a specific tool can process or analyze the filter before converting.

4 Analysis of the Results

The experiments involved the following steps:

a. Installed one of the email forensics tool on the machine
b. Launched the tool and created a case
c. Imported email files from the selected dataset folders
d. Ran keyword search on email files and within an email content
e. Generated report
f. Analyzed report

In the next subsections, we describe the behavior of each of the selected email forensics tools when we applied the above process.

fridaya@uef.fi

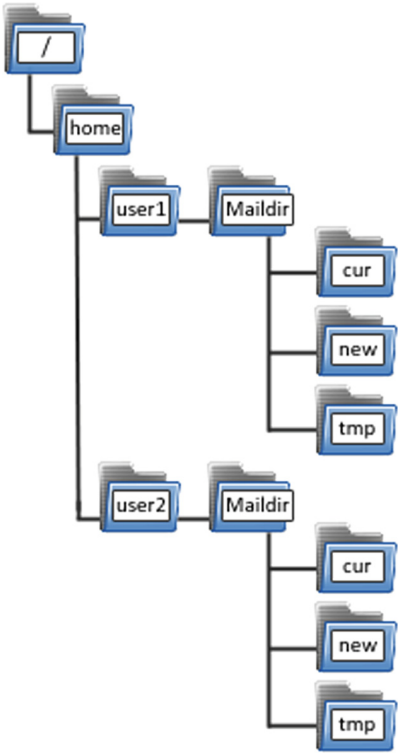


Fig. 1. Maildir internal structure

A. *Add4Mail*

We observed the following properties for *Aid4Mail*.

- It removes email-subjects from all reports and conversion process.
- Supports over 40 email formats and mail client programs, as well as many popular webmail services and remote accounts through IMAP.
- Can work on the *trash* folder.
- Capable of processing large email files
- Is able to search the computer for email file, folder name or identity name (this option not available when using *maildir* format).
- Allows the investigator to select a folder and then search subfolders for email files.
- Converts an email details to a file and the converted files treated as reports
- Supports search and filter options before the convert process.
- Can search all supported mail formats
- Can filter the emails based on text, time, date, keywords, logical operators, and regular expressions

fridaya@uef.fi

We imported the email files to *Aid4mail* and ran the tool. The generated report included the following information:

- The name of domain that email are sent or received from
- Timeline of email exchange of the three folders (time and date)
- Extracted picture
- Extracted PDF file
- Email addresses that is typed in the body of email messages
- Exported email metadata to three format namely CSV, TSV, and XML
- Converted email file to CSV and XML formats

The report result for domain name, day, and month are shown in Fig. 2a–c, respectively.

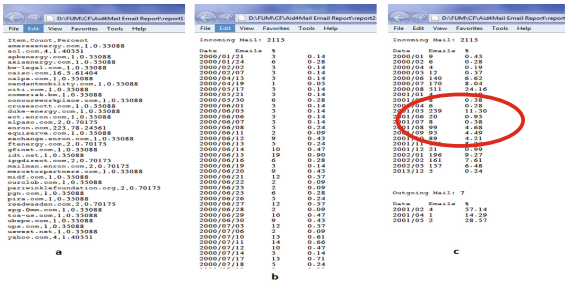


Fig. 2. Report generated by *Add4Mail*

One of the useful function of *Aid4Mail* is that we can use script or import new scripts into the software and use this script for filtering report. Some of these scripts are export only emails from Gmail address, mails with JPG/MOV/AVI files, process deleted emails, skip duplicates on export, etc. We used these filters for searching emails that was sent from 2000/04/14 until 2000/04/20. Figure 3 shows the filters we used and Fig. 4 shows the results.

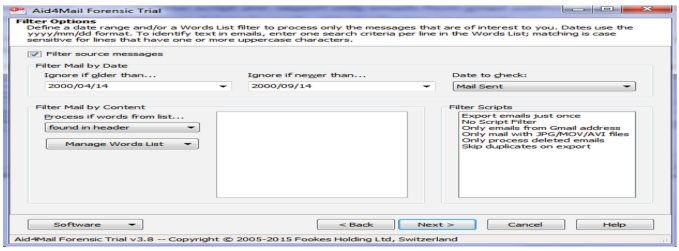


Fig. 3. Filtered page for emails between 4/14/2000 and 4/20/2000

fridaya@uef.fi

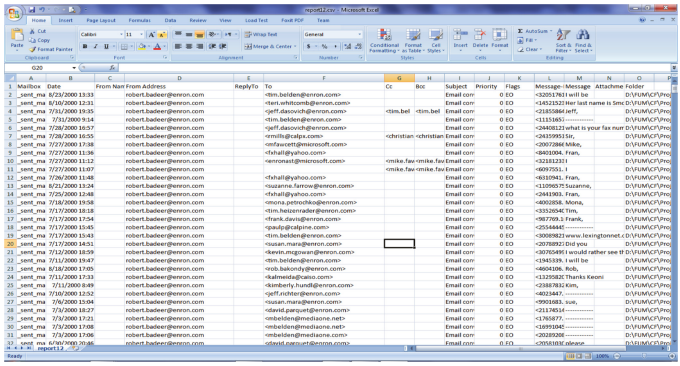


Fig. 4. Filtering email sent between 4/1/4/2000 and 4/19/2000

Our experiment shows that *Add4Mail* has some limitations. First, the Graphical User Interface (GUI) of the tool is very simple and offers no flexibility to the investigator. Second, we need other tools like web browsers; PDF readers or MS Excel for analyzing and searching the results of the report. This is because these applications are not integrated into the *Aid4Mail*.

B. *Paraben Email Examiner (EMX)*

Based on our research experiment, *Paraben*

- Can search all email files or whole database folders
- Has comprehensive keyword search features, bookmarking, advanced Boolean searching, and searching within attachments
- Supports for various languages including Unicode
- Can examine email headers and bodies, provides information based on the search filter (including contents from attachments)
- Accepts *maildir* format and several other formats
- Can show all the email files of a selected folder in a separate window
- Shows the results of analysis of the selected emails within the *Paraben*
- Can generate a report of the search results
- Can convert *maildir* email format to other formats like EML, EMX, MHT, PST. The conversion process supports filtering (based on text, hex search, locate code page, etc.) and the converted file represents a report. In addition, it has the option to define a scope for converting the email: subject, body, headers, etc.
- Allows selecting the information we want to be included in the reports. Reports in the HTML Evidence format do not support choosing data included in the report and generates a defined report. The output can be filtered (search terms, dates, or other criteria) before exporting. This feature helps the investigator in preserving privacy.

Since *Paraben* accepts *maildir* format we did not have to convert emails to a different format. Figure 5 shows the result of report generation. As can be seen from the figure, we can see the email data and the header.

fridaya@uef.fi

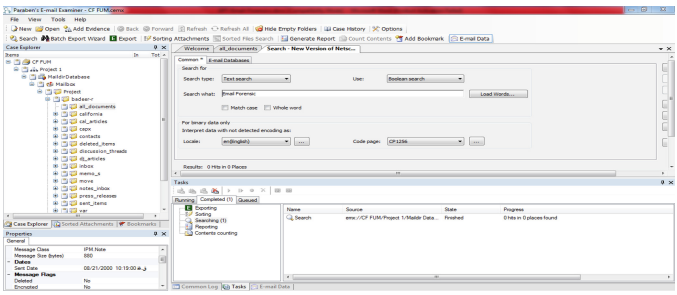


Fig. 5. Search option in Paraben for emails data

In addition, we used several keyword search and searched for the word “ENRON” in the three email folders we used in this research. The result is shown in Fig. 6.

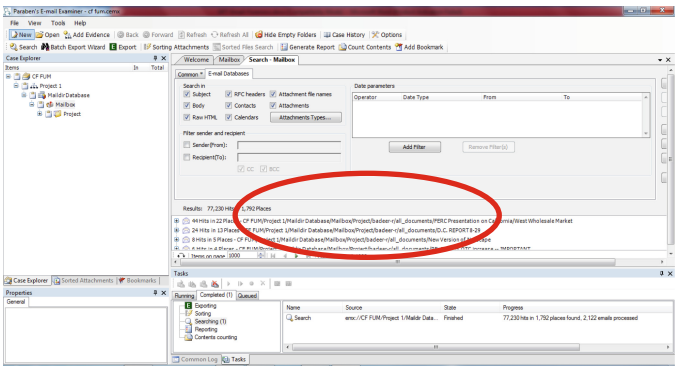


Fig. 6. Result of search for word “ENRON” in Paraben

We performed another search for the word “weekend”. The HTML report for this search is shown in Fig. 7. We can filter the search to show a number of character before and after the word.

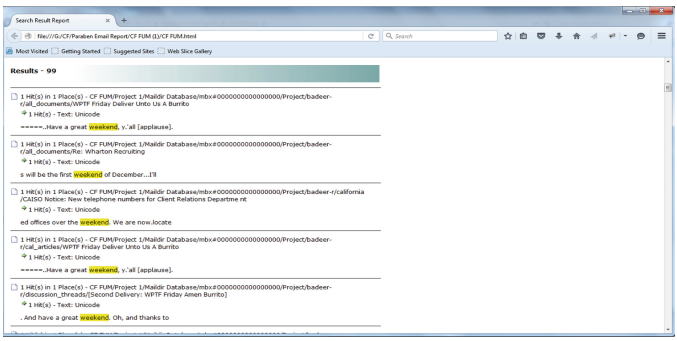


Fig. 7. HTML report generated from search result

fridaya@uef.fi

Paraben can also generate report in various formats including HTML Investigative, Simple Text, CSV Text, HTML Evidence, and HTML email message. Figure 8 show a CSV report created from email database folder.

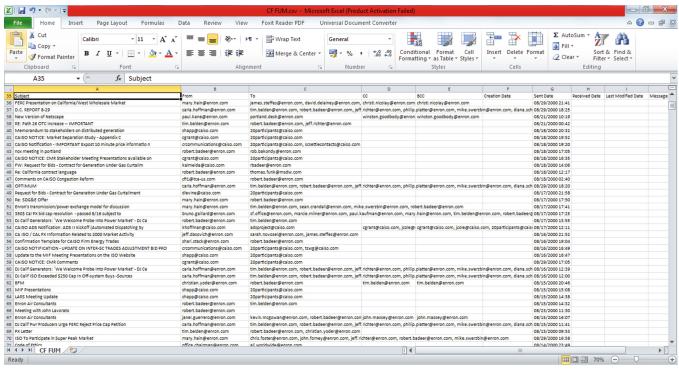


Fig. 8. Paraben CSV report generated From Enron emails

C. *eMailTrackerPr*

The results of working with this tool showed that, *eMailTrackerPro*

- Is different from the other three email forensic tools that we used, in that there is no option for importing emails file, folder or database
- Requires that we insert email header manually, connect this software to your email account, or suspect's email account
- Allows us to connect multiple accounts to this tool simultaneously
- Requires connecting to a live email account, doing keyword filtering search and generating support.

To experiment with this tool, we created a Yahoo email account, i.e. http://www.FUM_CF@YAHOO.COM/ and connected this account to the tool (see Fig. 9). If we want to trace an email that was received to this account, we can select that email and

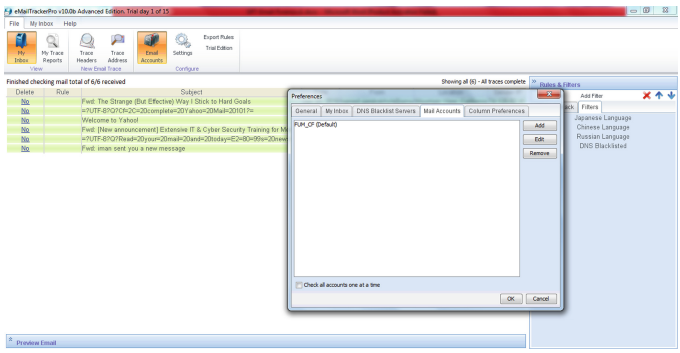


Fig. 9. Connected email account to EmailTrackerPro

fridaya@uef.fi

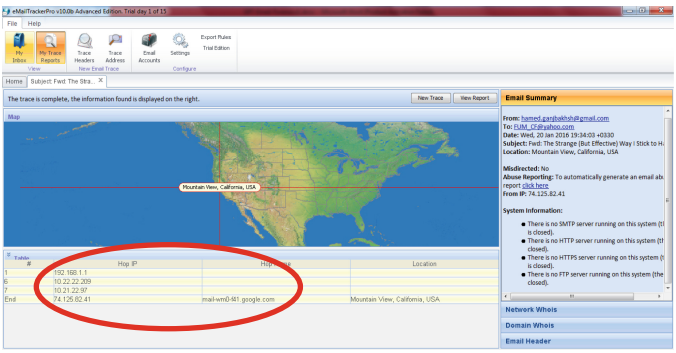


Fig. 10. Trace of an email in eMailTrackerPro

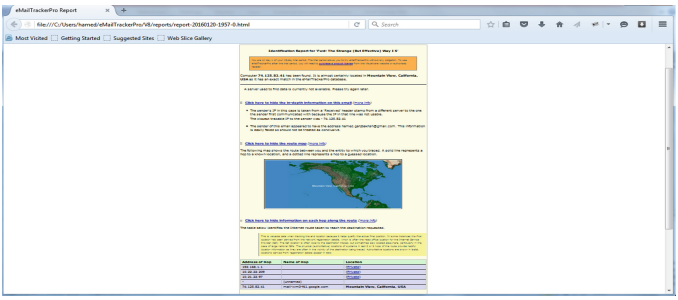


Fig. 11. HTML trace report created with eMailTrackerPro

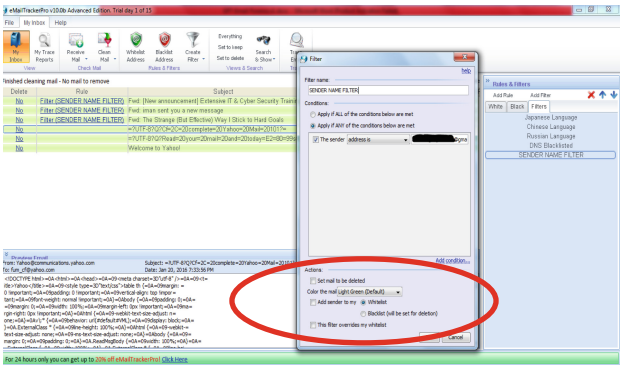


Fig. 12. Filter email sending by a specific sender

then trace it. After performing this task, we noticed that the results are shown in a separate page and include information like Whois (Network and Domain), sender and receiver name, IP address hop for this email and email header (see Fig. 10). We saved this information as a HTML report (see Fig. 11). For searching specific emails, we created specific filters (see Fig. 12).

fridaya@uef.fi

The create filter feature is capable of filtering by subject, sender's email address, to address, DNS blacklist, sender IP address, IP block, etc. Figure 13 shows a create filter window that we created. The filtering is to choose emails with sender name as http://www.Yahoo@communications.yahoo.com/

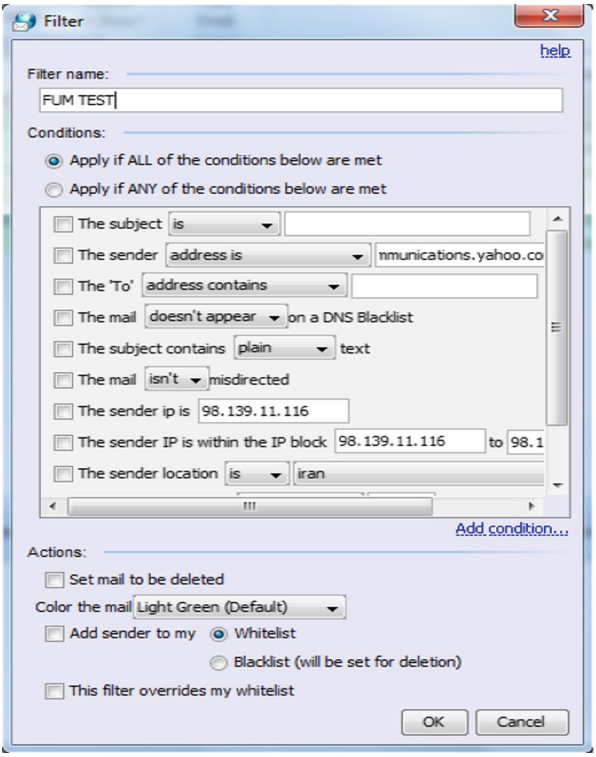


Fig. 13. Filtering email with name "FUM TEST"

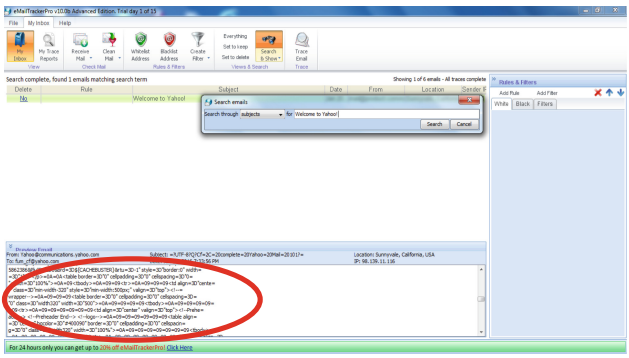


Fig. 14. Search result for email subject

fridaya@uef.fi

Another feature of this tool is that it allows the investigator to automatically send emails to blacklist or to whitelist. One downside of this tool is that it does not have many search options and investigator can only search through subject, From address, Send address, and sender IPs (see Fig. 14).

D. *MailXaminer*

The last email forensics tool that we experimented with is *MailXaminer* that:

- Supports most email file formats (36 file type, e.g. *.ost, *.pst *.olm, EML/EMLX, *maildir*, MBOX, *.mbx, *.edb, etc.)
- Facilitates both online and offline forensics (see Fig. 15).
- Allows previewing the details of all emails including mail body, HEX format of the email, properties, message header, MIME, email Hop, HTML format, RTF and attachments
- Implements keyword search filtering and analysis as well as allowing us to insert a CSV file containing keyword search.

To be consistent, we used the offline feature of this tool as well. Similar to what we did with the previous tools, we imported the emails file into this tool and began the experiment. Figure 16 shows an example of message preview in *MailXaminer*.

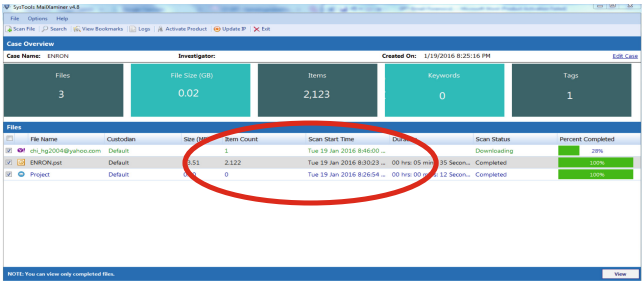


Fig. 15. Offline and online email support in *MailXaminer*

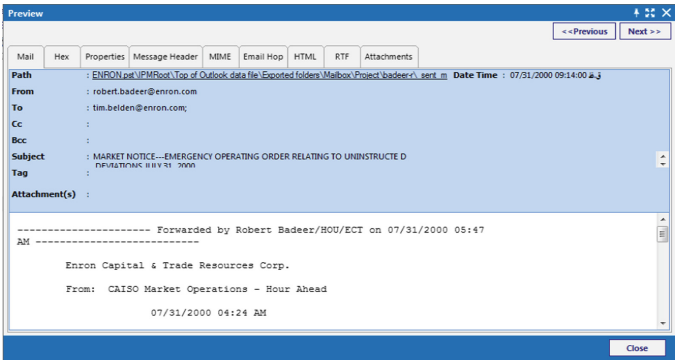


Fig. 16. Email preview on *MailXaminer*

fridaya@uef.fi

The search feature works based on a name, keyword, and date search. *MailXaminer* allows searching of emails containing specific names or emails that have been exchanged between the suspect and the other parties. Figure 17 show a search that we ran on our dataset for word “Enron”.

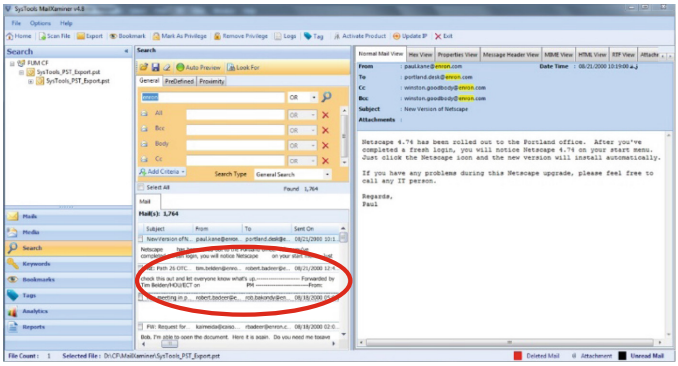


Fig. 17. Search for word “enron” in our dataset

A timeline can be created and the email exchanges within that particular date range will help in narrowing the range of emails that need to be examined. Figure 18 shows the result of a search we performed for the name “bob” and are in the timeline between 1/21/2000 and 9/22/2000.

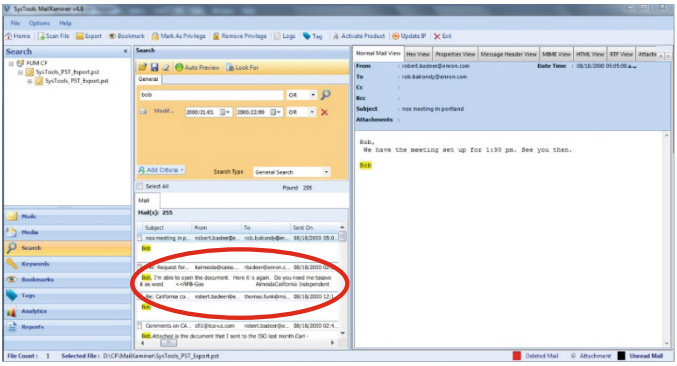


Fig. 18. Search for emails that have word “bob” and has this timeline 1/21/2000 and 9/22/2000

This tool has a Predefined tab that can search emails based on the Regular Expressions search algorithm. This search option helps the investigator to detect email message patterns with the use of category and sub category searches such as phone numbers, URLs, addresses, postal code, and country. The bookmarks section contains a record of case related evidences bookmarked which can further be exported, sent for

fridaya@uef.fi

review, deleted, or comments can be added to it. For example, Fig. 19 shows emails related to CAISO Corporation and we bookmarked Emails from that domain for further examination.

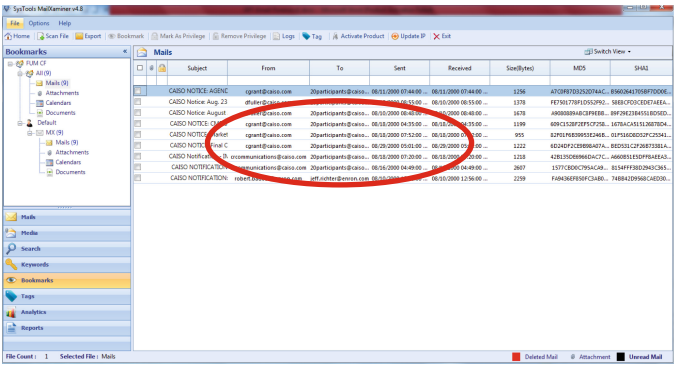


Fig. 19. Bookmarked emails under investigation

During the investigation, we may find some important emails; we can export those emails to 10 different format for analyzing by other tools or create report. The formats that are supported include Concordance, CSV, EML, HTML, HTML Reporter, MSG, PDF, PRINT, PST, TIFF. In this option, there is no filter or search option and the entire email file are exported as a report. Figure 20 shows the HTML report from emails that bookmarked in the previous step.

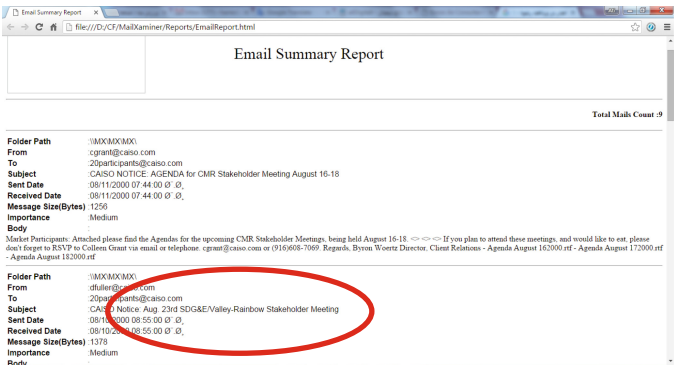


Fig. 20. HTML report from emails that investigator bookmarked

fridaya@uef.fi

5 Conclusion

In this research, we experimentally examined the capabilities of four email forensic tools, namely, *Add4Mail, Paraben Email Examiner (EMX), eMailTrackerPro,* and *MailXaminer* We used the offline version of all these tools and imported part of the Enron email dataset file for detailed analysis. We evaluated the keyword search, report generation, and other features of these tools. *Aid4Mail* can perform limited search and can create reports based on those searches but cannot analyze the generated email reports. This is because we need other applications such as PDF reader, MS Word, CVS, etc. for analyzing the results of any email search. Another limitation of this tool is that it has static and inflexible GUI. There is no dedicated search or report generation option in the GUI.

Paraben on the other hand, has comprehensive keyword search features, bookmarking, advanced Boolean searching, and searching within attachments. The search results can be analyzed within the tool and can generate reports based on keyword search results. It also has several search filters that help investigator to find items of interests from the emails. The tool can also examine email headers and bodies, provide information based on the search filter including contents from attachments. The tool has a very helpful and easy to use GUI.

The third tool, *eMailTrackerPro*, is different from the other tools in that it only supports live investigation. This tool does not support importing email files and folder so we must connect the tool to the email accounts or import header of an email manually to analyze that email. The only offline feature available in this tool is importing email headers manually to the tool and trace the email path. It is a powerful tool for searching in the emails file online and has the ability to create report from search results.

The forth tool, *MailXaminer* supports both online and offline forensics. You can import emails file into this tool and then perform keyboard search. The search feature works rapidly on the name search, keyword search and date search. Investigators can search emails, searching of emails containing specific names or emails that have been exchanged between the suspect and receivers. It supports various report generation formats.

In conclusion, some tools have better performance in online analysis and some other ones perform better offline. Overall, since each tool has some feature that is better than others are, selection of a tool depends on the needs of the investigators and the circumstances of the case under investigation.

6 Future Work

This research can be extended in several ways. First, additional tools and methodologies can be used for analyzing the emails. Second, repeat the same experiment but with proprietary version of the tools to examine additional strengths of the email forensics tools.

fridaya@uef.fi

Acknowledgment. The MCCB funding for professional development supported this research. The author would like to thank my department head, Dr. Ash Mady and the dean of the college, Dr. Dona Mayo for their support.

References

1. Aid4Mail: Email Forensic. http://www.aid4mail.com/
2. Armknecht, F., Dewald, A.: Digital Invest. **14**, 127–136 (2015)
3. Banday, M.T.: Techniques and tools for forensics investigation of emails. Int. J. Netw. Secur. Appl. (IJNSA) **3**(6), 227–241 (2011)
4. Chhabra, G.S., Bajwa, D.S.: Review of e-mail system, security protocols and email forensics. Int. J. Comput. Sci. Commun. Netw. **5**(3), 201–211 (2015)
5. Cohen, F.: Bulk email forensics. In: IFIP International Conference on Digital Forensics. Springer, Chapter 4, pp. 51–67 (2009)
6. Devendran, V.K., Shahriar, H., Clincy, V.: A comparative study of email forensic tools. J. Inf. Secur. **6**(2), 111–117 (2015)
7. EMailTrackerPro. http://www.emailtrackerpro.com/
8. Enron Email Dataset. https://www.cs.cmu.edu/~enron/
9. Hadjidj, R., Debbabi, M., Lounis, H., Iqbal, F., Szporer, A., Benredjem, D.: Towards an integrated e-mail forensic analysis framework. Digital Invest. **5**, 124–137 (2009)
10. Hatole, P.P., Bawiskar, S.K.: Literature review of email forensics. Imperial J. Interdiscip. Res. (IJIR). **3**(4), 1436–1439 (2017)
11. Khan, S.R., Nirkhi, S.M., Dharaskar, R.V.: E-mail data analysis for application to cyber forensic investigation using data mining. Int. J. Appl. Inf. Syst. 1–4 (2013)
12. MailXaminer. http://www.mailxaminer.com/
13. Meghanathan, N., Allam, S.R., Moore, L.A.: Tools and techniques for network forensics. Int. J. Netw. Secur. Appl. (IJNSA) **1**(1), 14–25 (2009)
14. Paglierani, J., Mabey, M., Ahn, G.-J.: Towards comprehensive and collaborative forensics on email evidence. In: 9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing, pp. 11–20 (2013)
15. Paraben (Network) E-mail Examiner. http://www.paraben.com/email-examiner.html
16. Stolfo, S.J., Hershkop, S.: Email mining toolkit supporting law enforcement forensic analyses. In: National Conference on Digital Government Research, Atlanta, Georgia, USA (2005)

fridaya@uef.fi

Random Sampling on an High-Dimensional Sphere for Solving SVP

Ling Qin[1], Xue Zhang[2], Xiaoyun Wang[2], Zhongxiang Zheng[3,4(✉)], and Wenfeng Qi[1]

[1] National Digital Switching System Engineering & Technological Research Center, Zhengzhou, People's Republic of China
ling_qin18@163.com
[2] Institute for Advanced Study, Tsinghua University, Beijing, People's Republic of China
[3] Department of Computer Science and Technology, Tsinghua University, Beijing, People's Republic of China
zhengzx13@mails.tsinghua.edu.cn
[4] Science and Technology on Communication Security Laboratory, Chengdu, People's Republic of China

Abstract. In this paper, we present a random sampling algorithm to solve the shortest vector problem (SVP). The idea of our algorithm comes from the Gaussian hypothesis, which is one of the most important and commonly used hypotheses in solving SVP. The main idea of our method is to generate the points which approximately satisfy Gaussian hypothesis of a given lattice. To find the shortest vector, we set some rules to make the result of this sampling algorithm get shorter progressively. We evaluate our algorithm by statistical and geometrical tools and give a basic analysis of the efficiency of algorithm. According to our analysis, parameters can be optimized to achieve high efficiency. Practical experiments have been conducted to validate our result where we solve SVP instances with dimension from 80 to 100.

Keywords: Computational geometry · Lattice · SVP · Sampling reduction

1 Introduction

Let R^m be the m-dimensional Euclidean space. Lattice $\mathcal{L}$ is a discrete additive subgroup of R^m which is the set of all integral combinations of n linearly independent vectors $\mathbf{b}_1, ..., \mathbf{b}_n$ in $R^m (m \geq n)$. The integer n is the dimension of $\mathcal{L}$ and the matrix $\mathbf{B} = (\mathbf{b}_1, ..., \mathbf{b}_n)$ is a basis of lattice $\mathcal{L}$. Computational

Supported by Foundation of Science and Technology on Information Assurance Laboratory (No. 61421120103162112008).

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 529–549, 2019.
https://doi.org/10.1007/978-3-030-22868-2_39

fridaya@uef.fi

problem is a classic field of lattice study. One of the most famous computational problems in lattice is the shortest vector problem (SVP): Find a shortest nonzero lattice vector of a given lattice. SVP was proved to be NP-hard under randomized reduction by Ajtai [1]. Due to its hardness, SVP is widely used in cryptography and other mathematical applications. Thus, algorithms for solving SVP have attracted intensive attention of scholars. Algorithms for solving SVP can be divided into two types generally:

1. Exact algorithms: Output the shortest or nearly shortest vector of a given lattice. Algorithms in this type can be classified by their space consumption. One is space-efficient enumeration algorithm [2–8]. The other is exponential-space algorithm such as sieve algorithm and its variants [9–12]. The latter have a better asymptotic running time while the former perform better in practice.
2. Approximation algorithms: Approximating the shortest lattice vector within an approximation factor c means to find a nonzero lattice vector with at most c-times the minimal possible length. One remarkable work is the Lenstra-Lenstra-Lovasz (LLL) [13] algorithm which is the first practical approximation algorithm to solve SVP. This algorithm outputs shortest lattice vector within an approximation factor $((1+\varepsilon)\sqrt{\frac{4}{3}})^{(n-1)/2}$ where ε is a positive constant. Over the 30 years after the LLL algorithm proposed, scholars have significantly improved the approximation algorithms by using some new reduction and enumeration techniques [4,14].

In a high dimensional basis, the cost of exact algorithm becomes expensive. Thus only approximation algorithm can be used to solve SVP. New techniques are proposed to combine the advantages of the two where approximation algorithms use exact algorithms as subroutines, and exact algorithms use approximation ones as preprocessing. The best approximation algorithm known in practice for high dimensional SVP is Block Korkin-Zolotarev (BKZ) algorithm [5] proposed by Schnorr and Euchner in 1994. BKZ algorithm combines the original LLL algorithm and enumeration algorithm. Although there is no guaranteed time bound for BKZ algorithm, it works better than the LLL algorithm in practice. In theory, the best approximation algorithm is Gama-Nguyen's reduction [15]. Another important idea for solving SVP is random sampling proposed by Schnorr. Sampling algorithm tries to generate points in a lattice which are short. More precisely, sampling algorithm estimate the length of vectors by observing their coordinates under orthogonal basis. If each dimension of these coordinates is small (around zero), they believe the vector is short with high probability. In 2013, Fukase and Kashiwabara proposed a variant of sampling algorithm [16] with two main changes. First, they choose the coordinates under orthogonal basis more flexible which means they can sample some points which can not be sampled by Schnoor's algorithm. Second, they observe the sum of the Gram-Schmidt orthogonalized vectors can be used as a measure to evaluate results based on which they give a new reduction algorithm to make the basis have smaller sum of the Gram-Schmidt orthogonalized vectors. Both of the two algorithms perform

fridaya@uef.fi

well in practical. One of the biggest advantages of sampling algorithms compared with other SVP solving algorithms is that it can work in parallel efficiently.

The main purpose of this paper is to propose a new sampling reduction algorithm based on the Gaussian heuristic. Our design is motivated by the random sampling and its other variants. In order to make the Gaussian heuristic close to practical case, we need to constantly optimize the basis by the power of reduction algorithm and our sphere sampling. The output of reduction algorithm is often a short basis which conforms to the GSA in some degree. To find some much better basis, we make the following two improvements. In the first aspect, we try to get a better basis by 'insertion reduction' which is widely used in reduction and sampling algorithms. Sampling algorithm outputs a specific vector and inserts this specific vector into the basis. The reduction process gets a set of vectors (may not a basis) as an input then outputs a basis by the reduction algorithm (BKZ or LLL). If the insertion vector is positive to the reduction algorithm, insertion reduction will feedback a more ideal basis (be close to the Gaussian heuristic). In the second aspect, we try to find a method to improve the efficiency of our algorithm. Another technique used in our algorithm is the sparse representation of shortest vectors with respect to the Gram Schmidt basis. This technique can be dated from Schnorr's random sampling and was expanded independently by Ding, Zhu, Wang in a genetic algorithm [17]. After that, the idea of sparse representation was combined with many other SVP algorithms and got many good results, such as orthogonalized lattice enumeration [8] proposed by Zheng, Wang, Xu et al. (an algorithm combines the sparse representation with enumeration) and restricting reduction algorithm [16] proposed by Fukase and Kashiwabara (a variant of sampling algorithm). Different from the former results, the main idea of our algorithm focuses on generating lattice vectors with norm shorter than a given value, but not the form the representations of short lattice vectors. Two benefits can be brought by the new idea. First, the range of our sampling algorithm can be calculated theoretically based on the lattice basis while it is given experimentally in the former algorithms. Second, the efficiency of our sampling can be adjusted by the choice of parameters. And we can find a set of appropriate parameters according to practical need based on our theoretical explanations.

Our Contributions. Our main contribution is to give a new sampling algorithm which tries to find the short vector progressively by adjusting the sampling parameter. The ideas of random sampling and its variants try to sample vectors in some subsets of lattice. Vectors in these subset have representations which are likely to be short. And the choices of sampling sets are based on the experiments. By utilizing Gaussian heuristic as well as sparse representation techniques into sampling algorithms, we propose a new sampling method with theoretical analysis. According to the Gaussian Heuristic, short vectors are likely to uniformly distribute on a sphere, this indicates that we can search for short vectors by sampling from those close to the chosen sphere. The hardness of generating a lattice near the sphere often relates to the basis. In other words, a 'good' basis with relatively short base vectors seems to fit better with Geometric Series Assumption. And sampling from a 'good' basis is likely to output vectors that are close

fridaya@uef.fi

to sample sphere. As previously mentioned, insertion reduction always outputs a better basis when given a good insertion vector. Insertion reduction calls BKZ reduction as a subroutine. Two observations can be obtained as follow: (1) In general, finding a good enough vector is easier (cost fewer time) than calling a BKZ reduction. (2) The different insertion vectors used in the insertion reduction usually give different output bases. According to these two facts, we need not only good vectors but also the vectors which can be effectively used in the insertion reduction to output a good basis. Our method reaches a balance between the quality of basis and efficiency of algorithm. In brief, we give up some good vectors and only vectors which likely reduce the basis better will participate in insertion reduction process. So the key of this part is how to choose the parameters to generate the vectors we need. We find there is a relationship between the number of these vectors and radius. On the one hand, the smaller radius we choose the shorter vector we could get. On the other hand, smaller radius make sample algorithm generate non-zero points with a much smaller probability. So a trade off should be considered between sample radius and the number of distinct vectors which we can generate. In our paper, we observe this phenomenon by experiment and give a description in theory.

2 Preliminaries

2.1 Basic Notation

The bold fonts in our paper denote vectors while the capital letters denote the matrices. Row representations of matrices are used here. For example, $\mathbf{B} = (\mathbf{b}_1, ..., \mathbf{b}_n)$ is a matrix, all $\mathbf{b}_i$ are row vectors. For a real number m, $\|m\|$ means the absolute value of m while $\lceil m \rfloor$ is an integer which is the closest integer around the real number m. For a vector $\mathbf{v} = (v_1, ..., v_n)$, $\|\mathbf{v}\|$ is the norm of vector. In this paper, norm means Euclidean norm. Rounding a vector $\mathbf{v}$ means rounding every element of $\mathbf{v}$, which says $\lceil \mathbf{v} \rfloor = (\lceil v_1 \rfloor, ..., \lceil v_n \rfloor)$. We denote the n-dim Euclidean ball of radius R as $Ball_n(R)$. Its volume $V_n(R) = R^n \cdot \frac{\pi^{n/2}}{\Gamma(n/2+1)}$. The n-dim sphere of radius R is denoted by $S(R)$. Let R^m be m-dim linear space. We use $\mathcal{L}$ to denote an n-dim lattice in R^m. Another important geometry we used in our paper is hyperellipsoid. If the hyperellipsoid has function $\frac{x_1^2}{a_1^2} + ... + \frac{x_n^2}{a_n^2} \leq 1$, the volume $Ve_n = \prod_{i=1}^{n} a_i \cdot \frac{\pi^{n/2}}{\Gamma(n/2+1)}$. Some distributions are used in our paper. The sign D denote normal Gaussian distribution, and U denote uniform distribution. we use $\leftarrow$ to denote a sampling process.

2.2 Lattice

Lattice is a discrete additive subgroup of some R^m. Given a set of n linearly independent vectors $\mathbf{B} = [\mathbf{b}_1, ..., \mathbf{b}_n] \in R^{n \times m}$, lattice spanned by $\mathbf{B}$ is the set $\mathcal{L}(\mathbf{B}) = \{\mathbf{B}(\mathbf{x}) | \mathbf{x} \in \mathbb{Z}^n\}$ which is an integral combination of $\mathbf{b}_i$. We say $\mathbf{B}$ is a basis of $\mathcal{L}$. The integer n is the dimension of $\mathcal{L}$. Obviously, $n \leq m$ and when

fridaya@uef.fi

$n = m$, $\mathcal{L}$ is called full-dimension of lattice. $\mathcal{L}$ could be represented by different bases, but has a fixed $vol(\mathcal{L})$ or $det(\mathcal{L})$. Actually, $vol(\mathcal{L})$ and $det(\mathcal{L})$ are two observations but the same quantity from geometry and algebra. Parallelepiped of a lattice is the section $\mathcal{P}(\mathbf{B}) = \{\sum_{i=1}^{n} x_i \mathbf{b}_i \ , 0 \leq x_i < 1\}$. $vol(\mathcal{L})$ can be viewed as volume of parallelepiped. It is easy to see that $vol(\mathcal{L}) = det(\mathcal{L}) = \sqrt{\mathbf{B}\mathbf{B}^{\mathcal{T}}}$. Define the Hermite factor as $\|b_1\|/vol((\mathcal{L}))$.

2.3 Shortest Vector

A non-zero vector with the smallest Euclidean norm in a lattice $\mathcal{L}$ is called the shortest vector of $\mathcal{L}$. The length of the shortest is called the first minimum and written as $\lambda_1(\mathcal{L})$. So we have $\lambda_1(\mathcal{L}) = min_{\mathbf{v} \in \mathcal{L}, \mathbf{v} \neq \{\prime\}} \|\mathbf{v}\|$. Given a lattice $\mathcal{L}$, the Shortest Vector Problem (SVP) is to find a shortest vector. An estimation of the length of the shortest vector is given by Minkowski [7] where $\lambda_1(\mathcal{L}) \approx GH(\mathcal{L}) = vol(\mathcal{L})^{1/n} \cdot V_n(1)^{-1/n}$. Another bound is given by Minkowski. His theorem shows that $\lambda_1(\mathcal{L}) \leq 2GH(\mathcal{L})$ for any lattice $\mathcal{L}$. SVP challenge [21] of Darmstadt calls for finding vectors with norms less than $1.05GH(\mathcal{L})$.

2.4 Orthogonalization

Orthogonalization is an important tool when discussing the hard problem in lattice. An $n \times m$ basis $\mathbf{B} = (\mathbf{b}_1, ..., \mathbf{b}_n)$ can be written uniquely as $B = \mu \cdot D \cdot Q$ where $\mu = (\mu_{i,j})$ is $n \times n$ lower-triangular with unit diagonal, D is $n \times n$ positive diagonal, and Q is $n \times m$ with orthogonal row vectors. $\mathbf{B}^* = DQ = (\mathbf{b}_1^*, ..., \mathbf{b}_n^*)$ is the Gram-Schmidt orthogonalization of the basis. The sign π_i denotes the orthogonal projection over $(\mathbf{b}_1, ..., \mathbf{b}_{i-1})^{\perp}$.

2.5 Heuristic and Assumption

In this part we give some heuristics and assumptions which are used in our paper. Most of these heuristics have not been proven or have counter-example in some cases. But generally, we regard the heuristics hold under random basis.

Gaussian Heuristic. The Gaussian heuristic is used to estimate the number of vectors in a lattice. It assumes that the number of points in a set is related to its volume. Given a lattice $(\mathcal{L})$ and a measurable subset $S \subseteq R^m$, the number of points in $\mathcal{L} \cap \mathcal{S}$ is approximately $vol(S)/vol(\mathcal{L})$.

Heuristic 2. The distribution of the coordinates of the shortest vector v, when written in normalized Gram-Schmidt basis $(\mathbf{b}_1^*/\|\mathbf{b}_1^*\|, ..., \mathbf{b}_n^*/\|\mathbf{b}_n^*\|)$ of the input basis, looks like those of a uniformly distributed vector of norm $\|\mathbf{v}\|$.

Heuristic 3. The distribution of the normalized Gram-Schmidt orthogonalization $(\mathbf{b}_1^*/\|\mathbf{b}_1^*\|, ..., \mathbf{b}_n^*/\|\mathbf{b}_n^*\|)$ of a random reduced basis $(\mathbf{b}_1, ..., \mathbf{b}_n)$ looks like that of uniformly distributed orthogonal matrix.

fridaya@uef.fi

The Geometric Series Assumption (GSA). The output basis of reduction algorithm has series below: Let $\mathbf{b}_1, ..., \mathbf{b}_n$ be the output of reduction algorithm with its orthogonal basis denoted as $\mathbf{b}_1^*, ..., \mathbf{b}_n^*$. We have $\frac{\|\mathbf{b}_i\|^2}{\|\mathbf{b}_{i+1}\|^2} \approx q$, $q > 1$ for $i = 1, ..., n-1$. Here q is a constant which relates to the algorithm. Consider the Hermite factor $\|\mathbf{b}_1\|/vol(\mathcal{L})$, we have $\|\mathbf{b}_1\|/vol(\mathcal{L}) = \|\mathbf{b}_1^*\|/\prod_{i=1}^{n} \mathbf{b}_i^* = q^{(n-1)n/2}$. In other word, Hermite factor also relates to the algorithm. For example, BKZ's Hermite factor is different with LLL algorithm. For the different bases but same algorithm, the Hermite factor are usually the same.

2.6 Integer Representation

A lattice vector $\mathbf{v}$ can be written as $\mathbf{v} = t_1\mathbf{b}_1 + t_2\mathbf{b}_2 + ... + t_n\mathbf{b}_n = u_1\mathbf{b}_1^* + u_2\mathbf{b}_2^* + ... + u_n\mathbf{b}_n^*$. And the vector $(\lceil u_1 \rfloor, ..., \lceil u_n \rfloor)$ is the integer representation of $\mathbf{v}$. For every lattice vector $\mathbf{v}$, there is one and only one integer representation. When given a lattice vector, we can get the only integer representation within polynomial steps. And we can also recover the lattice vector according to its integer representation within polynomial steps.

2.7 Random Sampling Reduction

The earliest sampling technique was proposed by Schnorr named Random Sampling Reduction. The random sampling reduction algorithm includes two parts: Random Sampling and Insertion Reduction. Given some $u < n$, the random sampling generates lattice vectors which satisfy:

$$\|v_j\| = \begin{cases} \frac{1}{2}, \text{ for } j < n-u, \\ 1, \text{ for } n-u \leq j \leq n, \end{cases} \tag{1}$$

The number of different lattice vectors which can be generated by sampling algorithm is 2^u. The candidates of a lattice vector under lattice basis are integers so that the first $n-u$ candidates have only one value satisfying $\|v_j\| \leq \frac{1}{2}$ while the last u candidates have two possible value satisfying $\|v_j\| \leq 1$. Consider a lattice vector $\mathbf{v} = \sum_{j=1}^{n} v_j\mathbf{b}_j^*$ where $\mathbf{b}_1^*, ..., \mathbf{b}_j^*$ are orthogonal, so the norm of the vector has the relation

$$\|\mathbf{v}\|^2 = \sum_{j=1}^{n} \|v_j\mathbf{b}_j^*\|^2$$

Given $\mathbf{b}_1, ..., \mathbf{b}_n$, let every $\|v_j\mathbf{b}_j\|$ be small. $\|\mathbf{v}\|$ seems to be small with high probability. According to GSA, the norms of an orthogonal basis, output by a BKZ algorithm, are close to a geometric series where $\|\mathbf{b}_j^*\|/\|\mathbf{b}_{j+1}^*\| \approx q$ and $q > 1$ which tends to be smaller in a better basis. It is seen that $\|\mathbf{b}_j^*\|$ with small index j are much larger and contribute a lot in $\|\mathbf{v}\|$. When the basis is not good enough, the sampling algorithm hardly finds the shortest lattice vector directly. To deal with this situation, a sampling algorithm is often combined with a reduction algorithm. The goal of the sampling algorithm is to generate some vectors which can be used to reduce basis by insertion reduction.

fridaya@uef.fi

The reduction algorithm will be called when sampling algorithm choose vector $\mathbf{v}$ satisfying $\|\pi_j(\mathbf{v})\| \leq \|\mathbf{b}_j^*\|$ from all 2^u distinct lattice vectors. Since the reduction algorithm usually costs more time in high dimensional situations and a vector used to update basis can be generated easily by sampling algorithm, a reasonable thought is to filter and choose the vectors which can be efficiently used in the insertion algorithm. In general, if insertion vector holding the inequations above for a small integer $j > 0$, output of insertion algorithm are likely to be short. In a practical algorithm, a bound of the subscript j should be given. In Schnorr's paper this bound is 10. In other words, only vectors generated by sampling algorithm with $\|\pi_j(\mathbf{v})\| \leq \|\mathbf{b}_j^*\|$ for some $1 \leq j \leq 10$, will be used in insertion reduction.

The reduction part of Schnorr's algorithm just seems like most other reduction algorithms. By inputting a basis with some inserted vectors, the reduction algorithm outputs a new basis which is better than the given basis. Reduction and sampling will be called alternately until a vector with short norm is found.

Algorithm 1. Sampling Algorithm

Input:
B: a lattice basis $\mathbf{B} = [\mathbf{b}_1, ..., \mathbf{b}_n]$
u: an integer subject to $1 \leq u < n$
Output:
$\mathbf{v}$: a lattice vector satisfying Eq.(1).

compute $\mu_{i,j}$ subject to $\mu_{i,j} =< \mathbf{b}_i, \mathbf{b}_j^* > / < \mathbf{b}_j^*, \mathbf{b}_j^* >$
$V := b_n$
for $j = 1, ..., n-1$ **do**
  $\mu_j := \mu_{n,j}$
end for
for $i = n-1, ..., 1$ **do**
  select $y \in Z$ randomly subject to $|\mu_i - y| \leq \begin{cases} \frac{1}{2}, & \text{if } i < n-u \\ 1, & \text{if } i \geq n-u \end{cases}$
  $\mathbf{v} := \mathbf{v} - y\mathbf{b_i}$
  for $j = 1, ..., i-1$ **do**
    $\mu_j := \mu_j - y\mu_{i,j}$
  end for
end for

3 Sphere Sampling Reduction

In these part, we propose a new sampling algorithm which is called Sphere Sampling Reduction. The idea of sphere sampling is to use the Gaussian Heuristic to restrict the execution of sampling algorithm. According to heuristics given in Sect. 2.3, it indicates $\mathbf{v}$ should distribute uniformly in a sphere whose radius is very close to the norm of the shortest vector. Different Shortest vectors $\mathbf{v}$ always have the same norm $\|\mathbf{v}\| = \lambda_1(\mathbf{B})$. So if we can estimate the norm of $\mathbf{v}$, we can search sphere with radius $\|\mathbf{v}\|$ and find the points of given lattice.

fridaya@uef.fi

But the idea may not work so well because of the following two reasons. Firstly, the estimation of the shortest vector in a given lattice is not accurate, only a upper bound can be given. Secondly, although we get a random point from sphere, how to find a lattice vector which is closest to the given point (Closest Vector Problem, CVP) is still a hard problem. Existing CVP algorithm or its variants need a basis and a point $u \in R^m$ as input and output a lattice point which is "near" u. The output is related with the given basis. A high quality basis usually means much smaller distance between u and the output lattice vector. In practice, two problems above are reflected in the sampling stage. Giving a small radius and random points on this sphere is always easy. But the output of CVP is zero vector with a very high probability. Although this idea may not work but can give us some enlightenments that efficiency of the algorithm is mainly decided by the sampling area. A sphere sampling may have a better efficiency than other sampling. The only problem is how to make the sampling output enough number of none-zero vectors. To solve this problem, we have to focus on two aspects, appropriate radius and high quality basis.

In our paper, we use progressive method to solve these problems. Instead of giving a small radius or an ideal basis, we combine sampling and reduction algorithm to reach them progressively.

(1) In the first step we use the BKZ with a small blocksize to reduce the basis which makes the basis better.
(2) We input our basis and determine the parameters under this basis and start sphere sampling algorithm.
(3) Insert vector to the basis and call BKZ algorithm.
(4) Repeat step (2) and (3) times until a shortest vector is found.

Our sampling algorithm needs a new parameter r which named sampling radius as a measure to control sampling area and norms' expectation of output vectors. The parameter is related to current basis and will be renewed when the basis is updated. To be more specific, our algorithm tries to sample the points of lattice which are near a sphere with radius which can be calculated by a given basis. Like random sampling reduction algorithm, we choose some points that satisfy $\|\pi_j(\mathbf{v})\| \leq \|\mathbf{b}_j^*\|$ to renew our basis. When we get some appropriate points, the reduction stage will start, where a set of vectors are input into BKZ algorithm and then a basis is output. This process actually is a variant of BKZ. BKZ terminates when the combination of vectors in every single block can not be used to reduce the basis. But there are still many linear combinations of vectors contained in different blocks which can be used to renew basis. Our sampling can find these vectors and use them in reduction algorithm to make the basis better. Then we can decrease our sampling radius based on the better basis. These steps can be repeated times, and the output basis will become shorter and more conform to heuristics. So we can get a short enough lattice point which has the norm very close to the shortest vector. The algorithm has following subroutines: GenSphere, GauSampling, GenRepv, FindVec.

fridaya@uef.fi

The details of these subroutines will be shown as below:

Algorithm 2. Sphere Sampling Reduction Algorithm

Input: **B**:a lattice basis $\mathbf{B} = [\mathbf{b}_1, ..., \mathbf{b}_n]$.
insertsign:a sign to record where to update.
Blocksize:the blocksize of BKZ
bound:the bound of subscript which is used to control insertion vectors, $bound \leq n$.
Output: **v**:a short lattice vector satisfying $\|\mathbf{v}\| \leq 1.05GH(\mathcal{L})$, $1.05GH(\mathcal{L})$ is an estimation of the shortest vector.

```
while ||b_1^*|| ≥ 1.05GH(L) do
    calculate the radius r of the ball
    insertsign = 0
    while insertsign = 0 do
        call the GenSphere(r) to generate a random point s on a sphere with radius r.
        get representation repv from s according to GenRepv(s).
        generate lattice vector v from FindVec(repv).
        for j = bound, ..., 1 do
            if ||π_j(v)|| ≤ ||b_j^*|| then
                insertsign = j
            end if
        end for
    end while
    use BKZ(B, v, insertsign) to renew the basis.
end while
```

Random Sampling on a Sphere. Random sampling on a sphere is the major part of our sampling algorithm. This algorithm will output points distributed uniformly on a given sphere.

GenSphere is based on a very basic thought. The method was first suggested by G.W. Brown in 1956 and is still one of the most useful methods to get random points on n-sphere ($n \geq 3$). Suppose that a random point, uniformly distributed on the surface of a sphere in n-dimension, is required or, equivalently, that a random vector is required with uniform distribution in angle. If n independent Gaussian deviates $x_1, x_2, ..., x_n$ are obtained, the direction of the vector will satisfy the requirement. The GenSphere algorithm has two inputs. One is radius r of sphere which is the range of points' generation, while the other one is the dimension of space. In our algorithm the dimension is often the same as lattice's. But in practice the dimension can be much smaller than lattice's dimension. We will describe in detail the practical algorithm in Sect. 5.3.

Gaussian Sampling. The key of GenSphere is how to sample from Gaussian distribution. It is a traditional problem in math and computer science. Many methods were proposed to improve the precision and efficiency of Gaussian sampling. We use Polar method for normal deviates, which can generate a pair of independent normally distributed variables every time. It was proposed by M.E.

fridaya@uef.fi

Algorithm 3. Random sampling on a sphere (GenSphere)

Input:
r:the radius of a sphere.
n:the dimension of sphere.
Output:
s:a point on this sphere

1: **for** $i = n, ..., 1$ **do**
2: $x_i \leftarrow D$
3: $sum = sum + x_i^2$
4: **end for**
5: $\mathbf{s} = (x_1, ..., x_n)$
6: $\mathbf{s} = r/\sqrt{sum} \cdot \mathbf{s}$

Muller and G. Marsaglia in 1958 [20]. This algorithm has a loop which can give an output only when $S < 1$, and it is easy to get that the expectation number of rounds is 1.27 times, with a standard deviation 0.59. Actually, our Gaussian sampling algorithm may be not the most optimal in efficiency and can be replaced by other discrete Gaussian sampling algorithms that can achieve either better time complexity or space complexity. For example, if storage is enough, a Cumulative Distribution Tables (CDT) [18] method which stores all probabilities and output samples by table look-up may achieve a better efficiency.

Algorithm 4. Gaussian Sampling (GauSampling)

Output:
x_1, x_2: a pair of independent normally distributed variables

1: $S = 1$
2: **while** $S \geq 1$ **do**
3: $u_1 \leftarrow U, u_2 \leftarrow U$
4: $v_1 = 2u_1 - 1, v_2 = 2u_2 - 1$
5: $S = v_1^2 + v_2^2$
6: **end while**
7: **if** $S = 0$ **then**
8: $x_1 = 0, x_2 = 0$
9: **else**
10: $x_1 = v_1\sqrt{\frac{-2lnS}{S}}, x_2 = v_2\sqrt{\frac{-2lnS}{S}}$
11: **end if**

Find Near Lattice Point. In our algorithm, we need to generate some lattice vectors which are near sample vector on the sphere. Finding lattice vectors which are closest to the given vector is one of the basic hard problems in lattice (CVP). The result of CVP algorithm is influenced by quality of basis. In our algorithm, we use integer representation to solve this problem. It can be seen as a variant of

fridaya@uef.fi

Babai's round-off algorithm and has been used to solve lattice problems in former work [19]. When given a vector, we can get a lattice vector's representation by the algorithm below:

Algorithm 5. Generate Integer Representation (GenRepv)

Input:
s:a vector.
Output:
repv:a representation of lattice vector which is near the given vector.

1: $\mathbf{repv} = (\lceil \mathbf{s}_1/\|\mathbf{b}_1^*\| \rfloor, ..., \lceil \mathbf{s}_n/\|\mathbf{b}_n^*\| \rfloor)$

First, divide the elements of vector on sphere by norm of basis. Then, after rounding the coordinate calculated in the first step, we get an integer representation of a lattice vector which is close to the given vector. And an integer representation gives a round off of coordinate of lattice point. This correspondence is a bijection from lattice to integer representation. Because the reduction algorithm needs to estimate the quality of lattice vector generated in our sampling algorithm (in sphere sampling algorithm, a judgment of whether lattice vector can be used to renew basis is necessary). An algorithm which recovers the lattice vector according to its integer representation is needed. An ordinary result was shown as below. Let a basis $\mathbf{B} = (\mathbf{b}_1, ..., \mathbf{b}_n)$ to its orthogonal basis is $\mathbf{B}^* = (\mathbf{b}_1^*, ..., \mathbf{b}_n^*)$. For every lattice vector $\mathbf{v}$, we have

$$\begin{aligned} v &= t_1\mathbf{b}_1 + t_2\mathbf{b}_2 + ... + t_n\mathbf{b}_n \\ &= u_1\mathbf{b}_1^* + u_2\mathbf{b}_2^* + ... + u_n\mathbf{b}_n^* \\ &= \mathbf{b}_1^*(t_1 + \frac{<\mathbf{b}_2, \mathbf{b}_1^*>}{<\mathbf{b}_1^*, \mathbf{b}_1^*>}t_2 + ... + \frac{<\mathbf{b}_n, \mathbf{b}_1^*>}{<\mathbf{b}_1^*, \mathbf{b}_1^*>}t_n) + ... + \mathbf{b}_n^*t_n \end{aligned}$$

Thus,

$$\begin{aligned} u_1 &= t_1 + \frac{<\mathbf{b}_2, \mathbf{b}_1^*>}{<\mathbf{b}_1^*, \mathbf{b}_1^*>}t_2 + ... + \frac{<\mathbf{b}_n, \mathbf{b}_1^*>}{<\mathbf{b}_1^*, \mathbf{b}_1^*>}t_n \\ u_2 &= t_2 + \frac{<\mathbf{b}_3, \mathbf{b}_2^*>}{<\mathbf{b}_2^*, \mathbf{b}_2^*>}t_3 + ... + \frac{<\mathbf{b}_n, \mathbf{b}_2^*>}{<\mathbf{b}_2^*, \mathbf{b}_2^*>}t_n \\ &\vdots \\ u_{n-1} &= t_{n-1} + \frac{<\mathbf{b}_n, \mathbf{b}_{n-1}^*>}{<\mathbf{b}_{n-1}^*, \mathbf{b}_{n-1}^*>}t_n \\ u_n &= t_n; \end{aligned}$$

fridaya@uef.fi

And the integer representation of vector is $(repv_1, ..., repv_n) = (\lceil u_1 \rfloor, ..., \lceil u_n \rfloor)$. We note the last $\lceil u_n \rfloor = u_n = t_n = repv_n$. To get the original $(t_1, ..., t_{n-1})$, we have observation as below:

$$\begin{cases} (repv_{n-1} - 0.5) - \frac{<\mathbf{b}_n, \mathbf{b}^*_{n-1}>}{<\mathbf{b}^*_{n-1}, \mathbf{b}^*_{n-1}>} t_n \le t_{n-1} \le (repv_{n-1} + 0.5) - \frac{<\mathbf{b}_n, \mathbf{b}^*_{n-1}>}{<\mathbf{b}^*_{n-1}, \mathbf{b}^*_{n-1}>} t_n \\ \vdots \\ (repv_1 - 0.5) - (\frac{<\mathbf{b}_2, \mathbf{b}^*_1>}{<\mathbf{b}^*_1, \mathbf{b}^*_1>} t_2 + ... + \frac{<\mathbf{b}_n, \mathbf{b}^*_1>}{<\mathbf{b}^*_1, \mathbf{b}^*_1>} t_n) \le t_1 \le (repv_1 + 0.5) - (\frac{<\mathbf{b}_2, \mathbf{b}^*_1>}{<\mathbf{b}^*_1, \mathbf{b}^*_1>} t_2 + ... \\ \qquad\qquad + \frac{<\mathbf{b}_n, \mathbf{b}^*_1>}{<\mathbf{b}^*_1, \mathbf{b}^*_1>} t_n) \end{cases}$$

All t_i can be solved when given a representation. Because t_i is the only integer which satisfies the ith inequation. We have the algorithm below:

Algorithm 6. Find Lattice Vector(Findvec)

Input
repv:an integer representation.
Output
t:coordinate under the lattice basis B.
v:a lattice vector whose integer representation is $repv$.

1: $t_n = repv_n$.
2: **for** i=n-1,...,1 **do**
3: $t_i = \lceil repv_i - 0.5 - \sum_{j=i+1}^{n} \frac{<\mathbf{b}_j, \mathbf{b}^*_i>}{<\mathbf{b}^*_i, \mathbf{b}^*_i>} t_j \rceil$
4: **end for**

4 Running Time and Success Probability Analysis

The key to analyze the success probability is to find a good distribution to describe the sampling process. To estimate the success probability, a direct idea is to calculate the area of sampling section $S_{sampling}$ and goal region S_{goal} first. Then $S_{goal}/S_{sampling}$ is an estimation of success probability. In general, the area of $S_{sampling}$ is hard to calculate. A viable solution is to find some areas which can be calculated easily to replace $S_{sampling}$.

To study the behavior of algorithm, we divide this algorithm into three parts:

In the first part, we generate a vector $\mathbf{s} = (\frac{x_1}{\sqrt{x_1^2+...+x_n^2}} r, ..., \frac{x_n}{\sqrt{x_1^2+...+x_n^2}} r)$ uniformly distributed on a sphere with radius r.

In the second part, a representation of a lattice vector is given by transforming the vector $\mathbf{s}$. This transform is a two-step process. The first step divides the i_{th} element of $\mathbf{s}$ by $\|b_i^*\|$ respectively. The second step rounds off the output vector. We get an integer representation.

In the third part, vectors are recovered from integer representations. And the distribution of these recovered vectors is the goal of our calculation.

From the first part, it is obvious that $\mathbf{s}$ is a point on a sphere which conforms $x_1^2 + ... + x_n^2 = r^2$. This sphere uniform distribution is the basic of our research.

Back to the second part, first step transformer process will generate a vector on an ellipsoid which has the function $\frac{x_1^2}{r^2/\|\mathbf{b}_1^*\|^2} + ... + \frac{x_n^2}{r^2/\|\mathbf{b}_n^*\|^2} = 1$:

fridaya@uef.fi

Theorem 1. *Given a lattice basis* $\mathbf{B} = (\mathbf{b}_1, ..., \mathbf{b}_n)$*, there is a one-to-one mapping between vectors on a sphere with radius* r *and vectors on an ellipsoid which satisfies* $\frac{x_1^2}{r^2/\|\mathbf{b}_1^*\|^2} + ... + \frac{x_n^2}{r^2/\|\mathbf{b}_n^*\|^2} = 1$*. If sampling process samples vectors uniformly distributed on sphere, the images of this mapping will uniformly distribute on ellipsoid.*

Proof. The sphere with radius r and center origin has the function $x_1^2 + x_2^2 + ... + x_n^2 = r^2$. For a vector $\mathbf{v} = (v_1, ..., v_n)$ on the sphere, we have $v_1^2 + ... + v_n^2 = r^2$. Considering the vector $\mathbf{v}^* = (v_1\|\mathbf{b}_1^*\|, ..., v_n\|\mathbf{b}_n^*\|)$, it is obvious that $\mathbf{v}^*$ is a vector on ellipsoid $\frac{x_1^2}{r^2/\|\mathbf{b}_1^*\|^2} + ... + \frac{x_n^2}{r^2/\|\mathbf{b}_n^*\|^2} = 1$. We say the mapping $\varphi : \mathbf{v} \to \mathbf{v}^*$. To prove φ is a one-to-one mapping, we need to prove every vector on ellipsoid has one and only one preimage. If there are two preimage $\mathbf{v}_1, \mathbf{v}_2$ for the image $\mathbf{v}^*$, then $\mathbf{v}_1 = (v_{11}\|\mathbf{b}_1^*\|, ..., v_{1n}\|\mathbf{b}_n^*\|) = (v_{21}\|\mathbf{b}_1^*\|, ..., v_{2n}\|\mathbf{b}_n^*\|) = \mathbf{v}_2$. For every vector on sphere, there will be an image on ellipsoid. When the sampling algorithm generates vector on sphere with probability p, the image of this vector also has the same probability p.

The theorem above tells that the vectors distribute uniformly on an ellipsoid after the first transformation. This clear geometric relationship is a great help for us to discuss the success probability. After the step one, the output of step one is rounded off. And the theorem below tells us the range where the coordinates under orthogonal basis of lattice vectors locate.

Theorem 2. *The coordinates under orthogonal basis of lattice vectors generated by sphere sampling algorithm locate in* e_1 *where* e_1 *is an ellipsoid which has function* $\frac{x_1^2}{(\lceil r/\|\mathbf{b}_1^*\|\rfloor+0.5)^2} + ... + \frac{x_n^2}{(\lceil r/\|\mathbf{b}_n^*\|\rfloor+0.5)^2} = n$*.*

Proof. Define the representation of lattice in our algorithm as $\mathbf{repv} = (repv_1, ..., repv_n)$ which can be written as $\mathbf{repv} = (\lceil \frac{x_1}{\sqrt{x_1^2+...+x_n^2}} \cdot \frac{r}{\|\mathbf{b}_1^*\|} \rfloor, ..., \lceil \frac{x_n}{\sqrt{x_1^2+...+x_n^2}} \cdot \frac{r}{\|\mathbf{b}_n^*\|} \rfloor)$. According to the relationship between representation and coordinate under basis of lattice vector, we have $repv_i + 0.5 \geq \mu_i \geq repv_i - 0.5$, where $repv_i$ is ith element of representation and μ_i is the corresponding element of coordinate under orthogonal basis. For every i, $(|repv_i| + 0.5)^2/(\lceil r/\|\mathbf{b}_i^*\|\rfloor + 0.5)^2 < 1$ holds, so inequation $\frac{(|repv_1|+0.5)^2}{(\lceil r/\|\mathbf{b}_1^*\|\rfloor+0.5)^2} + ... + \frac{(|repv_n|+0.5)^2}{(\lceil r/\|\mathbf{b}_n^*\|\rfloor+0.5)^2} \leq n$ holds. This means coordinates under orthogonal basis of lattice vectors generated by sphere sampling algorithm locate in e_1.

We will analyze probability under following assumption.

Assumption 1. The coordinate of lattice vector under orthogonal basis uniformly distributes in the e_1.

Consider a lattice vector v with $v = \mu_1 b_1^* + ... + \mu_n b_n^*$. We use the directions of $b_1^*, ..., b_n^*$ as the positive directions of coordinate axis and the length 1 as the unit length to establish coordinate system named A which is a rectangular coordinate system. Because the coordinates under orthogonal basis of lattice vectors generated by sphere sampling algorithm locate in e_1, we can get the theorem below.

fridaya@uef.fi

Theorem 3. *Under the coordinate system A, the lattice vector will distribute in the ellipsoid* $\frac{y_1^2}{((\lceil r/\|\mathbf{b}_1^*\|\rfloor+0.5)\|\mathbf{b}_1^*\|)^2}+...+\frac{y_n^2}{((\lceil r/\|\mathbf{b}_n^*\|\rfloor+0.5)\|\mathbf{b}_n^*\|)^2}=1$ *uniformly.*

Proof. Under the coordinate system A, there is a map from $\mathbf{u}=(\mu_1,...,\mu_n)$ to $\mathbf{v}'=(\mu_1\|\mathbf{b}_1^*\|,...,\mu_n\|\mathbf{b}_n^*\|)$ which is a bijection. On the one hand, Theorem 3 guarantees $\frac{\mu_1^2}{(\lceil r/\|\mathbf{b}_1^*\|\rfloor+0.5)^2}+...+\frac{\mu_n^2}{(\lceil r/\|\mathbf{b}_n^*\|\rfloor+0.5)^2}\leq 1$. Obviously, $\frac{y_1^2}{((\lceil r/\|\mathbf{b}_1^*\|\rfloor+0.5)\|\mathbf{b}_1^*\|)^2}+...+\frac{y_n^2}{((\lceil r/\|\mathbf{b}_n^*\|\rfloor+0.5)\|\mathbf{b}_n^*\|)^2}\leq 1$ holds. On the other hand, if there exists some $\mathbf{v}'=(y_1,...,y_n)=$ which satisfies the relation $\frac{y_1^2}{((\lceil r/\|\mathbf{b}_1^*\|\rfloor+0.5)\|\mathbf{b}_1^*\|)^2}+...+\frac{y_n^2}{((\lceil r/\|\mathbf{b}_n^*\|\rfloor+0.5)\|\mathbf{b}_n^*\|)^2}\leq 1$, it is obviously that vector $\mathbf{u}=(\mu_1,...,\mu_n)=(y_1/\|b_1^*\|,...,y_n/\|b_n^*\|)$ satisfies the inequation $\frac{\mu_1^2}{(\lceil r/\|\mathbf{b}_1^*\|\rfloor+0.5)^2}+...+\frac{\mu_n^2}{(\lceil r/\|\mathbf{b}_n^*\|\rfloor+0.5)^2}\leq 1$. So there is a one-to-one mapping from points in the e_1 to those in e_2. According to the assumption, we have the lattice vector will distribute in the ellipsoid $\frac{y_1^2}{((\lceil r/\|\mathbf{b}_1^*\|\rfloor+0.5)\|\mathbf{b}_1^*\|)^2}+...+\frac{y_n^2}{((\lceil r/\|\mathbf{b}_n^*\|\rfloor+0.5)\|\mathbf{b}_n^*\|)^2}=1$ uniformly.

This theorem indicates that to estimate the success probability, we can calculate the area of sampling section $S_{sampling}$ and goal region S_{goal} first. Then $S_{goal}/S_{sampling}$ is an estimation of success probability.

Lemma 1. *For* $r>0$ *and* $a_i>0$, *denote* $f_1(\mathbf{x})=x_1^2+,...,+x_n^2$, $f_2(\mathbf{x})=\frac{x_1^2}{a_1^2}+...+\frac{x_n^2}{a_n^2}$, $f_3(\mathbf{x})=\frac{x_1^2}{min(a_1,r)^2}+,...,+\frac{x_n^2}{min(a_n,r)^2}$. *If* $S_1=\{\mathbf{x}|\mathbf{x}$ *satisfies* $f_1(\mathbf{x})\leq r^2$ *and* $f_2(\mathbf{x})\leq 1\}$, *while* $S_2=\{\mathbf{x}|\mathbf{x}$ *satisfies* $f_3(\mathbf{x})\leq 1\}$ *the vectors which are in* S_2 *must be in* S_1.

Proof. For some $\mathbf{v}=(v_1,...,v_n)\in S_2$ which has $f_3(\mathbf{v})\leq 1$. For the terms $min(a_i,r)=r$, $\frac{x_i^2}{a_i^2}\leq\frac{x_i^2}{min(a_i,r)^2}$. Hence, $f_2(\mathbf{v})\leq f_3(\mathbf{v})\leq r^2$. Similarly, $f_1(\mathbf{v})\leq f_3(\mathbf{v})\leq r^2$. These two inequations indicate the vector v is also in S_1.

The area of S_1 is hard to calculate. But we still have S_2 is contained in S_1 and can be calculated by formula. Use S_2 to replace S_{goal}, we get the theorem below, which gives us a lower bound of success probability:

Theorem 4. *The success probability of sphere sampling Pr with radius* $r=m\|\mathbf{b}_1^*\|$ *is greater than* $\prod_{i=1}^{n}\frac{min(\lceil\frac{r}{\mathbf{b}_i^*}\rfloor+0.5,\|\mathbf{b}_1^*\|/\|\mathbf{b}_i^*\|)}{\lceil\frac{r}{\mathbf{b}_i^*}\rfloor+0.5}$. *If lattice basis* $\mathbf{B}$ *satisfies the GSA and Hermite factor is q, this probability can be written as* $Pr\geq\prod_{i=1}^{n}\frac{min(\lceil mq^{i-1}\rfloor+0.5,q^{i-1})}{\lceil mq^{i-1}\rfloor+0.5}$

Proof. We have $Pr=S_{goal}/S_{sampling}$. According to Lemma 1, $Pr\geq S_2/S_{sampling}$ where S_2 defined in Lemma 1 is an ellipsoid which has the parameters $min(\lceil\frac{r}{\mathbf{b}_i^*}\rfloor+0.5,\|\mathbf{b}_1^*\|/\|\mathbf{b}_i^*\|)$.

fridaya@uef.fi

5 The Analysis of Sphere Sampling Reduction

In Sects. 3 and 4 we have described our algorithm and its success probability. In this section we will analyze our algorithm and give out the rules of choosing parameters.

5.1 The Range of Sphere Sampling

We have described how to conduct random sampling on sphere. In most sampling algorithms, the sampling range is small. For example, the random sampling reduction always samples vectors which have specific integer representations. Their representations have a lot of 0 in the head and the elements which are not zero appear in a low proportion and only exist in the tail part of these representations. In our new algorithm, although we generate representation in a distinctive way, similar phenomenon still exists. In this section we try to give some rules of our algorithm by using the tool of probability theory.

We generate vector $\mathbf{s}$ on a sphere with radius r where it has the form $\mathbf{s} = (s_1, ..., s_n) = (\frac{x_1}{\sqrt{x_1^2+...+x_n^2}}r, ..., \frac{x_n}{\sqrt{x_1^2+...+x_n^2}}r)$, and $x_1, ..., x_n$ are samples from a normal distribution. Note that the integer representation which is output by algorithm is $\mathbf{repv} = (repv_1, ..., repv_n) = (\lceil \frac{x_1}{\sqrt{x_1^2+...+x_n^2}} \cdot \frac{r}{\|\mathbf{b}_1^*\|} \rfloor, ..., \lceil \frac{x_n}{\sqrt{x_1^2+,...,x_n^2}} \cdot \frac{r}{\|\mathbf{b}_n^*\|} \rfloor)$. The ith element of integer representation is not zero if and only if $\frac{x_i}{\sqrt{\sum_{j=1}^n x_i^2}} \cdot \frac{r}{\|\mathbf{b}_i^*\|} \geq 0.5$ or $\frac{x_i}{\sqrt{\sum_{j=1}^n x_i^2}} \cdot \frac{r}{\|\mathbf{b}_i^*\|} \leq -0.5$. To calculate the probability, when given the r and basis $\mathbf{B}$, we should calculate the probability $-0.5\frac{\|\mathbf{b}_i^*\|}{r} < \frac{x_i}{\sqrt{\sum_{i=1}^n x_i^2}} < 0.5\frac{\|\mathbf{b}_i^*\|}{r}$. Since all $\frac{x_i}{\sqrt{\sum_{i=1}^n x_i^2}}$ have the same distribution, denote random variable $X = \frac{x_i}{\sqrt{\sum_{i=1}^n x_i^2}}$.

$$
\begin{aligned}
Pr(-0.5\frac{\|\mathbf{b}_i^*\|}{r} < X < 0.5\frac{\|\mathbf{b}_i^*\|}{r}) &= Pr(X^2 < 0.25(\frac{\|\mathbf{b}_i^*\|}{r})^2) \\
&= Pr(\frac{x_i^2}{\sum_{i=1}^n x_i^2} < 0.25(\frac{\|\mathbf{b}_i^*\|}{r})^2) \\
&= Pr(1/(1 + \frac{\sum_{j\neq i}^n x_j^2}{x_i^2}) < 0.25(\frac{\|\mathbf{b}_i^*\|}{r})^2) \\
&= Pr(1/(1+Y) < 0.25(\frac{\|\mathbf{b}_i^*\|}{r})^2)
\end{aligned}
$$

It's easy to see that Y is a kind of F-distribution, which confirms $Y \sim F(n-1, 1)$. Hence, $Pr(-0.5\frac{\|\mathbf{b}_i^*\|}{r} < X < 0.5\frac{\|\mathbf{b}_i^*\|}{r}) = Pr(Y > 1/(\frac{\|\mathbf{b}_i^*\|}{2r})^2 - 1)$ can be given by checking the distribution of $F(n-1, 1)$. In practice, $r = c\|\mathbf{b}_1^*\|$ where c is some real number. According to the geometric series, $\|\mathbf{b}_1^*\|/\|\mathbf{b}_i^*\| = q^{i-1}$. So the $Pr(-0.5\frac{\|\mathbf{b}_i^*\|}{r} < X < 0.5\frac{\|\mathbf{b}_i^*\|}{r}) = Pr(Y > 1/(\frac{1}{2cq^{i-1}})^2 - 1)$, which means the range of integer representation generated by our algorithm can be certainly determined during the algorithm.

This analysis helps us improve our algorithm in advance. We will give out the modified algorithm in Sect. 6.

fridaya@uef.fi

5.2 The Choice of Radius

How to choose a fit radius of sphere is the main question. Our goal is to find some vectors which satisfy $\|\pi_i(\mathbf{v})\| < \|\mathbf{b}_i^*\|$ for some i. If the radius r is large, vectors generated by the algorithm are much likely to be "bad" and hardly hold the condition. But if radius r is too small, there will be too many zero representations which are useless for our SVP solution. An obvious standard we can get is $\|\mathbf{b}_i^*\|$. $\|\mathbf{b}_1^*\|$ is the length of shortest vector of reduction basis. We give a factor m to control the radius $r = m\|\mathbf{b}_1^*\|$. Back to the algorithm, there are two 'if-else' which affect the efficiency of algorithm.

On the one hand, the basis will be updated if and only if the lattice vector generated is small in some "direction". We need to adjust m to make our sampling algorithm have enough outputs which can be used to update our basis. According to the heuristic, r should uniformly distribute in different dimensions. To make $\lceil \frac{r}{\sqrt{n}\|\mathbf{b}_n^*\|} \rfloor \|\mathbf{b}_n^*\| < \|\mathbf{b}_n^*\|$, since $\lceil \frac{r}{\sqrt{n}\|\mathbf{b}_n^*\|} \rfloor \|\mathbf{b}_n^*\| \leq (\frac{r}{\sqrt{n}\|\mathbf{b}_n^*\|} + 0.5)\|\mathbf{b}_n^*\|$, if inequation $(\frac{r}{\sqrt{n}\|\mathbf{b}_n^*\|} + 0.5)\|\mathbf{b}_n^*\| \leq \|\mathbf{b}_n^*\|$ holds, $\lceil \frac{r}{\sqrt{n}\|\mathbf{b}_n^*\|} \rfloor \|\mathbf{b}_n^*\| < \|\mathbf{b}_n^*\|$ must hold. We have $m \leq 0.5\sqrt{n}\frac{\|\mathbf{b}_n^*\|}{\|\mathbf{b}_1^*\|}$. Similarly, by using relationship $\|\pi_i(\mathbf{v})\| < \|\mathbf{b}_i^*\|$ for different $1 \leq i \leq n$, we can get n different formulas, and every formula gives an estimation of m. To be convenient, in most of cases, $i = n$ can give us a good estimation.

On the other hand, the sampling part will start again when the representation is zero representation. So our algorithm should generate seldom zero representation. Recall our representation which can be denoted as $(\lceil \frac{y_1}{\|\mathbf{b}_1^*\|} r \rfloor, ..., \lceil \frac{y_n}{\|\mathbf{b}_n^*\|} r \rfloor)$, where $y_i = \frac{x_i}{\sqrt{\sum_{j=1}^{n} x_j^2}}$, for $i = 1, ..., n$. $\lceil \frac{y_i}{\|\mathbf{b}_i^*\|} r \rfloor = 0$ if and only if $-0.5 \leq \frac{y_i}{\|\mathbf{b}_i^*\|} r < 0.5$. We use a geometric method to discuss this problem. Consider the vector $(y_1, ..., y_n)$, which is a sample of random sampling on a unit sphere. When the vector is multiplied by r, it is a random sampling on sphere with radius r, if we multiply $\|\mathbf{b}_i^*\|$ in the both sides of inequations. The n different inequations for $i = 1, ..., n$ actually constitute a range of hypercube, and vectors inner the hypercube will be mapped to the zero representation when the algorithm executes.

The problem above can be seen to find some r which makes the proportion of sphere in the hypercube to whole sphere be small. Take a 3-dimension situation as an example. There will be three cases determined by the relationship between hypercube and sphere. (1) Hypercube is in the sphere like Fig. 2. In this case, the radius of sphere have the relation $r > (\sum_{i=1}^{n}(0.5\|\mathbf{b}_i^*\|)^2)^{\frac{1}{2}}$. Because our algorithm generates the vectors on the sphere which means all these representation generated by our algorithm is non-zero. (2) Hypercube is inscribed on the sphere. In this case, we have radius $r = (\sum_{i=1}^{n}(0.5\|\mathbf{b}_i^*\|)^2)^{\frac{1}{2}}$. And the 2^n vertexes of hypercube are on the sphere. If and only if the process gets these 2^n points, we will get zero representation. (3) The hypercube is out of the sphere as it is shown in Fig. 1. In fact, this case is the most general in sampling algorithm. The estimation of lattice short vector is also in this section. It is clear that the bigger the radius is, the fewer zero representation will be generated by our algorithm. However, as we have described in the last paragraph, the radius should be small enough to make the basis update efficiently. So it is a balance between

fridaya@uef.fi

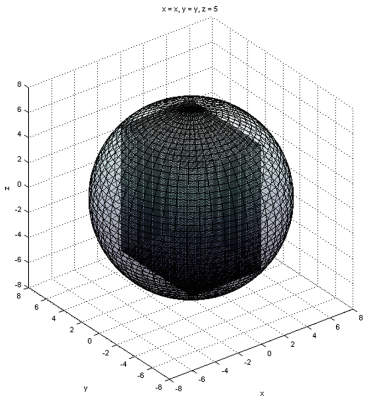


Fig. 1. Sphere intersect Box

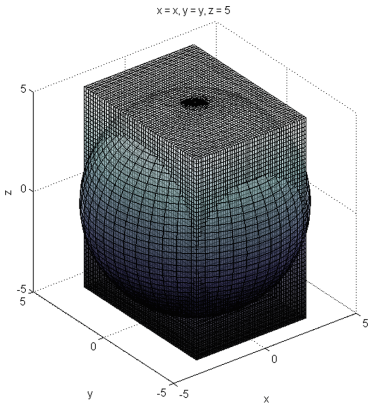


Fig. 2. Box in the sphere

the long radius to make the zero representation fewer and the short length to make the basis update efficiently. The change of probability to generate zero representation can be given by calculating the area intersection of sphere and hypercube.

Define the area of sphere which is contained in the box is S_{in}, and the area of whole sphere is S_{sphere}. We denote the proportion S_{in}/S_{sphere} as P_b which means the proportion of the area of sphere in the box. When the basis is given, the area of box is determined. Obviously, P_b decreases when radius r increases. And we can get the curve of P_b and r. There are several specific values which are the break points of this curve. Except $(\sum_{i=1}^{n}(0.5\|\mathbf{b}_i^*\|)^2)^{\frac{1}{2}}$, there are some other specific values which should be concerned. For every i, $0.5\|b_i^*\|$ is an inflection point of curve. In these points, the tendency of curve will change a lot. The most important length is $0.5\|\mathbf{b}_1^*\|$. It is the first point which will make the probability change a lot in the third case. In that length, the radius is not too long to update

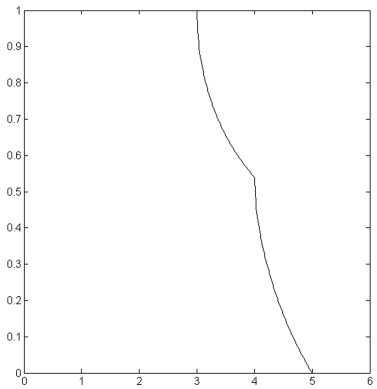


Fig. 3. Proportion of zero representation with radius r in dimensions 2

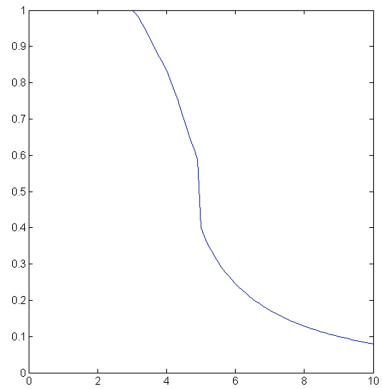


Fig. 4. Proportion of zero representation with radius r in dimensions 3

fridaya@uef.fi

the basis but can keep zero representation in a low probability. In this case, the estimation of $m = 0.5$. The pictures above Figs. 3 and 4 show the proportion with radius r in 2 and 3 dimensions.

Two estimations above give us a range where the radius should be chosen. In our algorithm we choose the radius between these two estimations. Next subsection we will describe the advanced algorithm and the parameters in practice.

5.3 Practical Sphere Algorithm

After the analysis, we give an improvement variant of our algorithm. Compared with the algorithm in Sect. 3, the modified algorithm has two differences. First, the process to find integer representations is simplified: the elements of integer representations which are zero in our sampling algorithm will not be calculated in this variant. Second, the parameter radius r floats in the range given in Sect. 5.2. This modified algorithm is given in Appendix A.

This modified algorithm adjusts the radius gradually by a counter. If there are more than 10000 vectors sampled with some radius r but don't call the reduction algorithm, the radius will change. In practice, this algorithm works well, because all this parameter can be adjusted automatically to make sure the algorithm run efficiently.

6 Experiments in SVP Challenge

In this part, we use our algorithm to solve the SVP challenge in 80 and 100 dimensions. The original basis is generated by the SVP challenge website (see Appendix B). BKZ algorithm was used in our algorithm as a tool to renew basis. To save the time in running the BKZ algorithm, we choose a small block-size which is 20. Our procedure returns vectors (see Appendix B) with length smaller than $1.05GHL$ within 50 rounds in 100-dimensions and 30 rounds in 80-dimensions. The codes and basis of our procedure are shown in Github [22].

7 Conclusion

In this paper, we describe a new sampling algorithm for generating short vectors which can be used to renew basis by reduction algorithms such like BKZ and LLL. Compared with other sampling algorithm, our algorithm restrict the length of vectors by some parameters of sampling sphere, which means we can adjust and observe the performance by choosing appropriate parameters. To check the efficiency and availability, we solve 80 to 100 dimension random lattice challenge by our sphere sampling reduction.

A A SVP Challenge

80 Short Vector: [–311 98 19 209 –42 –76 –28 51 202 –180 –36 –80 –392 75 –67 –158 261 –407 278 –226 52 –252 397 294 –194 –7 –394 –285 177 –73 18 114 259

fridaya@uef.fi

109 22 29 408 58 491 319 –162 –67 5 –344 119 –604 7 –55 –517 –106 –271 430 –264 128 –190 –25 –506 –339 –277 105 133 152 –6 347 –147 53 –452 –245 397 –89 –131 –279 66 128 –41 –224 414 –4 157 –170] 100 Shortest Vector: [–11 96 226 –103 –345 196 –135 265 330 –110 248 –104 –237 422 337 121 331 –459 90 –15 588 –165 87 –155 –340 269 38 291 480 –221 –386 –134 247 –317 –244 –101 –151 –138 –59 –187 –78 –423 –473 142 185 69 –50 303 264 –220 –274 84 –346 90 162 –31 174 –83 81 576 –488 –279 64 –14 –42 –358 264 18 –45 –273 34 268 –134 443 70 –178 558 –356 532 115 –8 634 193 –157 –23 250 –321 –270 –30 122 63 101 –93 –196 1 –467 –67 474 –343 –119]

B Modified Algorithm

Algorithm 7. Modified Sphere Sampling Reduction Algorithm

Input:
B:a lattice basis $\mathbf{B} = [\mathbf{b}_1, ..., \mathbf{b}_n]$.
insertsign:a sign to record whether or where to update.
droppr: probability which is gave up in integer representations.
$listF_d$:a matrix vector memorizes the F-distribution list, it has two rows,while the first row is the probability and the second row is corresponding value of variable.
Blocksize:the blocksize of BKZ
bound:the bound of subscript which is used to control insertion vectors,$bound \leq n$.
Output:
v:a short lattice vector satisfying $\|\mathbf{v}\| \leq 1.05GHL$, $1.05GHL$ is an estimation of the shortest vector.

while $\|\mathbf{b}_1^*\| \geq 1.05GHL$ **do**
  calculate the low bound factor $m = 0.5\sqrt{n}\frac{\|\mathbf{b}_n^*\|}{\|\mathbf{b}_1^*\|}$
  while $m \leq 0.5$ **do**
    $r = m\|b_1^*\|$
    calculate the vector $repvmask = r/\|b_1^*\|, ..., r/\|b_n^*\|$
    for $rr = 1, ..., n$ **do**
      if $1/(\frac{\|\mathbf{b}_i^*\|}{2r})^2 - 1 > vx$ **then**
        $repvmask(rr) = 0$
      else
        $repvmask(rr) = 1$
      end if
    end for
    $insertsign = 0$
    while $insertsign = 0$ **do**
      $num = num + 1$
      if num=10000 **then**
        $m = m + 0.01$
        $num = 0$
      end if
      call the MoGenSphere$(r, repvmask)$ to generate a random point **s** on a sphere with radius r.
      get representation **repv** from **s** according to GenRepv(**s**).
      generate lattice vector **v** from FindVec($repv$).
      for $j = bound, ..., 1$ **do**
        if $\|\pi_j(\mathbf{v})\| \leq \|\mathbf{b}_j^*\|$ **then**
          $intersign = j$
        end if
      end for
    end while
  end while
  use $BKZ(\mathbf{B}, \mathbf{v}, insertsign)$ to renew the basis.
end while

fridaya@uef.fi

Algorithm 8. Modified Random Sampling on a Sphere(MoGenSphere)

Input:
r:the radius of a sphere.
n:the dimensions of sphere.
$repvmask$: vector to sign the calculate relation.
Output:
s:a point on this sphere

for $i = n, ..., 1$ **do**
  $x_i \leftarrow D$
  $sum = sum + x_i^2$
end for
$\mathbf{s} = (x_1, ..., x_n)$
for $j = 1, ..., n$ **do**
  if $repvmask(j) = 1$ **then**
    $\mathbf{s}(j) = r/\sqrt{sum} \cdot \mathbf{s}(j)$
  else
    $\mathbf{s}(j) = 0$
  end if
end for

References

1. Ajtai, M.: The shortest vector problem in L2 is NP-hard for randomized reductions (extended abstract). In: Thirtieth ACM Symposium on Theory of Computing, pp. 10–19. ACM (1998)
2. Fincke, U., Pohst, M.: A procedure for determining algebraic integers of given norm. In: Computer Algebra, EUROCAL 1983, European Computer Algebra Conference, London, England, 28–30 March 1983. Proceedings. DBLP, pp. 194–202 (1983)
3. Pohst, M.: On the computation of lattice vectors of minimal length, successive minima and reduced bases with applications. ACM (1981)
4. Kannan, R.: Improved algorithms for integer programming and related lattice problems. In: ACM Symposium on Theory of Computing, 25–27 April 1983, Boston, Massachusetts, USA, pp. 193–206. ACM (1983)
5. Schnorr, C.P., Euchner, M.: Lattice basis reduction: improved practical algorithms and solving subset sum problems. Math. Program. **66**(1–3), 181–199 (1994)
6. Stehl, D., Watkins, M.: On the extremality of an 80-dimensional lattice. In: Algorithmic Number Theory, International Symposium, ANTS-IX, Nancy, France, 19–23 July 2010. Proceedings, pp. 340–356 (2010)
7. Gama, N., Nguyen, P.Q, Regev, O.: Lattice enumeration using extreme pruning. In: International Conference on Theory and Applications of Cryptographic Techniques, pp. 257–278. Springer (2010)
8. Zheng, Z., Wang, X., Guangwu, X.U., et al.: Orthogonalized lattice enumeration for solving SVP. Sci. China (Inf. Sci.) **61**(3), 1–5 (2018)
9. Ajtai, M., Kumar, R., Sivakumar, D.: A sieve algorithm for the shortest lattice vector problem. In: ACM Symposium on Theory of Computing, pp. 601–610. ACM (2001)
10. Nguyen, P.Q., Vidick, T.: Sieve algorithms for the shortest vector problem are practical. J. Math. Cryptol. **2**(2), 181–207 (2008)
11. Pujol, X., Stehl, D.: Solving the shortest lattice vector problem in time 2 2.465n. IACR Cryptology Eprint Archive, 2012, 2009
12. Rckert, M., Schneider, M.: Estimating the security of lattice-based cryptosystems. IACR Cryptology Eprint Archive, 2010

fridaya@uef.fi

13. Lenstra, A.K., Jr Lenstra, H.W., Lovász, L.: Factoring polynomials with rational coefficients. Math. Ann. **261**(4), 515–534 (1982)
14. Gama, N., Nguyen, P.Q.: Finding short lattice vectors within Mordell's inequality. In: ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 2008. DBLP, pp. 207–216 (2008)
15. Schnorr, C.P.: Lattice reduction by random sampling and birthday methods. In: STACS, vol. 2607, pp. 145–156 (2005)
16. Masaharu, F., Kenji, K.: An accelerated algorithm for solving SVP based on statistical analysis (preprint). J. Inf. Process. **23**(1), 67–80 (2015)
17. Zhu, G., Wang, X.: A genetic algorithm for searching the shortest lattice vector of SVP challenge. In: Conference on Genetic and Evolutionary Computation, pp. 823–830. ACM (2015)
18. Ppelmann, T., Ducas, L., Gneysu T.: Enhanced lattice-based signatures on reconfigurable hardware. In: Cryptographic Hardware and Embedded Systems C CHES, pp. 353–370. Springer, Heidelberg (2014)
19. Babai, L.: On Lovaiasz' lattice reduction and the nearest lattice point problem. Combinatorica **6**(1), 1–13 (1986)
20. Marsaglia, G., Bray, T.A.: A convenient method for generating normal variables. SIAM Rev. **6**(3), 260–264 (1964)
21. Website of SVP Challenge. http://www.latticechallenge.org/svp-challenge/
22. Github. https://github.com/audreycauchy/Sphere-Sampling-reduction/tree master

fridaya@uef.fi

Hybrid Dependencies Between Cyber and Physical Systems

Sandra König[1(✉)], Stefan Rass[2], Benjamin Rainer[1], and Stefan Schauer[1]

[1] Center for Digital Safety and Security, Austrian Institute of Technology GmbH, Vienna, Austria
{sandra.koenig,benjamin.rainer,stefan.schauer}@ait.ac.at
[2] Institute of Applied Informatics, System Security Group, Universität Klagenfurt, Klagenfurt, Austria
stefan.rass@aau.at

Abstract. Situational awareness is often a matter of detailed local information and proportionally limited view on the bigger picture. Conversely, the big picture avoids complicating details, and as such displays the system components as atomic "black boxes". This work proposes a combination of local and global views, accounting for a common practical division of physical and cyber domains, each of which have their own group of experts and management processes. We identify a small set of data items that is required about the physical and cyber parts of a system, along with a high-level description of how these parts interoperate. From these three ingredients, which we call physical, cyber and hybrid "awareness" (meaning just knowledge about what is there), we construct a simulation model to study cascading effects and indirect implications of distortions in a cyber-physical system. Our simulation model is composed from coupled Mealy automata, and we show an instance of it using a small cyber-physical infrastructure. This extends the awareness from "knowing what is" to "knowing what could happen next", and as such addresses a core duty of effective risk management. Manifold extensions to this model are imaginable and discussed in the aftermath of the definition and example demonstration.

Keywords: Security · Cascading effects · Probabilistic automata · Critical infrastructure

1 Introduction

Situational awareness is a major issue in enterprise security and risk management. It provides an overview on the current situation within an organization and indicates potential threats and emerging risks. Problems in the context of situational awareness may partly be attributed to insufficient knowledge about prospective consequences of certain behavior or implications of incidents. This lack of information is often due to opaque details about the system dynamics and complex interactions between system components. Since the world becomes

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 550–565, 2019.
https://doi.org/10.1007/978-3-030-22868-2_40

fridaya@uef.fi

more and more interconnected, the interdependencies between information systems are often hard to understand, difficult to model and challenging to predict.

Prominent examples for the importance of situational awareness in security and risk management are Critical Infrastructures (CIs) such as, i.e., airports, maritime ports or power systems. Situational awareness is a crucial and non-trivial task to achieve in such CIs because they include a large landscape of system components which are in many cases highly heterogeneous with respect to physical and cyber (software services) assets. Being aware and maintaining an up-to-date picture of what happens is a core duty of risk management and a complex task. Research in the field of situational awareness in risk management goes beyond assessing the current situation by also trying to predict how an incident may probably "spread" (cascade) in the CI at hand and which assets may be affected over time [5,11]. Predicting possible cascading effects can be tremendously helpful for decision makers providing them a possible picture of "possible and likely scenarios". However, it lies in the nature of such models that their prediction accuracy highly depends on the available information about the CI, on the possibility to account for the current prevalent situation (e.g., states of assets, types of incidents happening) and last but not least on the capabilities of the model itself.

Our contribution relates to enterprise risk management in general and security awareness therein in particular. Specifically, we describe a simulation framework to study the implications of security incidents, including cascading effects on related components by simulating the indirect effects of signals and alarms within a system. In a nutshell, we can delineate the model as a system of coupled probabilistic Mealy automata, which logically amounts to a division of the dynamics of the system compound, from the dynamics of each system component themselves. Thus, we treat incidents as signals and the cascading effects thereof as a response of the coupled automata which model the dependent assets of the CI. Further, we adopt a *local view* on each system component, which is separate from the *global view* that treats the system's parts as black boxes and describes only their interplay. The goal of this work is extending this awareness by "what could happen next", using simulation. In particular we propose a (simulation) model that

- allows capturing the signal–response nature of occurring incidents and their cascading effects among assets in a CI;
- covers a wide range of other models (e.g., Markov chain based models, percolation theory based models, etc.);
- is computational feasible even for large CIs.

Organization: Sect. 2 describes an intentionally small core-base of knowledge required to set up our model; ideally, the model that Sect. 3 then formally introduces, will be definable in an automated way to the widest extent possible, and using mostly information that already exists in an enterprise. Applications of simulation are manifold, and the related work in Sect. 1.2 summarizes a selection of past approaches to the issue in the literature. Section 1.1 puts the goals

fridaya@uef.fi

and assumptions of this work into more compact terms, thus outlining the applications thereof, which we extend in the discussion Sect. 5 at the end of the paper.

1.1 Problem Description

In order to gain insights we separate the set of assets of a CI into *physical* and *cyber* parts. Assets belonging to either area usually are subject to their own risk management processes and domain expert groups (who are not necessarily disjoint). The connection between physical and cyber parts of the system, however, creates its own realm of incidents and risk landscape; these dependencies, which we call *hybrid* (because they put physical and cyber assets into relation) lead to the need of *hybrid risk awareness*. For example, if a physical intrusion is supported by a cyber attack on an electronic door lock, is a hybrid incident in our terminology. Likewise, if the physical check clock at a building's exit records an employee leaving, but the (logical) logging records a login after the person left, this logical inconsistency could trigger an alarm and raise actions by the responsible staff. This kind of incident is hybrid in the sense of being a combination of physical and cyber information, and the models in this work are designed towards simulating the potential implications of such events. Other parts of the system, e.g., social or purely technical information on the system, can also be integrated in the analysis but we refrain from doing so here to keep the focus on the construction of the hybrid model.

Despite the huge knowledge that often exist in a company, this knowledge is typically not collected or even combined. Rather, there often exists a group of experts with only limited communication between them, yielding an insufficient awareness on the overall situation. We account for this "division of knowledge" by speaking about Physical Situational Awareness (PSA) and Cyber Situational Awareness (CSA) to summarize both domains, not yet spanning any connection between the two. However, the existence of connections is evident, as a simple example shows.

Example 1. Suppose that a physical system (part of the PSA) has a fire detector going off, on which the cyber system (part of the CSA) must react by informing staff members (e.g., raise an alarm to leave the building) and perhaps activate backup services (as known in the CSA) to compensate a shutdown of servers that may be affected by the fire. The fact that a physical machine may be affected by the fire, however, is a matter of physical circumstances, and hence an information found in the PSA. The real sequence of events is far more complex, and typically involves human interactions in the chain. For example, an alarm may not automatically be forwarded to external forces (e.g., fire fighters), unless a human operator confirms that it was no false alarm.

While Example 1 is simple enough to admit a manual analysis of implications, more complex incidents may not be as trivial to study. For instance, if the system detects anomalies encountered in access patterns to services or servers, this may be an indication of some hacking activities, along which privilege escalations

fridaya@uef.fi

could have happened. The potential impact of such an incident may run far deeper and, in its entirety, cannot be fully assessed by experts from either domain (physical or cyber).

The main problem is thus the limited exchange and information about how consequences of a failure in one system trigger subsequent events and effects in other systems, since the overall (hybrid) situational awareness is insufficient. Thus, we develop a model that describes:

- cascading effects and how the state of an asset evolves after an alarm from either PSA or CSA;
- the hybrid situational awareness in order to get an overview about the "state" of the whole system, e.g., described by a network where each node representing an asset is either green, orange or red, depending on how much it is affected (representing the situations "complete failure", "partial damage" or "not affected").

1.2 Related Work

An early approach to model and investigate cascading effects was the Cross Impact Analysis (CIA) [4] and an alternative approach provided in [20]. The CIA allows describing how the dependencies between events would affect events in the future. This method is extended by the Cross Impact Analysis and Interpretative Structural Model (CIA-ISM) [1] which is applied in emergency management to analyze the effects between critical events and to obtain a view on future consequences. One may consider these models as the predecessors of nowadays stochastic models.

A comprehensive overview on existing models on cascading effects in power systems is given in [5] and for CIs in [11]. In both reviews, methodologies are classified according to their main feature, contribution and use case/application. Additionally, cascading failure models are compared and their advantages and disadvantage are summarized. Cascading effects in CIs are also analyzed in [8] and [10].

Cascading effects in interconnected networks have been investigated based on physical models such as percolation [2,6] or on topological properties and network analysis [12]. The model we introduce generalizes percolation models to a certain extent (cf. also Sect. 5) such that we can mimic at least original percolation models. A core insight when investigating cascading effects is that it is impossible to exactly predict the future development due to many unknown and hidden factors and, not least, because humans are involved. Thus, it is indispensable to include some stochastic components in the applied models used. Hence, Markov chains with memory m are used in [22] to model situations where the transition probabilities do not only depend on the current state but on the past m ones (i.e., a Markov chain of order m). In this setting, the transition matrix turns into a tensor of order $m+1$. While the model allows to consider the states of neighboring assets (because of the common state space) it is not intended to have events (i.e., alarms, signals) which may trigger state changes

fridaya@uef.fi

themselves. Our model does extend this by providing the possibility to react on external events (coming from other assets or even from an asset itself).

One way to handle the complexity of Markov models is to use an abstract state space whose states contain all the information relevant for the dynamics of the system. This approach was taken in [17]. The method of representing infrastructures by single states is extended by the Interdependent Markov Chain (IDMC) model [16]. This model allows describing cascading failures in interdependent infrastructures. Therein, each system is described by one discrete-time Markov chain and the interdependencies between these systems are represented by coupling the Markov chains. This can also easily be achieved by our model because it generalizes Markov chains (each automaton can be a separate Markov chain) and does also inherently support the coupling of these chains. Further, a conditional Markov transition model has been applied to describe cascading failures in electric power grids [21]. Here, the transition probabilities are derived from a stochastic for flow redistribution. Even this can be covered by our model by extending the state transitions within the automata to be time-dependent.

Another approach to model cascading effects in CIs is application of branching processes that are typically used to describe growth of a population. The applicability of branching processes on modeling cascading effects is investigated in [3,14].

2 Basic Hybrid Dependency Model

Our basic dependency model is a graph consisting of assets (e.g., cyber and physical assets) represented by vertices and various dependencies between them represented by directed edges. We consider a situation where information about both the physical and the cyber situation of an infrastructure (such as a port) are available. In this section we describe a way how to combine these two views into a hybrid one that includes information about the interplay between them. The main benefit of such a hybrid view is its potential to investigate cascading effects and to detect threats arising from a combination of physical and cyber events.

Our model building will start from what we believe to be a minimal base of information needed to describe

- the assets themselves,
- the mutual dependencies and relations between assets,
- the events and alarms related to the assets.

However, the attributes that are taken into account can be arbitrarily extended to suit any needs. Typically in risk management, *threats* would naturally extend this list, but to avoid complicating matters, we let threats correspond to alarms (if a threat becomes relevant, exploited, active, ...).

fridaya@uef.fi

2.1 Definitions and Notation

Let the infrastructure contain *assets* that are mutually dependent on one another. Furthermore, each asset carries a set of attributes, e.g., it can be *cyber* (e.g., a server, application, service) or *physical* (e.g., a video camera, a restricted area), have a certain criticality, and many more.

We define *situational awareness* as the entirety of information available to the risk analyst a priori and a posteriori, obtained from a simulation analysis. This is to reflect the fact that we can typically expect the relevant information to come as a collection of assets (as, for example, in a Configuration Management Database (CMDB) or similar) and that the risk analysis will have its feedback enter the same risk database. Likewise, and for model building, we will thus simply refer to the underlying collection of information as the *situational awareness* and divide it into parts related to physical, cyber and hybrid situational awareness; the latter of which describes the dependencies between physical and cyber assets.

In this sense, let PSA and CSA be collections of all information known and relevant for physical and cyber assets, respectively. This information includes, but may practically not be limited to, an attribute describing the type of asset, its criticality (meaning or value for the overall infrastructure), and its current status. Thus, the Hybrid Situational Awareness (HSA) sets physical and cyber assets into relation, based on their corresponding attributes.

A state change of assets, most of the time a hazardous external or internal influence, can be described by an incident (i.e., an *alarm* describing the details of the occurred incident), which extends the CSA, PSA, and HSA, correspondingly. Each alarm is characterized by a *type* (such as fire, flood or similar), criticality and a *status* that describes its current state. An alarm may involve any number of physical or cyber assets (or both). Alarms are categorized either as "cyber" or "physical", depending on the origin of the corresponding alarm. Additionally, an alarm can be considered as "hybrid" if it is raised due to a critical combination of cyber and physical incidents. The criticality of an alarm describes its impact on the corresponding assets (e.g., "complete failure", "partial damage" or "not affected"). We assume that the number of occurring alarms is not limited for any type of asset.

Finally, we define a *scenario* as the *temporal* situational awareness that relates to and involves any number of alarms and assets. In being a time series, we assume that it carries a notion of time (e.g., time-stamp), plus an information about the alarm or involved asset, and the status of the respective entity involved. This value will later be defined more rigorously along the description of the simulation model in Sect. 3.

The overall information embodied in the infrastructure's assets can be put into a graphical representation. This is essentially a directed graph that displays dependencies between assets, as far as the given information induces them. In such a graph model, alarms can be included as special purpose nodes that are connected to those assets they affect. We will postpone a formal definition of this dependency model until later, when we have described the asset behavior in more detail in Sect. 3.

fridaya@uef.fi

2.2 Preliminary Considerations and Design Choices

Towards a description of the asset dynamics under simulated incidents, we assume that each asset (regardless of cyber or physical) is in one of a fixed number of predefined states that represent their functionality, say "complete failure", "partial damage" or "not affected". This scale is the same for all assets, be it physical or cyber.

If we think of a component to individually be in one out of three possible states, it seems convenient to apply a Markov chain model to describe the changes inside each node. Unlike as for conventional Markov chains, the transition between states depends on the state of related components.

For example, if the door to a room where a critical server is installed is open, then the server has a decent chance to switch from "not affected" into "partial damage" or even "complete failure", depending on what an adversary does. Likewise, if some other server controls an electronic door lock, then a "complete failure" server may cause the door lock to switch from "not affected" into "partial damage", if the server answers all requests from the door lock positive to signalize openness.

Formally, where a conventional Markov chain would have transition probabilities $p_{ij} = \Pr(X_{t+1}|X_t)$ from the state of component $X_t \in \{\text{complete failure, partial damage, not affected}\}$ at time t, a generalized Markov model could extend this transition towards conditioning on the state Y_t of another component at time t, i.e., we now have $p_{ij} = \Pr(X_{t+1}|X_t, Y_t)$. In that sense, the Markov chains for components X, Y are *coupled*, and the overall model would be a huge Markov chain with an exponentially large state space, since there would be at least $\geq 2^{O(N)}$ states in the chain for N assets, each of which can individually take its own state. This motivates a different approach stemming from the same intuition yet taking a different route, namely,

(i) an outer dependency structure that specifies which components influence one another, i.e., which transitions depend on the states of other components,
(ii) inner transition regimes for each component,

similarly to a percolation model [7,9].

Besides these computational problems, there are important issues that a classical Markov chain model does not capture. In particular, the transitions do not necessarily occur at fixed points in time (as they would for a Markov chain), but are mostly due to external triggers, such as alerts. Letting the transition depend on an external force or signal essentially transforms the Markov chain into a probabilistic automaton. However, incorporating such external signals into our so-far described model is easy since they can just treat an alarm as another random variable, affecting (or not affecting) the transition likelihood accordingly. Formally, if Z_k denotes an alarm of kind k (just to distinguish multiple alarms originating from different incidents), we can make a transition depend on this alarm variable, too, i.e., we can refine the inner state transition regime of a component to $p_{ij} = \Pr(X_{t_1}|X_{t<t_1}, Y_{t<t_1}, Z_k(t_1))$. Note that our notation is somewhat sloppy here but meant to illustrate an important assumption: the alarm

fridaya@uef.fi

Z_k comes in at time t_1, and exactly the *change* of the random variable Z_k at this point in time triggers the transition (with its specified likelihood), depending on the past (now current) state of X before time t_1 (denoted as $X_{t<t_1}$) and the state of its relative Y before time t_1 (again, denoted as $Y_{t<t_1}$).

As this gets quite complex with an increasing number of factors influencing the transitions, we consider a different model to describe the inner dynamics of an asset, based on finite automata. Applications of probabilistic automata to security problems have been discussed in [13,19] but so far probabilistic automata have rarely been applied to security problems and especially to risk management. Since a state transition is not only triggered upon external events but may itself cause subsequent cascading effects, the automata will be endowed with an output facility to forward notifications and alarms to other assets. Our choice is here on Mealy automata to describe an asset's behavior. It is important to note that Mealy automata perfectly fit our assumption that a *successful* transition from one state in another will trigger the output, in contrast to Moore automata where an output is triggered by a state.

3 The Simulation Model

Together with the notation introduced in Sect. 2.1 we now get the following model. We let $\mathcal{MA}$ denote the set of all finite Mealy-automata, each of which is defined in the familiar way, where we shall instantiate the components of the automaton suitably to describe the state of the critical infrastructure.

Definition 1 (Mealy-automaton). *A* Mealy-automaton *is a six-tuple* $M = (S, \Sigma_{in}, \Sigma_{out}, \delta, \lambda, s_0)$ *that satisfies all of the following conditions:*

1. S *is a nonempty and finite set of states.*
2. Σ_{in} *is a nonempty and finite input alphabet.*
3. Σ_{out} *is a nonempty and finite output alphabet.*
4. $\delta : S \times \Sigma_{in} \to S$ *is a transition function.*
5. $\lambda : S \times \Sigma_{in} \to \Sigma_{out}$ *is an output function (an output occurs if and only if a state transition happens).*
6. $s_0 \in S$ *is the initial state.*

The same asset may generally react differently even on the same alarm, say, if its reaction is uncertain or indeed undetermined (as dependent on hidden variables that we can neither capture nor control). It thus makes sense to replace the transition *function* of the automaton by a probabilistic state change, formally being a distribution over the next state, conditional on the input symbol. The respective change to Definition 1 is simple:

Definition 2 (Probabilistic Mealy Automaton). *A* probabilistic Mealy automaton *is a six-tuple* $M = (S, \Sigma_{in}, \Sigma_{out}, \delta, \lambda, s_0)$ *as in Definition 1, except that the transition regime is given by a mapping* $\delta : S \times \Sigma_{in} \to \Delta(S)$*, where* $\Delta(S) := \left\{ \boldsymbol{p} \in [0,1]^{|S|} : \|\boldsymbol{p}\|_1 = 1 \right\}$ *is the simplex, i.e., set of all discrete probability distributions, over the state space* S*. We let* $\mathcal{PMA}$ *denote the set of all probabilistic Mealy automata.*

fridaya@uef.fi

Thus, Definitions 1 and 2 differ by only one token (replacing S by $\Delta(S)$ in the definition of δ), which means that a transition from a state $s \in S$ upon an input symbol $\sigma \in \Sigma_{in}$ may be to any of the other states (including s itself), but happens with a given probability $\delta(s, \sigma) = (\Pr(s'|s, \sigma))_{s' \in S}$. Obviously, by defining δ to output only degenerate distributions, a deterministic Mealy automaton is recovered, thus $\mathcal{MA} \subset \mathcal{PMA}$. The strictness of this inclusion follows from the inclusion of finite automata in $\mathcal{MA}$ and results of Rabin [15].

Definition 3 (Infrastructure Model). *Our infrastructure model is a directed graph $G = (V, E)$ with $V \subset \mathcal{PMA}$ and $E \subseteq V \times V$. Each $M \in V$ corresponds to an asset. An edge (denoted as $M_1 \to M_2$)$\in E$ ($M_1 \to M_2$) means that the output symbols of M_1 become input symbols to M_2. This entails, among others, that all $M \in V$ have the following common characteristics:*

1. *$S = \{1, 2, 3\}$, i.e., an asset can be in one out of three states, corresponding to "not affected" ($s = 1$), "partial damage" ($s = 2$) or "complete failure" ($s = 3$).*
2. *$\Sigma_{in} = \Sigma_{out} =: \Sigma \subset \{0, 1\}^*$ is composed from strings that describe* alarms*. These unambiguously encode at least (among others) the following information:*
   - *alarm kind (type)*
   - *alarm criticality level (on a fixed finite scale)*
   - *time-stamp*

   Thus, alarms are understood by an asset, and it may (but not need to) output another alarm that acts as input symbol to another automaton. Without inducing ambiguities, we can ease our notation by writing Σ to denote the set of all alarm notifications.

The further components of the automata are intentionally left unspecified by Definition 3 and must be instantiated to fit the situation at hand. While it may not be too difficult to list the assets themselves, their description for the infrastructure model is exactly prescribed by the specification of the respective automata. The "to-do list" of artifacts to specify thus comes to:

1. List the assets, each of which will correspond to an automaton $M_1, \ldots, M_k$, each of which takes the form $M = (\{1, 2, 3\}, \Sigma, \Sigma, \delta, \lambda, 1)$, where Σ, δ and λ are defined next. The initial state is for all automata set to 1, corresponding that at the beginning of the simulation, all assets are in the "not affected" state.
2. Specify the possible alarms, by first specifying the properties or information items that they may carry. For example, if there is a set K of different alarm kinds, T is the time-stamp (strings), C is a criticality scale (e.g., $C = 1, 2, 3$), then the alarms are subsets $\Sigma \subset K \times T \times C \times \ldots$, depending on what (further) information is embodied in an alarm string.
   In order to process multiple alarms at the same asset, we shall assume a total ordering on Σ. If Σ can be written as Cartesian product $\Sigma \subset \Sigma_1 \times \Sigma_2 \times \ldots$, then it suffices that each Σ_i is totally ordered, and Σ can be ordered using

fridaya@uef.fi

the lexicographic order based on the orderings of the attributes. For example, alarms may come in sequence of time-stamps; upon identical time-stamp, the ordering is according to the criticality, and so on.

3. Define the transition regime δ: this amounts to pondering how likely an alarm $a \in \Sigma$ is putting the asset from state i into state j for $i, j \in \{1, 2, 3\}$. Essentially, δ can thus be specified as a family of stochastic matrices $\left\{\boldsymbol{P}_a \in [0,1]^{(3\times 3)} : a \in \Sigma\right\}$, one per alarm. If an alarm is not relevant for an asset, then we can use the identity matrix $\boldsymbol{P}_a = \boldsymbol{I}_3$ to represent $\delta(\cdot, a)$ as a function of the state.
4. Define the notification behavior of the asset, i.e., the output function λ: this means thinking about whether a state transition outward from $s \in S$ upon an alarm $a \in \Sigma$ may trigger further alarms b at other assets. If it does not, then we can define $\lambda(s, a) = \varepsilon$, where ε is the empty word, denoting "no further alarm".

   Note that this step naturally induces the dependency structure, as if there is an output of the automaton $M_i \in V$ defined by λ, we should naturally have a target node $M_j \in V$. If so, then we add the edge $M_i \to M_j$ to the set E in our graph dependency model.
5. Collect the entirety of automata in the set V and let the set E collect all dependencies identified during step 4 by specifying the output functions λ per automaton. The tuple $G = (V, E)$ is then the final infrastructure model.

We have to highlight a few remarks on this method. First, it is easy to see that the alphabets defined in step 2 grow exponentially with the amount of information that an alarm may carry, and indeed, this is a natural consequence of how alarm notifications look in reality. The entirety of these, however, may (and in practical cases will) boil down to a feasibly small set of variants for which proper mitigation actions will be known to the CI provider (otherwise, an automated enumeration of possible alarm kinds may dig up notifications that were never thought of, and as such already provides new information to the CI provider).

Second, it is worth pointing out that we can allow state transitions with the empty word as input. Note that neither the transition functions δ nor the output functions λ are required to be total, although they should be in practice. A transition upon the empty word then corresponds to a state change that happens *without* any external force. While this can make sense if we want to model "inner" problems that "just happen", like it would be the case for a Markov chain. In most cases, we will nonetheless specify $\delta(s, \varepsilon) := s$, meaning that the state will remain unchanged unless there is an external event relevant for the asset.

Third, each Mealy automaton can itself be described by a directed multigraph with labeled edges. That is, an automaton M may have multiple transitions between the same two states, in order to enable different forms of alarming other assets upon a transition. For example, if the transition from state 1 ("not affected") to state 2 ("partial damage") is caused by an alarm $a \in \Sigma$ in a (parent) asset M_1, then the automaton may notify others by sending an alarm (symbol)

fridaya@uef.fi

$b \in \Sigma$. If, on the contrary, the transition into the state "partial damage" has a different cause by another parent asset M_2, then the information to other assets may in turn be different, and the output symbol may be $c \in \Sigma$. Hence, there would be two transitions from "not affected" to "partial damage" in this example, being arrows $1 \xrightarrow{a/b} 2$ and $1 \xrightarrow{a/c} 2$, where the arrow annotation p/q denotes the transition triggered by the input symbol p and emitting the symbol q. For probabilistic transitions, we extend the notation as $1 \xrightarrow{p/q/r} 2$, telling that the transition will happen with probability r. Typically, the values r for the transitions would be constants, defined by domain experts. A generalization of the above model lets the probability r of a transition be conditional on the states of certain (relevant) other nodes and even time-dependent. That is, a state transition may not only depend on an external signal but also on the current states of other nodes and time. Formally, we can (generically) specify $r = \Pr(\cdot|N, t)$, where $N \subseteq V$ is a set of neighbors whose state is relevant for the transition, and t is the time-variable, e.g., if the likelihood for a transition into a failure state increases over time unless mitigation actions happen (i.e., backward transitions into the functional state).

Finally, it is easy to imagine that multiple alarms, say a_1, a_2 may simultaneously trigger transitions into distinct states. The transitions can, but do not need to be consistent, in the sense that a notification a_1 may put an asset back from state 2 into functional state 1, while another, concurrent, alarm a_2 could put the same asset from state 2 into failure state 3. The total (e.g., lexicographic) ordering that we assumed on Σ lets us process the alarms in a natural order and avoid issues of simultaneous state transitions.

4 Example

As an example, consider a small (sub-)network consisting of four cyber assets (server A connected to a database, a surveillance application and an authentication application) and 11 physical assets of which five are connected to the cyber network (i.e., video cameras are connected the surveillance application and surveying the LNG area). The whole scenario has a real life background coming from the SAURON project [18] and is depicted in Fig. 1. Assume that each node is in one of, say, three different states that can be interpreted as "complete failure", "partial damage" or "not affected". Figure 1 also depicts how we model a single node (e.g., server A) in this example as probabilistic Mealy automaton 2. In order to demonstrate how our model is able to find cascading effects we model two alarms, fire and a software virus for the provided example. We consider that the fire breaks out in room A and the software virus infects server A. The goal is to predict possible cascading effects when one (or both) of the described incidents happens.

Each node is modeled as probabilistic Mealy automaton (cf. Definition 2) with respect to the two alarms. The dependencies given in Fig. 1 show how we connect the single automatons with each other. In order to resolve the issue

fridaya@uef.fi

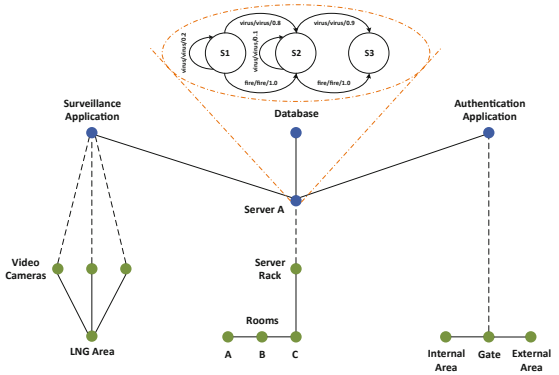


Fig. 1. Illustrative example showing the dependencies between physical and cyber assets and depicting how a vertex is modeled using Definition 2.

that an automaton has two alarms at once to process, we use a Fist-in-First-out (FIFO) queue as input sequence for each automaton. For each incident, we conduct 1000 simulation runs, each run includes at most 10k steps and then we average the obtained results. Figure 2 depicts the cascading effects if we assume a fire breaks out in room A. The nodes are colored from green to dark red, indicating their state at the end of the simulations. The darker a node is colored red, the more often it is in the state "complete failure" with the numbers next to the nodes indicating the percentages according to all simulation runs. We see that a fire starting in room A would affect the whole infrastructure. The node which indirectly connects the LNG area and the access gate is server A. Thus, if the fire arrives at the server room (cf. Fig. 1) and affects the server, all the service hosted by the server are going offline. This is also covered by the result of our simulation, i.e., it captures the cascading effects in this case quite well.

Fig. 2. The possible cascading effect of a fire outbreak in room A.

In contrast to the physical incident, let us assume that server A is infected by a software virus. In our example this cannot have an impact on the rooms and the

fridaya@uef.fi

rack where the server is mounted on. However, services hosted by server A may be directly affected, depending on the security measure taken to isolate services. In this context, we assume that the surveillance application is well protected in a secured container. Figure 3 depicts how the software virus "spreads". The virus has the most severe impact on the server and a weaker impact on the authentication application (which in return controls the physical access to an internal area).

Fig. 3. The possible cascading effect if server A gets affected by a software virus.

Figure 4 depicts a scenario in which both alarms are active at the same time. We notice a significant increase of the probability that cyber assets (and the physical assets that depend on them) are in the state "complete failure" compared to the case where only the fire (cf. Fig. 2) or the software virus (cf. Fig. 3) alarms are active, respectively.

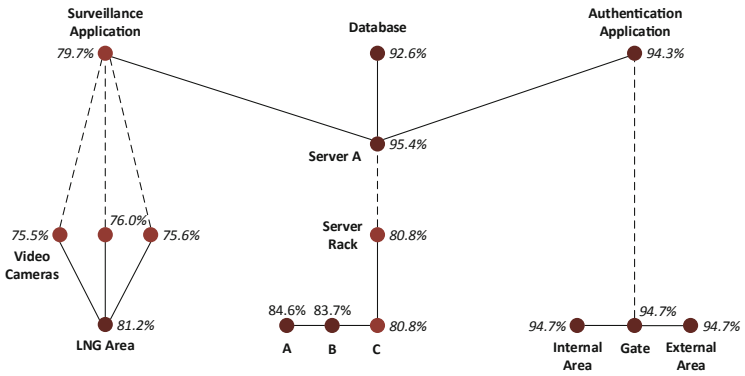


Fig. 4. The cascading effects if both alarms (fire and software virus) are active.

5 Discussion

Our dependency model as described here also includes some models from the literature as special cases. For example, the model can mimic percolations by

fridaya@uef.fi

letting the automata emit and react on a single alarm only with a fixed probability over all automata. This resembles a percolation model, and by allowing for distinct probabilities per transition, also covers a generalized percolation model [7]. Further, our model can also describe Markov chains, although they essentially differ from percolations, since the chain takes on exactly one (out of many possible) next states upon a transition, while the automata can (in generally) notify several related (connected) automata. If the dependency model, i.e., the graph G composed from the automata, is designed such that it admits all unique paths for alarms, then there will be no "branching" of alarms and the model behavior becomes that of a Markov chain. The state of the chain is then the most recent automaton that took action and signalized its neighbors. To enforce the activation of exactly one of them with a prescribed probability (coming from the Markov chain's transition matrix), we can use a specific alarm symbol to address a single automaton and emit this symbol with exactly the probability as prescribed for the chain, by adding the respective transition. Thus, our model can be used to describe both percolations and Markov chains.

In allowing for transitions upon the empty word ε, the model even extends its capabilities to the simulation of concurrent multiple incidents, since with ε-transitions, the automata become non-deterministic and can "move" independently of one another. In its pure form, the model can, to some extent (but not fully), also encode conditions like Petri nets would embody: suppose that an automaton has a transition that conditionally depends on the state of other automata. This is essentially the same as a transition and place in a Petri net. In the way as it is defined, however, the model *cannot* resemble a general Petri net: with the constraint of only three states throughout all automata, we would have a maximum of three markings per place (thinking of the model as a Petri net). Even without this limitation, the number of states would nonetheless need to be fixed and finite a priori, which is not required by a general Petri net.

Due to the complexity of the model, it is somehow difficult to study analytically, e.g., looking for steady state solutions, oscillating effects or similar. The fact that a Markov chain (in being a simplified instance of the model) already induces an exponentially large transition matrix in our setting, somewhat limits the model to simulation primary, and leaves analytical studies to automata and graph theoretic matters. These will be explored along future work.

6 Conclusion

In this article, we presented a novel approach towards modeling dependencies in highly interconnected infrastructures (such as CIs) and simulating the potential cascading effects of an incident. Therefore, we considered coupled Mealy automata, which describe the different operational states an asset can be in and how a state change due to an incident affects the entire asset network. We showed that due to this generic approach, our model is applicable not only in the cyber and in the physical domain, individually, but can also be used to describe effects across those domains.

fridaya@uef.fi

We believe the model to be a reasonable balance between expressiveness and feasibility of construction, since it originates from a practical context of port infrastructures and their highly interconnected physical and cyber assets (see the Acknowledgment), and has its underlying information structure come from discussions with domain experts. The power of the simulation, however, lies in digging up potential effects that the isolated view of an expert does not reveal unless the whole picture including related yet different expertise is taken into account. In this way, the results from the simulation can be used to support security officers in assessing the current situation in their infrastructure, going beyond separate physical and cyber views and towards a Hybrid Situational Awareness.

A more advanced model may also take into account information from social or purely technical parts of a network. The model presented here is generally able to work with more then two different parts and such extensions are subject to future work.

Acknowledgment. This work was supported by the European Commission's Project SAURON (Scalable multidimensional situation awareness solution for protecting European ports) under the HORIZON 2020 Framework (Grant No. 740477).

References

1. Bañuls, V.A., Turoff, M.: Scenario construction via delphi and cross-impact analysis. Technol. Forecast. Soc. Change **78**(9), 1579–1602 (2011). http://www.sciencedirect.com/science/article/pii/S0040162511000667. The Delphi technique: Past, present, and future prospects
2. Carreras, B.A., Newman, D.E., Gradney, P., Lynch, V.E., Dobson, I.: Interdependent risk in interacting infrastructure systems. In: 40th Annual Hawaii International Conference on System Sciences, HICSS 2007, p. 112, January 2007
3. Dobson, I.: Estimating the propagation and extent of cascading line outages from utility data with a branching process. IEEE Trans. Power Syst. **27**(4), 2146–2155 (2012)
4. Gordon, T., Hayward, H.: Initial experiments with the cross impact matrix method of forecasting. Futures **1**(2), 100–116 (1968). http://www.sciencedirect.com/science/article/pii/S0016328768800035
5. Guo, H., Zheng, C., Iu, H.H.C., Fernando, T.: A critical review of cascading failure analysis and modeling of power system. Renew. Sustain. Energy Rev. **80**, 9–22 (2017)
6. König, S., Gouglidis, A.: Random damage in interconnected networks, pp. 185–201. Springer, Cham (2018)
7. König, S., Schauer, S., Rass, S.: A stochastic framework for prediction of malware spreading in heterogeneous networks. In: Proceedings of NordSec Conference 2016 Secure IT Systems, Oulu, Finland, pp. 67–81. Springer (2016)
8. Kotzanikolaou, P., Theoharidou, M., Gritzalis, D.: Cascading effects of common-cause failures in critical infrastructures. In: Butts, J., Shenoi, S. (eds.) Critical Infrastructure Protection VII, pp. 171–182. Springer, Heidelberg (2013)
9. Massaro, E., Bagnoli, F.: Epidemic spreading and risk perception in multiplex networks: a self-organized percolation method. Phys. Rev. E **90**(5), 052817 (2014). https://doi.org/10.1103/PhysRevE.90.052817

fridaya@uef.fi

10. McGee, S., Frittman, J., James Ahn, S., Murray, S.: Implications of cascading effects for the hyogo framework. Int. J. Disaster Resil. Built Environ. **7**, 144–157 (2016)
11. Ouyang, M.: Review on modeling and simulation of interdependent critical infrastructure systems. Reliab. Eng. Syst. Saf. **121**, 43–60 (2014). http://www.sciencedirect.com/science/article/pii/S0951832013002056
12. Pagani, G.A., Aiello, M.: The power grid as a complex network: a survey. Phys. A: Stat. Mech. Appl. **392**(11), 2688–2700 (2013)
13. Paz, A., Rheinboldt, W.: Introduction to Probabilistic Automata. Elsevier Science, Burlington (2014). http://gbv.eblib.com/patron/FullRecord.aspx?p=1901751
14. Qi, J., Dobson, I., Mei, S.: Towards estimating the statistics of simulated cascades of outages with branching processes. IEEE Trans. Power Syst. **28**(3), 3410–3419 (2013)
15. Rabin, M.O.: Probabilistic automata. Inform. Control **6**(3), 230–245 (1963). https://doi.org/10.1016/s0019-9958(63)90290-0
16. Rahnamay-Naeini, M., Hayat, M.M.: Cascading failures in interdependent infrastructures: an interdependent Markov-chain approach. IEEE Trans. Smart Grid, pp. 1997–2006. (2016). https://doi.org/10.1109/tsg.2016.2539823
17. Rahnamay-Naeini, M., Wang, Z., Ghani, N., Mammoli, A., Hayat, M.M.: Stochastic analysis of cascading-failure dynamics in power grids. IEEE Trans. Power Syst., pp. 1767–1779. (2014). https://doi.org/10.1109/tpwrs.2013.2297276
18. SAURON: Scalable multidimensionAl sitUation awaReness sOlution for protectiNg european ports, December 2018. https://www.sauronproject.eu/
19. Troina, A.: Probabilistic Timed Automata for Security Analysis and Design. CreateSpace Independent Publishing Platform, Scotts Valley (2017)
20. Turoff, M.: An alternative approach to cross impact analysis. Technol. Forecast. Soc. Change **3**, 309–339 (1971). http://www.sciencedirect.com/science/article/pii/S0040162571800215
21. Wang, Z., Scaglione, A., Thomas, R.J.: A Markov-transition model for cascading failures in power grids. In: 2012 45th Hawaii International Conference on System Sciences. IEEE, January 2012. https://doi.org/10.1109/hicss.2012.63
22. Wu, S.J., Chu, M.T.: Markov chains with memory, tensor formulation, and the dynamics of power iteration. Appl. Math. Comput. **303**(C), 226–239 (2017). https://doi.org/10.1016/j.amc.2017.01.030

fridaya@uef.fi

Design and Implementation of a Lightweight Encryption Scheme for Wireless Sensor Nodes

Rutuja Salunke[1(✉)], Gaurav Bansod[1], and Praveen Naidu[2]

[1] Pune Institute of Computer Technology, Pune 411043, India
salunkerd96@gmail.com, gaurav249@gmail.com
[2] Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh 520007, India
praveennaidu468@gmail.com

Abstract. For the Internet engineering society, Internet of Things (IoT) is a new prospect as it links embedded devices to the Internet Protocol (IP) based networks. One of the concept in IoT is 6LoWPAN, which enables Internet Protocol (IP) to be applied to the smallest devices. 6LoWPAN architecture is more popular and known for its architecture in the constrained environment. 6LoWPAN is used in many applications like wireless sensor networks (WSNs) where the nodes operating have limited capabilities in terms of power, memory and footprint area. Currently in the 6LoWPAN stack, the encryption algorithm used is AES-128/256 which is more power consumptive to do encryption on any given platform. Moreover, its hardware implementation as IP core will also result in higher Gate Equivalents (GE's). This paper is all about replacing the encryption algorithm- Advanced Encryption Standard (AES-128/256) used in 6LoWPAN with a new lightweight encryption design NUCLEAR, which focuses on making the 6LoWPAN stack more lightweight. The lightweight encryption design can also be implemented as IP core as it results in very fewer gate counts as compared to AES-128/256 or any other lightweight cipher designs. We have proposed a Generalized Feistel Structure (GFS) cipher NUCLEAR which is highly efficient in terms of memory requirements, power dissipation and footprint area. The proposed design NUCLEAR satisfy all the constraints posed by the resource-constrained environment like IoT.

Keywords: Internet of Things · Lightweight encryption · GFS · CLEFIA

1 Introduction

Industry 4.0 which includes cyber-physical systems, the Internet of things (IoT), cloud computing and cognitive computing is the current trend of data exchange in manufacturing technologies. Among which Internet of Things (IoT) has been recorded as one of the six disruptive technologies [1]. Some groups from the Internet Engineering Task Force (IETF), an open standard organization have been working in the Internet area and are collectively named as 6LoWPAN. The IEEE 802.15.4 standard released in 2003 was the biggest factor which led to 6LoWPAN standardization. 6LoWPAN is the key to implement and analyze the environment of Wireless Embedded Internet.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 566–581, 2019.
https://doi.org/10.1007/978-3-030-22868-2_41

fridaya@uef.fi

6LoWPAN uses IPv6 in low-power, processing- limited embedded devices over low-bandwidth wireless networks.

6LoWPAN is connecting more things to the cloud. Low power, IP driven layers and large mesh network support make this technology a possible option for the internet of things (IoT) applications. 6LoWPAN stands for IPv6 over low-power wireless personal area network. The principle idea with 6LoWPAN is that the internet protocol packets can be carried efficiently within small link layer frames. It means that 6LoWPAN should be used with low power wireless network with memory constrained devices. It is adapted and used over a variety of other networking media including Sub1-GHz, low power RF, power line control and low power Wi-Fi. A 6LoWPAN network is normally composed of mini low power nodes in a self-healing mesh form. In the 6LoWPAN, data can be exchanged freely by using an Edge Router which acts as a simple gateway to the rest of the internet. Applications running on the node can pass IP packets to the server on the internet without thinking about the Edge Router which just releases the packets from the 6LoWPAN network to the internet.

There are three distinct advantages of a 6LoWPAN system. Firstly, 6LoWPAN uses Open IP Standards which includes many well-known standards such as TCP, UDP, HTTP, COAP, MQTT, and web-sockets that many developers are already familiar with. The lack of an application layer gateway is another advantage.

Secondly, mesh routing enables networks to become more robust with large size because more nodes can talk to one another. The performance can be optimized for either hop count or link quality and the network is self-healing. Thirdly, multiple physical layers as well as frequency bands are supported by 6LoWPAN. 6LoWPAN can be any low-power wireless node that needs to be connected to the internet. A distinct advantage of 6LoWPAN is that it uses IPv6 and this alone sets it aside from other standards such as Zigbee. 6LoWPAN is very advantageous for low power wireless sensor networks (WSNs) and other forms of low power wireless networks.

2 Related Work

Since the last decade, many ciphers were introduced but none of these ciphers were used to replace the AES-128/256 in 6LoWPAN protocol stack. This is because every cipher had some or the other drawback which was affecting resource constrained environment. CLEFIA [2] which is accepted by ISO/IEC 29192 International Standards is a very efficient GFS cipher as it has a higher throughput of 1.6 Gbps. But at the same time GE's required are 2499 and Flash memory is 4708 bytes. PRESENT [3] on the other hand has very fewer GE's as compared to CLEFIA i.e. 1339 GE's, but power consumption is higher i.e. 38mW. PICCOLO [4], TWINE [5], Camellia [6], BORON [7], RECTANGLE [8], mCrypton [9], PRIDE [10] and PRINCE [11] are some other existing ciphers whose Gate Equivalents and power dissipation are more. Therefore, to make the 6LoWPAN stack more lightweight, a lightweight cipher which is efficient in every design metric has to be designed. In this paper, we have proposed a new lightweight GFS cipher NUCLEAR, which is efficient on all design metrics and is having edge over all existing lightweight cipher designs. The proposed cipher design is

fridaya@uef.fi

capable of making the 6LoWPAN stack more lightweight by replacing AES-128/256 in software stack or as an IP core on any ASIC platform.

3 6LoWPAN and Security

The 6LoWPAN protocol stack is a simple IPv6 protocol stack with 6LoWPAN. It is almost identical to a normal IP stack with some differences as follows: First of all, 6LoWPAN supports only IPv6, therefore a small adaptation layer was introduced between IP stack's link and network layer to optimize IPv6 over IEEE 802.15.4 and enable transmission of IPv6 datagrams over IEEE 802.15.4 radio links. Maximum packet size of the IEEE 802.15.4 frame is 128 bytes. The Physical layer (IEEE 802.15.4 PHY) provides basic capabilities required for the communication in the medium. The MAC PIB (PAN Information Base) controls the security services which are provided by IEEE 802.15.4 MAC. Frame security is the critical function of IEEE 802.15.4 MAC. The Data Link layer enables the reliable, single-hop communication links between 6LoWPAN nodes.

The main component of 6LoWPAN is the adaptation layer. The most important function of this layer is to compress the TCP/IP header. It's second most important function is handling the fragmentation and reassembling of packets. Routing is another function of the adaptation layer. The sensor nodes are capable of internetworking due to the 6LoWPAN network layer. Addressing, mapping and routing protocols are most important concerns of this layer.

As shown in Fig. 1, 6LoWPAN stack uses AES-128/256 as encryption protocol to provide data integrity. This standard protocol is introduced at MAC layer of the software stack.

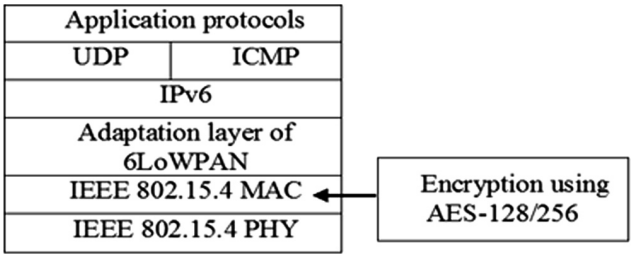


Fig. 1. The 6LoWPAN protocol stack using AES-128/256

The most important and vulnerable part is the security of 6LoWPAN and maintaining it. Unique challenges are possessed by 6LoWPAN and therefore traditional security techniques are not used directly. A few well-publicized security and privacy incidents can seriously damage the future of an emerging technology. Wireless networks are wide open to attackers, who may overhear and inject packets being exchanged, possibly using advanced antenna technology to mount their attacks from well outside the normal range of IEEE 802.15.4 devices. IEEE 802.15.4 nodes can

fridaya@uef.fi

operate in either secure mode or non-secure mode [12]. Two security modes are defined in the specification in order to achieve different security objectives: Access Control List (ACL) mode and secure mode.

Unless the security objectives have been defined that are being claimed to be fulfilled, the system cannot be called a secure system. Security objectives are grouped into three categories:

Confidentiality. It can be achieved by making the information unintelligible by cryptographic encryption.

Integrity. Integrity is achieved by adding cryptographic integrity checks to messages. A related security objective is that of authentication: a message actually is originating from the source that it claims to be. A message integrity check is therefore similar to message authentication.

Availability. Denial of service attacks from sources that cannot be easily controlled is very dangerous.

Without cryptographic support, the protection of messages from reading or detecting a tampered message is not possible. Security objectives that can only be met by protecting the conversation between two end-nodes are therefore best met by performing the cryptography at layer 3.

4 Proposed Scheme

Data link layer (IEEE 802.15.4 MAC) is responsible for the Confidentiality, Integrity, and Authentication of the data. Unauthorized users cannot access the information due to data confidentiality. Alteration of received data by an attacker during the transition is ensured by data integrity. Data Authentication is ensuring that the data is originated from a trusted source in case any attacker adds messages to the original data. This is possible because the IEEE 802.15.4 MAC currently uses AES-128/256 for encryption of the data in layer 2 (Data link layer). The 6LoWPAN stack uses AES (Advanced Encryption Standard) for encryption to make it secure but using AES has many drawbacks in the current scenario where a million numbers of small nodes are communicating over a network, and each node needs to be secured. As mentioned in [12], use of Lightweight ciphers in place of AES-128/256 is recommended to overcome the drawbacks of AES-128/256 and therefore making the 6LoWPAN stack more lightweight. AES-128/256 has more memory requirement, more execution time, and more power dissipation as compared to the existing lightweight cipher designs. Therefore, the use of Lightweight cipher will reduce the execution time and increase the throughput. The 6LoWPAN stack can be made lightweight in two ways: Firstly, in the software stack, by implementing the Lightweight cipher design instead of AES-128/256 as shown in Fig. 2.

Secondly, we can build it as an IP core on any ASIC platform, which helps in parallel execution of the stack and thus increases the speed of execution. This will not only make the process faster but it will also reduce the overhead of “doing encryption” on software stack. The same IP core can be of AES-128/256, but it will result in

fridaya@uef.fi

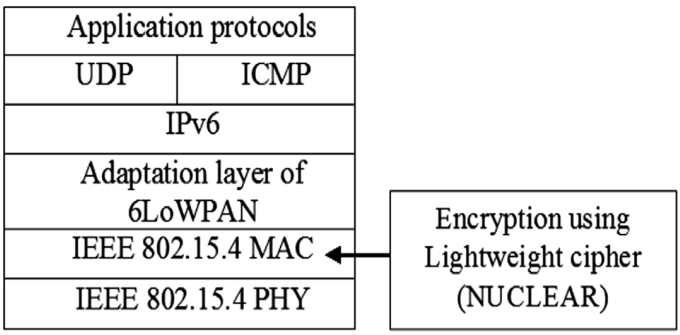


Fig. 2. The 6LoWPAN protocol stack using NUCLEAR

occupying larger area as Gate Equivalents for the AES design is very high around 2500. Definitely, this will make the stack lightweight but it possesses the drawbacks of implementation of AES-128/256 in hardware. Thus, the implementation of IP core using lightweight cipher is more efficient as shown in Fig. 3.

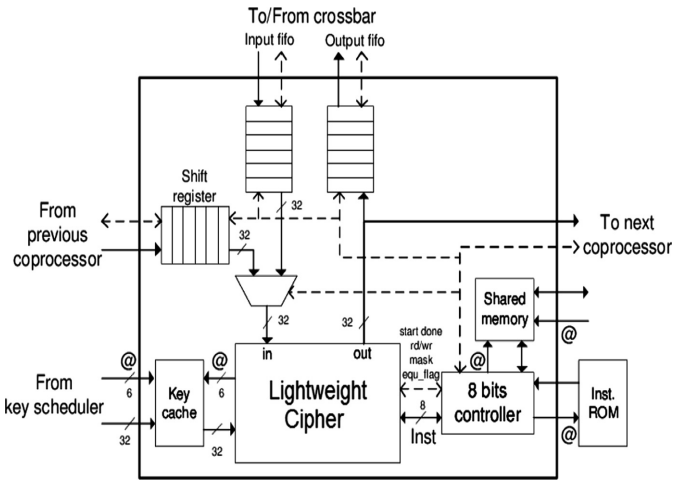


Fig. 3. IP core design of a lightweight encryption cipher.

5 Lightweight Cipher for 6LoWPAN

The 6LoWPAN network consists of the wireless sensor nodes. Each sensor node has its unique IPv6 address. These sensor nodes operate on the 8-bit microcontroller. Low bandwidth, small packet size and a large number of devices are some of the characteristics of 6LoWPAN. IPv6 allows transmission of packets with maximum size of 1280 bytes, which cannot be handled by the IEEE 802.15.4 which has a MTU of 128 bytes. Therefore, a fragmentation scheme is defined in 6LoWPAN which allows such

fridaya@uef.fi

packets. Big packet sizes should be avoided to prevent bad performance in lossy networks. The 6LoWPAN devices are considered as resource-limited. Constraints of 6LoWPAN networks are more computational power, less flash memory, storage and especially they are battery powered. This leaves no space available for the cryptographic protocol. The solution to this problem as mentioned above is to design a lightweight cipher with minimum flash memory which can fit into the 6LoWPAN stack and helps to overcome the constraints of 6LoWPAN. The lightweight cipher designed should result in fewer GE's, less execution time, more throughput and less Flash memory requirements.

6 Cipher Design

In this paper, we have proposed a lightweight GFS cipher which has optimum performance. NUCLEAR cipher is a GFS cipher. The plaintext of 64 bits is divided into four parts. The rationale behind the design of this cipher was to provide a GFS cipher with fewer GEs, less memory requirement, less power consumption and higher throughput. As compared to other GFS ciphers, NUCLEAR is an efficient encryption design with less complexity and resists every basic attack and few advance attacks mounted on the cipher.

We chose to design a GFS cipher because it divides the plaintext into four parts and therefore it is more complex as the attacker has to try every combination of each part of the plaintext. In NUCLEAR, the plaintext is divided into four parts of 16-bits each. Every block is operated through different data paths used in the cipher design as shown in Fig. 4. Two blocks of plaintext are being processed by the S-Box, at two different positions. Three EX-OR's are used with two of them EX-ORing with the key and with the plaintext. P-layer is not used in this cipher because it increases the flash memory requirements. Instead of that, by properly choosing S-Box in the design and proper use of circular shift operators provides the same security as using P-layer. Total 25 rounds are used in the design to produce a cipher text. For 25 rounds, 25 different keys are generated using key scheduling algorithm.

The design has nonlinearity because of use of S-box. The circular shift operators used are 3 and 6 as shown in Fig. 4 because it generates maximum number of active S-Boxes after performing linear and differential cryptanalysis on the cipher. For other circular shifts, less number of active S-Boxes was generated. The robust algorithm was written to find out the best circular shift values which can produce a maximum number of active S-boxes in minimum number of rounds Amongst the GFS ciphers, NUCLEAR is very efficient because it requires only 2028 bytes of flash memory, 171.77 usec of execution time, 372.591 kbps throughput, 1109 GE's and 27 mW of power consumption.

NUCLEAR is a secure GFS cipher because it resists the attacks mounted on it. Attacks like Linear and Differential cryptanalysis, Biclique attack and Avalanche effect were mounted on NUCLEAR.

We have not tested the cipher design against various advance attack as the scope of the paper is to introduce a design suitable for 6LoWPAN architecture, highlight its software and hardware design metrics and to also to prove resistance against popular

fridaya@uef.fi

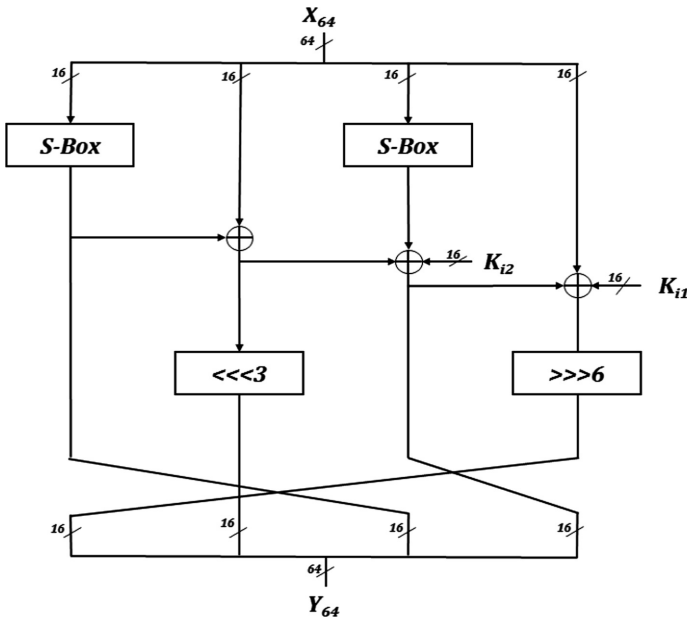


Fig. 4. NUCLEAR block cipher

basic and differential attacks. Cryptographic society can further take this work and test this cipher's resistance against more advance attacks. The NUCLEAR cipher has a nonlinear element in the design that is S-Box. The 4 × 4 S-Box used is given as follows (Table 1).

Table 1. S-box for NUCLEAR

X	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
S(X)	E	4	B	1	7	9	C	A	D	2	0	F	8	5	3	6

S-Box: S-Box is the only nonlinear element used in NUCLEAR cipher and an important component of symmetric key ciphers. It is important to design the S-Box which provides high nonlinearity and low autocorrelation [13] so that the block cipher resists every known and advance attacks. One such approach has been explained in [13]. The S-Box used in NUCLEAR is designed using Linear Approximation Table (LAT) [14] and Difference Distribution Table (DDT) [14]. Using this method we get specific values of CARLC and CARDC for the S- Box. The CARLC value is 4 and CARDC value is 4. These values depict cardinality of linear and distribution table derived from S- Box. The S-Box designed for NUCLEAR is resistant against linear and differential attacks. The S-Box designed and used in NUCLEAR is compact and robust as it resists linear and differential Cryptanalysis.

fridaya@uef.fi

7 Performance Analysis

The NUCLEAR cipher provides optimum performance on both the software as well as on the hardware platform. In this section, we have highlighted the cipher's performance on software and hardware platform.

7.1 Hardware Performance

The hardware performance analysis includes the number of Gate Equivalents (GE's) and architecture of the cipher (data path). In this section, we will discuss the parameters responsible for the hardware performance of the cipher.

Gate Equivalents: Gate Equivalents (GEs) required for NUCLEAR cipher are calculated using ARM 7 standard ASIC library IBM 8RF (0.130 μm). The GEs for various gates are given as follows: NOT 0.75, AND 1.25, OR 1.25, XOR 2.00, 2-1 MUX 2.25, and D flip-flop 4.25. Serialized architecture of NUCLEAR cipher is shown in Fig. 5, which requires fewer GE's than round based architecture. CLEFIA with 128-bit occupies 5,979 gates at a throughput of about 1.60 Gbps. The total number of GE's required for NUCLEAR cipher are 1109 in serialized architecture and 1266 in round based architecture, which is very less compared to other lightweight GFS ciphers. The calculation of GE's required for NUCLEAR is given in Table 2 which includes GE's calculation for the cipher as well as key scheduling (Fig. 6).

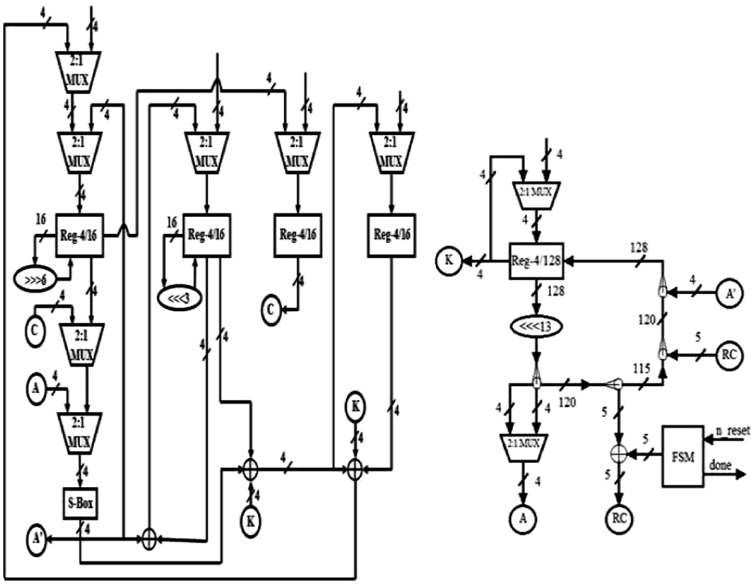


Fig. 5. Serialized architecture of NUCLEAR cipher

fridaya@uef.fi

Table 2. Calculation of GE'S for NUCLEAR

Data layer	GE's	Key layer	GE's
D Reg.	64 × 4.25	D Reg.	128 × 4.25
2:1 MUX	7 × 4 × 2.25	2:1 MUX	2 × 4 × 2.25
XOR	2 × 4 × 5	XOR	5 × 2
SBOX	28	Shift operators	0
Shift operators	0	FSM	134
Total	403	Total	706
Total No. of gates required for 128 bit key = 1109			

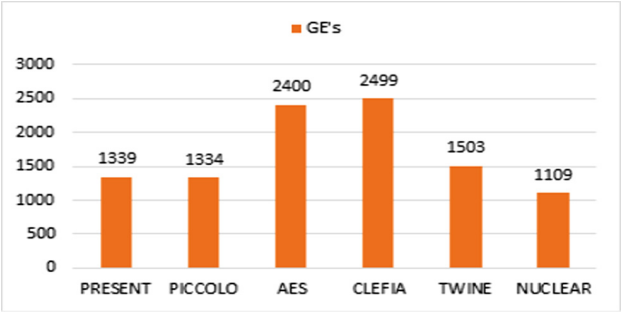


Fig. 6. GE's comparison of NUCLEAR with other existing ciphers

7.2 Software Performance

The software performance analysis includes power consumption, memory requirement, throughput and execution time. In this section, we will discuss the parameters responsible for the software performance of the cipher.

Memory Requirement: Amongst the lightweight GFS ciphers, NUCLEAR is the lightest cipher as it requires very less memory. Software performance of NUCLEAR cipher is tested on ARM 7 LPC2129 platform. The NUCLEAR cipher requires flash memory of 2028 bytes and RAM of 1320 bytes (Fig. 7).

Throughput and Execution Time: In software performance of NUCLEAR cipher, throughput and execution time are calculated at a frequency of 10 MHz on ARM 7 LPC2129 platform. The execution time of NUCLEAR cipher is very less as shown in Table 3. As the execution time is less, the throughput increases. The execution time of NUCLEAR cipher is 171.77 usec. The throughput of NUCLEAR cipher is 372.591 kbps.

Power Consumption: It is one of the most important issues and critical design metrics to be considered while designing a cipher. Every hardware used in the design contributes to the power consumption. The power dissipation by NUCLEAR cipher is 27 mW which is the least amongst the other existing GFS ciphers. The power

fridaya@uef.fi

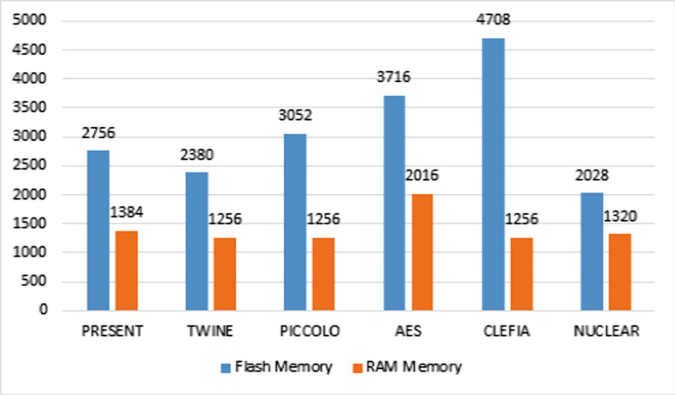


Fig. 7. Flash and RAM memory comparison of NUCLEAR with other existing ciphers

Table 3. Throughput and execution time

Cipher	Type	Block size	Key size	Execution time (u Sec)	Throughput In (Kbps)	No. of cycles
PICCOLO	Feistel	64	128	227.68	281	273
CLEFIA	Feistel	128	128	1048.01	122	12576.12
TWINE	Feistel	64	128	592.87	108	7114.44
NUCLEAR	**GFS**	**64**	**12**	**171.77**	**372.591**	**250**
PRESENT	SPN	64	128	2648.65	24.16	31783.8

consumption of NUCLEAR cipher is calculated with the help of X power tool in the ISE Design suit 14.7 (Xilinx) with 10 MHz frequency. Power consumption of other existing ciphers is also calculated on the same platform. The graph in Fig. 8 shows the comparison of power consumption of NUCLEAR with other existing ciphers.

7.3 Comparison

PRESENT [3] and CLEFIA [2] are block ciphers accepted by ISO/IEC 29192 International Standards. It is necessary to make the 6LoWPAN protocol stack lightweight using lightweight cryptography as mentioned in [12]. The Encryption algorithm (AES-128/256) used in the software stack can be replaced by PRESENT or CLEFIA, as both the ciphers are accepted by ISO/IEC 29192-2:2012 as suitable block ciphers for lightweight cryptography. This will definitely reduce the footprint area and increase the speed of the software stack. But NUCLEAR cipher compared to PRESENT and CLEFIA shows more efficiency and should be used to replace the AES-128/256 in the 6LoWPAN stack. The graph in Fig. 9 shows the comparison between hardware and software performance of PRESENT, CLEFIA, and NUCLEAR.

fridaya@uef.fi

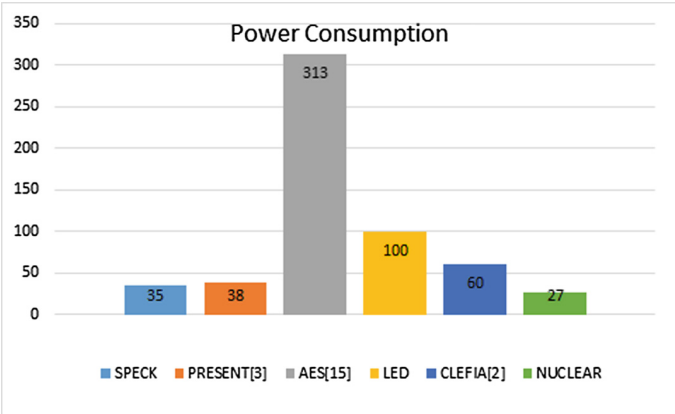


Fig. 8. Comparison of power consumption of NUCLEAR with other ciphers

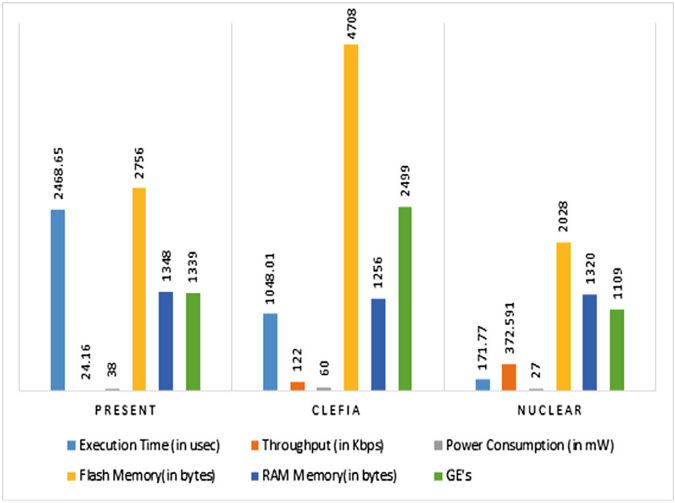


Fig. 9. Comparison of NUCLEAR with PRESENT and CLEFIA

8 Security Analysis

The philosophy behind the design of this cipher is based on offering a very high level of security. Cryptanalysis techniques show the strength of the cipher against the attacks. In this section, we will discuss the security analysis of NUCLEAR against some popular attacks, like Differential Cryptanalysis, Linear Cryptanalysis, Biclique Attack and Avalanche Effect.

fridaya@uef.fi

8.1 Linear Cryptanalysis

Linear Cryptanalysis is one of the two most widely used attacks on block ciphers. Matsui introduced [16] and successfully used linear cryptanalysis as a practical attack on Data Encryption Standard (DES) [17]. Linear Cryptanalysis is also called as known plaintext attack. In this cipher design, S-Box is the only nonlinear element. The Non-linear layer is tested by LAT (Linear Approximation Table). For Linear Cryptanalysis, the criterion for secure design is given by Matsui's Pilling-up lemma. For a cipher structure to resist the linear attack, the cipher should satisfy the condition, that it should have a maximum minimum number of active S-Boxes. Using Piling-Up Lemma, we can find out whether the cipher can resist the known plaintext attack or not. Table 4 shows the number of minimum active S-Boxes present for first four rounds of NUCLEAR cipher during linear cryptanalysis [15].

Table 4. Minimum number of active S-boxes for linear trails

Rounds	Minimum number of active S-boxes
1	0
2	1
3	3
4	8

In NUCLEAR, maximum bias for S-Box is 2^{-2}, which is calculated using LAT. For round 4, the total numbers of minimum active S-Boxes are 8. As per the Piling-Up Lemma principle the total bias for round 4 is:

$$= 2^{(\text{No. of active S}-\text{Boxes}-1)} \times (\text{Max. Bias})^{(\text{No. of Active S - Boxes})}$$
$$= 2^{(8-1)} \times (2^{-2})^{(8)}$$
$$= 2^{-9}$$

To find the total number of minimum active S-Boxes for NUCLEAR cipher after round 20, consider 20 rounds and calculate the minimum number of active S–Boxes. This is calculated by using the piling up lemma principle [16, 19]:

For 20 rounds the bias will be:

$$\varepsilon_{20} = 2^{(\text{Multiplier}-1)} \times (\text{total bias for round 4})^{(\text{Multiplier})}$$
$$\varepsilon_{20} = 2^{(5-1)} \times (2^{-9})^5 = 2^{-41}$$

The complexity of linear attack can be given by the formula NL = $1/(\varepsilon)^2$, so for NUCLEAR, the linear complexity is given as below [16, 19]:

$$N_L = 1/(\varepsilon)^2 = 1/\left(2^{-41}\right)^2$$
$$NL = 2^{82}$$

fridaya@uef.fi

As the linear complexity of NUCLEAR is greater than the defined limit i.e., 2^{64} (64 is the block size), we conclude that NUCLEAR cipher provides high resistance to linear cryptanalysis.

8.2 Differential Cryptanalysis

Differential Cryptanalysis is another most widely used attack on block ciphers. Biham and Shamir presented the differential cryptanalysis first at CRYPTO '90. Differential Cryptanalysis is called as chosen plaintext attack. The Non-linear layer is tested by DDT (Difference Distribution Table). For a cipher structure to resist the differential attack, the cipher should satisfy the condition, that it should have a maximum minimum number of active S-Boxes. Table 5 shows the number of minimum active S-Boxes present for first four rounds of NUCLEAR during differential cryptanalysis [18].

Table 5. Minimum number of active S-boxes for differential trails

Rounds	Minimum number of active S-boxes
1	0
2	1
3	3
4	7

The maximum bias is calculated by using DDT and in NUCLEAR cipher, the maximum bias is 2^{-2}.

The differential characteristic probability is given by P_d:

$$P_d = \left(2^{-2}\right)^{\text{No. of active S - Box}}$$

In NUCLEAR cipher, for round 4, minimum number of active S-Boxes are 7, therefore for 20 rounds, 35 minimum active S-Boxes are present. The total differential probability P_d as per lemmas [16, 20] is given as:

$$P_d = \left(2^{-2}\right)35 = 2^{-70}$$

The complexity of the differential attack can be calculated as $N_d = c/P_d$. Where N_d is the number of chosen plaintext pairs, $c = 1$ and $P_d = 2^{-70}$.

$$N_d = 1/2^{-70} = 2^{70}.$$

As the differential complexity of NUCLEAR is greater than the defined limit i.e., 2^{64}, we conclude that NUCLEAR provides sufficient security against differential cryptanalysis.

fridaya@uef.fi

8.3 Biclique Attack

Biclique attack is a well-known single-key attack. It is based on Meet-in-the-Middle (MITM) attack. Biclique attack is mounted on last four rounds of NUCLEAR. It is used to determine the data complexity of NUCLEAR. The dimension of the key used in biclique attack can be 2, 3 or 4 bits. For NUCLEAR, the dimension of the key is 4 bits. In biclique attack, the red key is used for forward computation and the blue key is used for backward computation. During forward and backward computation the blue key and red key should not coincide with each other. Also, both the computation should not start from same round. NUCLEAR satisfies both the conditions. As a result of biclique attack, the data complexity of NUCLEAR is 2^{60}.

MITM attack includes forward and backward re- computation. In the MITM attack, same keys which are used in biclique attack are used for forward and backward re-computation. Four-dimensional red and blue keys are used in MITM attack. It is used to calculate the total computational complexity of the cipher. The total computational complexity of any cipher should be greater than $2^{127.50}$, so that the cipher is secure. The total computational complexity of NUCLEAR is given as follows:

$$\begin{aligned} C_{total} &= 2^{k-2d}(C_{biclique} + C_{precomp} + C_{recomp} + C_{falsepos}) \\ &= 2^{127.824} \end{aligned}$$

Where,

Biclique computational complexity = $C_{biclique} = 2^{2.356}$
Pre-computational complexity = $C_{precomp} = 2^{3.748}$
Re-computational complexity = $C_{recomp} = 2^{7.585}$
Falsepos complexity = $C_{falsepos} = 2^4$
Number of key bits used for key scheduling = k = 128
Dimension of key selected = 4

The NUCLEAR cipher is secure with a computational complexity of $2^{127.824}$ and data complexity of 2^{60}.

8.4 Avalanche Effect

Avalanche effect is the most important property of any block cipher. In avalanche effect, if the input is changed by single bit then there should be a significant change in the output. The change should be at least 50% according to the strict avalanche criterion (SAC) and make the cipher secure. The cipher exhibiting avalanche effect to a significant degree has high randomization, which prevents the hacker from making predictions about the input if provided with the output. The Avalanche effect for NUCLEAR is given in Table 6.

fridaya@uef.fi

Table 6. Avalanche effect

Plaintext	0000 0000 0000 0000	No. of bits change
Key cipher text	0000 0000 0000 0000 0000 0000 0000 0000 eb8d 67b1 b4c1 3719	
Key cipher text	1000 0000 0000 0000 0000 0000 0000 0000 f492 1fef fd4f d403	35
Key cipher text	0000 0000 0000 0000 0001 0000 0000 0000 04a3 e5e5 0fa0 8962	37
Keeping the key constant and changing a bit in plaintext		
Key	0000 0000 0000 0000 0000 0000 0000 0000	
Plaintext cipher text	0000 0000 0000 0000 eb8d 67b1 b4c1 3719	
Plaintext cipher text	0000 0000 0000 0001 91e0 d24c 4971 ca02	43
Plaintext cipher text	0000 0000 0040 0000 6c44 e618 4a4a 57f6	34

9 Conclusion

In this paper, we have presented a new methodology to make the 6LoWPAN stack more lightweight. As the software stack uses AES-128/256 as the encryption algorithm more power is consumed and more footprint area is required. Therefore, to reduce power consumption and footprint area, AES-128/256 is replaced by Lightweight encryption design (NUCLEAR). This makes the software stack more lightweight and efficient. Another way to make 6LoWPAN stack lightweight is to implement the stack as an IP core on any ASIC platform and henceforth the speed of execution of the stack also increases. NUCLEAR is a compact cipher design which has low power consumption and lesser footprint area than the other existing GFS ciphers. It requires 55% lesser GE's, 45% less FLASH Memory and 34% less RAM Memory than AES-128/256. The NUCLEAR cipher has a robust S-Box and accurate shift parameters which makes it more secure and resists basic and advanced attacks mounted on it. For applications like IoT, Wireless sensor nodes where 6LoWPAN is used, where less power consumption, less execution time and GE's play a crucial role, we believe that NUCLEAR is the best GFS designed to be used in place of AES-128/256 in the software stack.

References

1. SRI Consulting Business Intelligence: Disruptive civil technologies: six technologies with potential impacts on US interests out to 2025. https://fas.org/irp/nic
2. Shirai, T., Shibutani, K., Akishita, T., Moriai S., Iwata, T.: The 128-bit blockcipher CLEFIA. In: FSE, LNCS 4953, pp. 181–195. Springer (2007)

fridaya@uef.fi

3. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: an ultra-lightweight block cipher. In: Paillier, P., Verbauwhede, I. (eds.) Cryptographic Hardware and Embedded Systems—CHES 2007. LNCS, vol. 4727, pp. 450–466. Springer, Berlin (2007)
4. Shibutani, K., Isobe, T., Hiwatari, H., Mitsuda, A., Akishita, T., Shirai, T.: PICCOLO: An ultra-lightweight blockcipher. In: Cryptographic Hardware and Embedded Systems—CHES 2011. LNCS, vol. 6917, pp. 342–357. Springer, Nara (2011)
5. Suzuki, T., Minematsu, K., Morioka, S., Kobayashi, E.: TWINE: a lightweight, versatile block cipher. In: Knudsen, L.R.,Wu, H. (eds.) Selected Areas in Cryptography. Lecture Notes in Computer Science, vol. 7707, pp. 339–354. Springer, Windsor (2012)
6. Aoki, K., Ichikawa, T., Kanda, M., Matsui, M., Moriai, S., Nakajima, J., Tokita, T.: Camellia: a 128-bit block cipher suitable for multiple platforms. In: *SAC '00*, LNCS 2012, pp. 41–54, Springer (2001)
7. Bansod, G., Pisharoty, N., Patil, A: BORON: an ultra lightweight and low power encryption design for pervasive computing. Frontiers (2016)
8. Zhang, W., Bao, Z., Lin, D., Rijmen, V., Yang, B., Verbauwhede, I.: RECTANGLE: a bit-slice ultra-lightweight block cipher suitable for multiple platforms. Cryptology ePrint Archive, Report 2014/084 (2014)
9. Lim, C.H., Korkishko, T.: mCrypton – a lightweight block cipher for security of low-cost RFID tags and sensors. In: Song, J.S., Kwon, T., Yung, M. (eds.) Information Security Applications. WISA 2005. Lecture Notes in Computer Science, vol. 3786. Springer, Berlin (2006)
10. Albrecht, M.R., Driessen, B., Kavun, E.B., Leander, G., Paar, C., Yalçın, T.: Block ciphers – focus on the linear layer (feat. PRIDE). In: Garay, J.A., Gennaro, R. (eds.) Advances in Cryptology – CRYPTO 2014. CRYPTO 2014. Lecture Notes in Computer Science, vol. 8616. Springer, Berlin (2014)
11. Borghoff, J., et al.: PRINCE – a low-latency block cipher for pervasive computing applications. In: Wang, X., Sako, K. (eds.) Advances in Cryptology – ASIACRYPT 2012. ASIACRYPT 2012. Lecture Notes in Computer Science, vol. 7658. Springer, Berlin (2012)
12. Park, S., Kim, K., Hadded, W., Chakrabarti, S., Laganier, J.: 6LoWPAN: security analysis (2011)
13. Clark, J.A., Jacob, J.L., Stepney, S.: The design of S-boxes by simulated annealing. In: Proceedings of the 2004 Congress on Evolutionary Computation (IEEE Cat. No.04TH8753), vol. 2, pp. 1533–1537 (2004)
14. Heys, H.M.: A tutorial on linear and differential cryptanalysis. Cryptologia **26**, 189–221 (2002)
15. Oliver, J.P., Acle, J.P., Boemo, E.: Power estimations vs. power measurements in Spartan-6 devices. In: 2014 IX Southern Conference on Programmable Logic (SPL), pp. 1–5 (2014)
16. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseth, T. (ed.) Advances in Cryptology, Proceedings of the Eurocrypt'93. LNCS 765, pp. 386–397. Springer (1994)
17. Coppersmith, D.: The data encryption standard (DES) and its strength against attacks. IBM Thomas J Watson Research Center technical report RC 18613 (81421), 22 December 1992
18. Biham, E.: New types of cryptanalytic attacks using related keys. J. Cryptol. **7**(4), 229–246 (1994)
19. Chabaud, F., Vaudenay, S.: Links between differential and linear cryptanalysis. In: Advances in Cryptology, EUROCRYPT'94. Lecture Notes in Computer Science, vol. 950, pp. 356–365. Springer, New York (1995)
20. Biham, E., Shamir, A.: Differential Cryptanalysis of the Data Encryption Standard. Springer, New York (1993)

fridaya@uef.fi

Deep Random Based Key Exchange Protocol Resisting Unlimited MITM

Thibault de Valroger(✉)

Paris, France

Abstract. We present a protocol enabling two legitimate partners sharing an initial secret to mutually authenticate and to exchange an encryption session key. The opponent is an active Man In The Middle (MITM) with unlimited computing and storage capacities. The resistance to such an unlimited MITM is obtained through the combined use of Deep Random secrecy, introduced in former work and proven as unconditionally secure against passive opponent for key exchange, and universal hash techniques. We prove the resistance to MITM interception attacks, and show that (i) upon successful completion, the protocol leaks as few residual information as desired about the current value of the shared secret to the opponent, and (ii) that any interception attempt results into unsuccessful completion and is detectable by the legitimate partners. We also discuss implementation techniques.

Keywords: Deep random · Zero knowledge authentication · Man In The Middle Attack · Secret key agreement · Key exchange protocol · Unconditional security · Quantum resistant

1 Introduction and Summary of Former Work

Modern cryptography mostly relies on mathematical problems commonly trusted as very difficult to solve, such as large integer factorization or discrete logarithm, belonging to complexity theory. No certainty exists on the actual difficulty of those problems. Some other methods, rather based on information theory, have been developed since early 90s. Those methods relies on hypothesis about the opponent (such as "memory bounded" adversary [6]) or about the communication channel (such as "independent noisy channels" [5]); unfortunately, if their perfect secrecy have been proven under given hypothesis, none of those hypothesis are easy to ensure in practice. At last, some other methods based on physical theories like quantum indetermination [3] have been described and experimented, but they remain complex to implement, with limitations of distance and specific infrastructures.

It has been proposed in [9] a new cryptographic technique, where Information Theoretical security against unlimited opponent can be reached. The security model relies on the fact that the passive opponent does not have the knowledge of the probability distributions used independently by each legitimate partner to generate their private information. This lack of knowledge has been translated into a mathematically defined opponent's limitation, namely, Deep Random Assumption that we will remind

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 582–596, 2019.
https://doi.org/10.1007/978-3-030-22868-2_42

fridaya@uef.fi

in this paper. Furthermore, while only passive unlimited opponents were considered in [9], we consider in this work active unlimited MITM opponents.

In our model of security, the legitimate partners of the protocol are using Deep Random generation to generate their shared encryption key, and the behavior of the opponent, when inferring from public information, is governed by Deep Random assumption. The legitimate partners have an initial shared authentication secret that is used only for authentication purpose, not for generating the shared encryption key.

Our protocol is original because it resists to opponents that are both unlimited and active MITM. A scheme like Wigner-Carter Authentication [11] is secured against unlimited opponent but not in a MITM scenario. WCA can however be secure against MITM in the specific case of QKD because of the non-cloning theorem. Public Key schemes can be resistant to MITM opponent but not if they are unlimited. Methods like tamper detection by latency examination [12] are not robust enough and not resistant to unlimited opponents.

1.1 Back on the Deep Random Assumption

Before presenting the Deep Random assumption, it is needed to come back on Prior probability theory.

If we denote $\Im_<$ the set of all prior information available to observer regarding the probability distribution of a certain random variable X ('prior' meaning before having observed any experiment of that variable), and $\Im_>$ any public information available regarding an experiment of X, it is then possible to define the set of possible distributions that are compatible with the information $\Im \triangleq \Im_< \cup \Im_>$ regarding an experiment of X; we denote this set of possible distributions as:

$$D_{\Im}$$

The goal of Prior probability theory [7, 8, 10] is to provide tools enabling to make rigorous inference reasoning in a context of partial knowledge of probability distributions. A key idea for that purpose is to consider groups of transformation, applicable to the sample space of a random variable X, that do not change the global perception of the observer. In other words, for any transformation τ of such group, the observer has no information enabling him to privilege $\varphi_{\Im}(v) \triangleq P(X = v|\Im)$ rather than $\varphi_{\Im} \circ \tau(v) = P(X = \tau(v)|\Im)$ as the actual conditional distribution. This idea has been developed by Jaynes [7].

We will consider only finite groups of transformation, because one manipulates only discrete and bounded objects in digital communications. We define the acceptable groups G as the ones fulfilling the two conditions below:

$(C1)$ Stability - Should $\varphi_{\Im}$ be any distribution in $D_{\Im}$, and should τ be any transformation in G, then $\varphi_{\Im} \circ \tau \in D_{\Im}$

$(C2)$ Convexity - Any distribution that is invariant by action of G does belong to $D_{\Im}$

fridaya@uef.fi

It can be noted that the set of distributions that are invariant by action of G is exactly:

$$R_{\Im}(G) \triangleq \left\{ \frac{1}{|G|} \sum_{\tau \in G} \varphi_{\Im} \circ \tau | \forall \varphi_{\Im} \in D_{\Im} \right\}$$

The set of acceptable groups as defined above is denoted:

$$\Gamma_{\Im}$$

For any group G of transformations applying on the sample space of X, we denote by $\Omega_{\Im}(G)$ the set of all possible conditional expectations when the distribution of X courses $R_{\Im}(G)$. In other words ($E[\cdot|\cdot]$ designating the conditional expectation, and $Z(\Im)$ designating the strategy of the opponent to evaluate X knowing $\Im$):

$$\Omega_{\Im}(G) \triangleq \left\{ Z(\Im) \triangleq E[X|\Im] | \forall \varphi_{\Im} \in R_{\Im}(G) \right\}$$

Or also:

$$\Omega_{\Im}(G) = \left\{ Z(\Im) = \int_F v \varphi_{\Im}(v) dv | \forall \varphi_{\Im} \in R_{\Im}(G) \right\}$$

The **Deep Random assumption** prescribes that, if $G \in \Gamma_{\Im}$, the strategy Z_ξ of the opponent observer ξ, in order to estimate X from the public information $\Im$, should be chosen by the opponent observer ξ within the restricted set of strategies:

$$\boldsymbol{Z_\xi \in \Omega_{\Im}(G)} \quad \text{(A)}$$

The Deep Random assumption can thus be seen as a way to restrict the possibilities of ξ to choose his strategy in order to estimate the private information X from his knowledge of the public information $\Im$. It is a fully reasonable assumption because the assigned prior distribution should remain stable by action of a transformation that let the distribution uncertainty unchanged. It enables to circumvent Shannon's impossibility result [1], because that impossibility result makes the assumption that the opponent knows the conditional probabilities for the secret information, and can thus perform the optimal Bayesian inference.

(A) suggests of course that Z_ξ should eventually be picked in $\bigcap_{G \in \Gamma_{\Im}} \Omega_{\Im}(G)$, but it is enough for our purpose to find at least one group of transformation with which one can apply efficiently the Deep Random assumption to the a protocol in order to measure an advantage distilled by the legitimate partners compared to the opponent.

1.2 Back on the Protocol Proposed in [9]

The following protocol has been presented in [9]. In order to shortly remind the notations, the sample space of the distribution of the private information (for A or B) is

fridaya@uef.fi

$[0,1]^n$. Considering $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ some parameter vectors in $[0,1]^n$ and $i = (i_1, \ldots, i_n)$ and $j = (j_1, \ldots, j_n)$ some Bernoulli experiment vectors in $\{0,1\}^n$, $\mathfrak{S}_n$ the symmetric group of permutations of n elements, we denote:

$x.y$ (resp. $i.j$) the scalar product of x and y (resp. i and j)

$$|x| \triangleq \sum_{s=1}^{n} x_s;\ |i| \triangleq \sum_{s=1}^{n} i_s$$

$\forall \sigma \in \mathfrak{S}_n$, $\sigma(x)$ represents $\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right)$

$\frac{x}{k}$ represents $\left(\frac{x_1}{k}, \ldots, \frac{x_n}{k}\right)$ for $k \in \mathbb{R}_+^*$

In that protocol, besides being hidden to any third party (opponent or partner), the probability distribution used by each legitimate partner also needs to have a specific property in order to prevent the opponent to takes advantage of the symmetry resulting from the independence between the choices of the legitimate partners for their respective distributions.

The specific property is that each probability distribution Φ (for A or B) must be "far" from its symmetric projection $\overline{\Phi}(x) \triangleq \frac{1}{n!}\sum_{\sigma \in \mathfrak{S}_n} \Phi \circ \sigma(x)$.

The technical details explaining those constraints are presented in [9]. The set of compliant distributions is denoted $\zeta(\alpha)$ where α is a parameter that measures the "remoteness" of a distribution from its symmetric projection.

For such a distribution Φ in $\zeta(\alpha)$, a tidying permutation, denoted σ_Φ, is a permutation that realizes the minimum

$$\sum_{u,v \in I_0 \times \bar{I}_0} M_{\Phi \circ \sigma_\Phi}(u,v) = \min_{\sigma \in \mathfrak{S}_n} \left(\sum_{u,v \in I_0 \times \bar{I}_0} M_{\Phi \circ \sigma}(u,v) \right)$$

where $I_0 \triangleq \{1, \ldots, n/2\}$. Minimizing the above criteria is an efficient way to ensure that $\Phi \circ \sigma_\Phi$ is driven away from its symmetric projection. Again, details are presented in [9].

Here are the steps of the proposed protocol:

A and B are the legitimate partners. The steps of the protocol $\mathcal{P}(\alpha, n, k, L)$ are the followings:

Step 1 – *Deep Random Generation*: A and B pick independently the respective probability distributions Φ and $\Phi' \in \zeta(\alpha)$, so that Φ (resp. Φ') is secret (under Deep Random assumption) for any observer other than A (resp. B) beholding all the published information. A draws the parameter vector $\mathrm{x} \in [0,1]^{\mathrm{n}}$ from Φ. B draws the parameter vector $\mathrm{y} \in [0,1]^{\mathrm{n}}$ from Φ'.

Step 2 – *Degradation*: A generates a Bernoulli experiment vectors $\mathrm{i} \in \{0,1\}^{\mathrm{n}}$ from the parameter vector $\frac{\mathrm{x}}{\mathrm{k}}$. A publishes i. B generates a Bernoulli experiment vectors $\mathrm{j} \in \{0,1\}^{\mathrm{n}}$ from the parameter vector $\frac{\mathrm{y}}{\mathrm{k}}$. B publishes j.

Step 3 – *Dispersion*: A and B also pick respectively a second probability distribution Ψ and $\Psi' \in \zeta(\alpha)$ such that it is also secret (under Deep Random assumption) for any observer other than A (resp. B). Ψ is selected also such that $\int_{|\mathrm{x}| \in [\mathrm{k}|\mathrm{i}| - \sqrt{\mathrm{n}}, \mathrm{k}|\mathrm{i}| + \sqrt{\mathrm{n}}]} \Psi(\mathrm{x})\mathrm{dx} \geq \frac{1}{2\sqrt{\mathrm{n}}}$ in order to ensure that $|\mathrm{i}|$ is not an unlikely value for

fridaya@uef.fi

$\approx \left|\frac{x}{k}\right|$ (same for ψ' by replacing x by y and i by j). Ψ (resp. Ψ') is used to scramble the publication of the tidying permutation of A (resp. B). A (resp. B) computes a permutation $\sigma_d[i]$ (resp. $\sigma'_d[j]$) representing the reverse of the most likely tidying permutation on Ψ (resp. Ψ') to produce i (resp. j). In other words, with i, $\sigma_d[i]$ realizes:

$$\max_{\sigma\in\mathfrak{S}_n}\int_x P(i|x)\Psi\circ\sigma_\Psi\circ\sigma^{-1}(x)dx$$

Then A (resp. B) draws a boolean $b \in \{0,1\}$ (resp. b') and publishes in a random order $(\mu_1,\mu_2) = t^b(\sigma_d[i],\sigma_\Phi)$, (resp. $(\mu'_1,\mu'_2) = t^{b'}(\sigma'_d[j],\sigma_{\Phi'})$) where t represents the transposition of elements in a couple.

Step 4 – *Synchronization*: A (resp. B) chooses randomly σ_A (resp. σ_B) among (μ'_1,μ'_2) (resp. (μ_1,μ_2)).

Step 5 – *Advantage Distillation*: A computes $V_A = \frac{\sigma_\Phi^{-1}(x).\sigma_A^{-1}(j)}{n}$, B computes $V_B = \frac{\sigma_B^{-1}(i).\sigma_{\Phi'}^{-1}(y)}{n}$. V_A and V_B are then transformed respectively by A and B in binary outputs $\widetilde{e_A}$ and $\widetilde{e_B}$ thanks to the sampling method:

$$\widetilde{e_A} = \left\lfloor\frac{(V_A-\rho)\sqrt{nk}}{K}\right\rfloor \bmod 2, \quad \widetilde{e_B} = \left\lfloor\frac{(V_B-\rho)\sqrt{nk}}{K}\right\rfloor \bmod 2$$

(where ρ is a public translation parameter randomly picked in $\left[0,\frac{2K}{\sqrt{nk}}\right[$ at each instance of the protocol). The multiplicative factor K is chosen such that:

$$\frac{1}{\sqrt{nk}} \ll \frac{K}{\sqrt{nk}} \ll \frac{1}{k}$$

We introduce the following notations corresponding to the canonical form of the opponent strategy to estimate V_A (obtained in the analysis of the opponent strategy under Deep Random assumption in [9]):

$$V_\xi(i,j,\sigma_{\xi,A},\sigma_{\xi,B}) \triangleq \frac{2k\left(\left(\sigma_{\xi,B}^{-1}(i)\cdot I_0\right)\left(\sigma_{\xi,A}^{-1}(j)\cdot I_0\right)+\left(\sigma_{\xi,B}^{-1}(i)\cdot \overline{I_0}\right)\left(\sigma_{\xi,A}'^{-1}(j)\cdot \overline{I_0}\right)\right)}{n^2}$$

$$\widetilde{e_\xi}(\sigma_{\xi,A},\sigma_{\xi,B}) \triangleq \left\lfloor\frac{(V_\xi(i,j,\sigma_{\xi,A},\sigma_{\xi,B})-\rho)\sqrt{nk}}{K}\right\rfloor \bmod 2$$

$$T_m = \left\{\widetilde{e_\xi}(\mu'_1,\mu_1)_m, \widetilde{e_\xi}(\mu'_1,\mu_2)_m, \widetilde{e_\xi}(\mu'_2,\mu_1)_m, \widetilde{e_\xi}(\mu'_2,\mu_2)_m\right\}$$

where $\sigma_{\xi,B}$ is the permutation chosen by ξ among (μ_1,μ_2) as σ_Φ and $\sigma_{\xi,A}$ is the permutation chosen by ξ among (μ'_1,μ'_2) as σ_A. The partners discard then the instance of the protocol if $|T_m| \neq 2$.

fridaya@uef.fi

Step 6: Classical Information Reconciliation [2] and Privacy Amplification [4] (IRPA) techniques then lead to get accuracy as close as desired from perfection between estimations of legitimate partners, and knowledge as close as desired from zero by any unlimitedly powered opponent, as shown in [4].

The choices of the parameters (α, n, k, λ) are theoretically discussed in proof of main Theorem in [9]. They are set to make Steps 5 and 6 possible.

The protocol can be heuristically analyzed as follows:

Regarding the legitimate partners:

(1) when B picks $\sigma_B = \sigma_\Phi$ and A picks $\sigma_A = \sigma_{\Phi'}$ (1/4 of cases), the choice of σ_A and σ_B remain independant from i, j, so that i and j remain draws of independent Bernoulli random variables, then allowing to apply Chernoff-style bounds with accuracy $O\left(\frac{1}{\sqrt{nk}}\right)$ $\left(E\left[\left(V_{A|\sigma_A=\sigma_{\Phi'}} - V_{B|\sigma_B=\sigma_\Phi}\right)^2\right]^{1/2} = O\left(\frac{1}{\sqrt{nk}}\right)\right.$ as shown in [9]).

(2) When A picks $\sigma_A = \sigma'_d[j]$ and/or B picks $\sigma_B = \sigma_d[i]$ (3/4 of cases), the Chernoff bound no longer applies and instead, and V_B or V_A become erratic, which will lead to an error probability of $\approx 1/2$.

On the other hand, the discarding condition $|T_m| = 2$ forced at step 5 imposes that the opponent has no better choice than to choose its bit $\widetilde{e_\xi}$ randomly within $\{0, 1\}$ with 50% chance. Thus, as soon as we can prove that condition $|T_m| = 2$ and condition (1) above coexist with a fair probability $\not\approx 0$, we have shown that an Advantage is created for the partners over the opponent. This fact is proven in [9] Theorem 1.

Below we explain the role of each step:

The Degradation transformations $x \mapsto \frac{x}{k}$ and $y \mapsto \frac{y}{k}$ with $k > 1$ at step 2 are necessary so that $\sigma_d[i](I_0) \neq \sigma_\Phi(I_0)$ (resp. $\sigma'_d[i](I_0) \neq \sigma_{\Phi'}(I_0)$). If $k = 1$, then $i = x$ and $j = y$ and this would cause $\sigma_d[i](I_0) \approx \sigma_\Phi(I_0)$ (resp. $\sigma'_d[i](I_0) \approx \sigma_{\Phi'}(I_0)$).

The Deep Random Generation at step 1 prevents the use of Bayesian inference based on the knowledge of the probability distribution. In particular, if Φ, Ψ (resp. Φ', Ψ') were not managed by Deep Randomness, ξ would be able discriminate σ_Φ among (μ_1, μ_2) (resp. $\sigma_{\Phi'}$ among $\left(\mu'_1, \mu'_2\right)$) by Bayesian inference.

Dispersion step 3 mixes σ_Φ within (μ_1, μ_2) with another permutation $\sigma_d[i]$ (and $\sigma_{\Phi'}$ within $\left(\mu'_1, \mu'_2\right)$ with another permutation $\sigma'_d[j]$). i is entirely determined by $|i|$ and a permutation, which explains the constraint and transformation applied on Ψ in step 3 to make σ_Φ and $\sigma_d[i]$ indisguishable knowing i (same with $\sigma'_d[j]$, $\sigma_{\Phi'}$, and j). $\sigma_d[i]$ is (1) indistinguishable from σ_Φ knowing $\Im_> = \left\{i, j, (\mu_1, \mu_2), \left(\mu'_1, \mu'_2\right)\right\}$, and (2) causes that $\sigma_d[i]^{-1}(i)$ and $\sigma_\Phi^{-1}(i)$ behave very differently due to the fact that $\sigma_d[i]^{-1}(i)$ only depends on $|i|$; indeed we remark that $\forall \mu \in \mathfrak{S}_n\ \sigma_d[\mu(i)]^{-1} \circ \mu = \sigma_d[i]^{-1}$, and therefore, $\sigma_d[i]^{-1}(i)$ is stable by action of $\mathfrak{S}_n$ on i. (resp. $\sigma'_d[j]^{-1}(j)$ only depends on $|j|$).

The synchronization step 4 is designed to overcome the independence between the choices of the legitimate partners for their respective distributions. It needs that the distributions to have special properties ($\in \zeta(\alpha)$) in order to efficiently play their role. It is efficient in 1/4 of cases (when B picks $\sigma_B = \sigma_\Phi$ and A picks $\sigma_A = \sigma_{\Phi'}$, which we will call 'favorable cases'). If σ_Φ and $\sigma_{\Phi'}$ were chosen randomly by A and B instead of being the respective tidying permutations of Φ and Φ', then ξ could, even without knowing Φ

fridaya@uef.fi

and Φ', use $V_{\xi,1}(i,j) \triangleq \frac{k[i][j]}{n^2}$ in order to estimate say V_A with accuracy $O\left(\frac{1}{\sqrt{nk}}\right)$, in which case $\widetilde{e_\xi} = \left[\frac{\left(V_{\xi,1}-\rho\right)\sqrt{nk}}{K}\right]$ would be as close to $\widetilde{e_A}$ as $\widetilde{e_B}$ in favorable cases.

The step 5 is called Advantage Distillation because at this step, thanks to the discarding condition $|T_m| = 2$, we have managed to create, under Deep Random Assumption, a protocol in which the error rate for the legitimate receiver is strictly lower than the error rate for the opponent.

We remark that the condition $|T_m| = 2$ plays also another role which is to protect against the unexpected predictability effects of using an error correcting code at Reconciliation phase. Indeed, unlike QKD protocols (protected by no cloning theorem) or Maurer's satellite protocol presented in [5], in our case the opponent can try the various possible combinations of choice for $\sigma_{\xi,A}$ and $\sigma_{\xi,B}$, and compare them in terms of matching a code word. To illustrate the issue, let's consider the non-optimal 'bit repeating' error correction method described by Maurer in [5]: the codeword v_A chosen by A can only be $(0,0,\ldots,0)_L$ or $(1,1,\ldots,1)_L$ depending on $e_A = 0$ or $e_A = 1$. B could then publicly discards all decoded sequence v_B that is not $(0,0,\ldots,0)_L$ or $(1,1,\ldots,1)_L$ and obviously decodes accordingly $e_B = 0$ if $|v_B| = 0$, and $e_B = 1$ if $|v_B| = L$. In such scenario, if, for one of the instance m within the sequence, we have $\sigma_A = \sigma_{\Phi'}$ and $T_m = (0,0,0,0)$ (or $T_m = (1,1,1,1)$), then it is clear that the code word is $(0,0,\ldots,0)_L$ (or $(1,1,\ldots,1)_L$). This example shows how, even with Deep Randomness indistinguishability applied locally to each instance, the opponent can still take advantage of an imprudent implementation of the error correcting method to break the secrecy. The condition $|T_m| = 2$ also has the advantage to avoid such predictability flaw.

2 Extended Protocol Against Active Opponent

The protocol $\mathcal{P}$ introduced in [9] and summarized in the former section is suitable to generate a common secret bit string S between the legitimate partners against a passive unlimited opponent. The parameters of $\mathcal{P}$ can be adapted so that, $\forall \varepsilon, \varepsilon' \in]0, 1/2]$

(i) $P(e_A \neq e_B) \leq \varepsilon$
(ii) $\left|P(e_\xi = e_B) - \frac{1}{2}\right| \leq \varepsilon'$

In order to generate S, the partners will execute a sequence of L rounds of $\mathcal{P}$ that can be serialized as shown in Fig. 1.

The purpose of generating S is of course to transmit a secret message M from A to B. The encryption function $E_S(M)$ can be simply $S \oplus M$ if one want the highest security offered by one time pad, or another symmetric key encryption scheme working with a shared secret key S. While we have shown that a passive opponent cannot gain knowledge of S, it is obvious that an active opponent can insert himself in the communication as shown in Fig. 2 below, first playing the role of B with A, which enables him to collect M, and then, optionally, to play the role of A with B to transmit M, or even a different message M' to B.

We will see now how to prevent this MITM attack with an extended version of the protocol $\mathcal{P}$.

fridaya@uef.fi

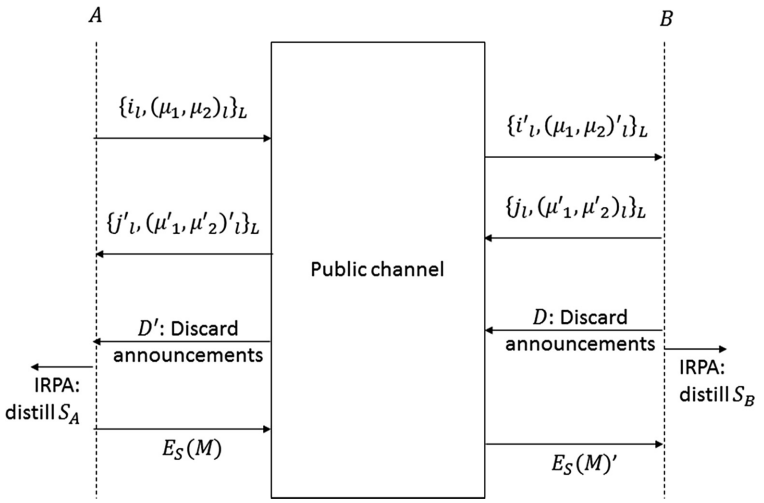


Fig. 1. Diagram of protocol $\mathcal{P}$ without authentication

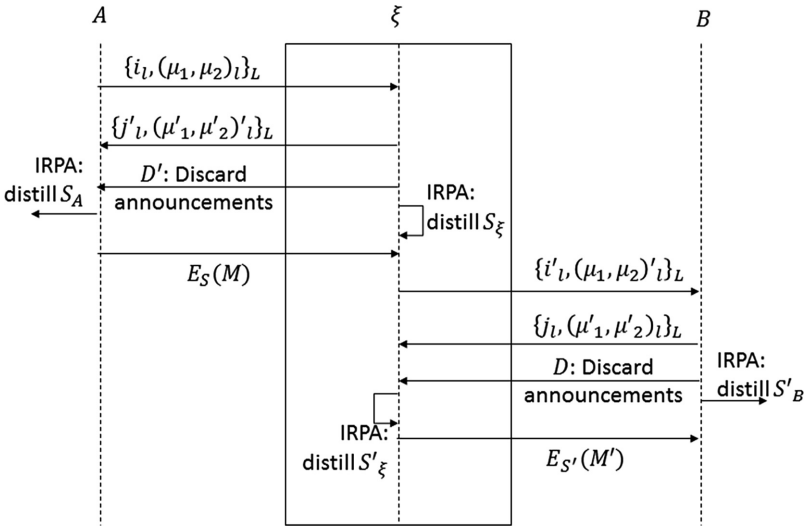


Fig. 2. Example of MITM attack on protocol $\mathcal{P}$ without authentication

2.1 The Security Model

In our 'active opponent scenario', ξ still has unlimited computation and storage power, but is also capable to (i) delete messages from the public channel, (ii) introduce new messages in the public channel at destination of either A or B; (i) + (ii) is also equivalent to the capacity to modify a message transiting from A to B or from B to A.

fridaya@uef.fi

A and B are equipped with a DRG, and also with a private 'wallet' that can contain a shared authentication secret s. This shared authentication secret can be updated at any time by the wallet holder. It is assumed as a pre-condition to the protocol that A and B have initially the same value s_0 in their wallet.

It is assumed that the opponent ξ has no access to (i) the content of a private wallet, (ii) anything computed or stored within a DRG.

The goal of the extended protocol is to continue to ensure that, $\forall \varepsilon, \varepsilon' > 0$

(i) $P(e_A \neq e_B) \leq \varepsilon$
(ii) $\left|P(e_\xi = e_B) - \frac{1}{2}\right| \leq \varepsilon'$

even with such an active opponent.

2.2 Description of the Extended Protocol $\mathcal{P}^*$

The extended protocol that we propose is possible only because of a specific non-reversibility property of $\mathcal{P}$. Indeed it is shown in [9] (Lemma 3) that, when the public information (i,j) is given, with the assumption that the distributions are synchronized, then it is impossible to reversely determine x or y with an accuracy equal to the one of a partner's estimation. $\mathcal{P}^*$ is original because, unlike any other protocol that does not benefit from the non-reversibility property of Deep Random Secrecy, it resists to opponents that are both unlimited and active MITM.

The extension then consists in adding a mutual verification phase after the bit string S has been generated, and before transmitting the secret message M. This verification is performed by:

(i) dividing the shared secret into 2 independent piece (s_A, s_B) (with $H(s_A) = H(s_B) = H(s)/2$)
(ii) sending from A to B (resp. B to A) a verification code

$$c_A = h\left(s_A, \left\{i_l, (\mu_1, \mu_2)_l\right\}_L, \left\{j'_l, \left(\mu'_1, \mu'_2\right)'_l\right\}_L\right)$$

(resp.)

$$c_B = h\left(s_B, \left\{i'_l, (\mu_1, \mu_2)'_l\right\}_L, \left\{j_l, \left(\mu'_1, \mu'_2\right)_l\right\}_L\right)$$

(the characteristics of the verification function h will be discussed hereafter)
(iii) verifying by B (resp. by A) that the local computation

$$v_A = h\left(s_A, \left\{i'_l, (\mu_1, \mu_2)'_l\right\}_L, \left\{j_l, \left(\mu'_1, \mu'_2\right)_l\right\}_L\right)$$

(resp.)

$$v_B = h\left(s_B, \left\{i_l, (\mu_1, \mu_2)_l\right\}_L, \left\{j'_l, \left(\mu'_1, \mu'_2\right)'_l\right\}_L\right)$$

is equal to the received code ($c'_A = v_A$) (resp. ($c'_B = v_B$)

fridaya@uef.fi

(iv) using part of the entropy of S to renew (sending $R(s')$) a new value of the shared authentication secret s' for a next round

(v) storing by A (resp. B) the new value of shared authentication secret s' in its private wallet.

The enhanced protocol is represented in Fig. 3 below.

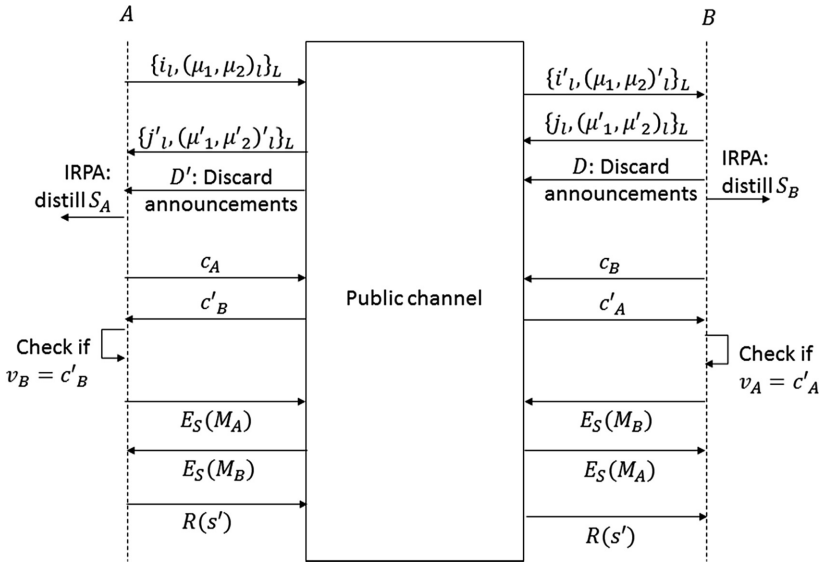


Fig. 3. Diagram of protocol $\mathcal{P}$ enhanced with authentication

Assuming that $E_S(M) = S \oplus M$, M having a length bounded at each round by $|M| \leq L_M$, the common shared bit string must have length:

$$|S| \geq H(s) + L_M$$

in order to be able to fully renew the shared authentication secret s after each round.

2.3 Design Considerations for Security

The verification function h must have at least the following characteristics:

Property 1:

if $\{i_l, (\mu_1, \mu_2)_l\}_L, \left\{j'_l, \left(\mu'_1, \mu'_2\right)'_l\right\}_L$ *are known, then*

$$\#\left\{s \middle| h\left(s, \{i_l, (\mu_1, \mu_2)_l\}_L, \left\{j'_l, \left(\mu'_1, \mu'_2\right)'_l\right\}_L\right) = h\left(s_A, \{i_l, (\mu_1, \mu_2)_l\}_L, \left\{j'_l, \left(\mu'_1, \mu'_2\right)'_l\right\}_L\right)\right\} \gg 1$$

fridaya@uef.fi

Property 2:
if $\{i_l,(\mu_1,\mu_2)_l\}_L, \left\{j'_l,\left(\mu'_1,\mu'_2\right)'_l\right\}_L$ *are known, and* $s \neq s'$

$$P\Big(h\Big(s,\{i_l,(\mu_1,\mu_2)_l\}_L,\left\{j'_l,\left(\mu'_1,\mu'_2\right)'_l\right\}_L\Big) = h\Big(s',\{i_l,(\mu_1,\mu_2)_l\}_L,\left\{j'_l,\left(\mu'_1,\mu'_2\right)'_l\right\}_L\Big)\Big) = o_{H(s)} \tag{1}$$

Due to the non-reversibility property of $\mathcal{P}$, it is not possible for ξ to keep $i_l,(\mu_1,\mu_2)_l$ (resp. $j_l,\left(\mu'_1,\mu'_2\right)_l$) unchanged when transmitting towards B (resp. A) because then it is impossible to generate x and σ_Φ (resp. y and $\sigma_{\Phi'}$) that would enable ξ to play the role of A vis à vis B (resp. B vis à vis A) in determining $S_\xi = S_A = S_B$ with unchanged transmitted public information $i_l,(\mu_1,\mu_2)_l$ and $j_l,\left(\mu'_1,\mu'_2\right)_l$. Therefore, ξ has no other choice than to execute fairly $\mathcal{P}$ with both A and B in order to share a secret with either of them. ξ can try to adjust the value s_ξ of a guessed shared authentication secret, but due to the fact that there exists many possible values of s_ξ such that

$$h\Big(s_{\xi,A},\{i_l,(\mu_1,\mu_2)_l\}_L,\left\{j'_l,\left(\mu'_1,\mu'_2\right)'_l\right\}_L\Big) = h\Big(s_A,\{i_l,(\mu_1,\mu_2)_l\}_L,\left\{j'_l,\left(\mu'_1,\mu'_2\right)'_l\right\}_L\Big)$$

and

$$h\Big(s_{\xi,B},\{i'_l,(\mu_1,\mu_2)'_l\}_L,\{j_l,\left(\mu'_1,\mu'_2\right)_l\}_L\Big) = h\Big(s_B,\{i'_l,(\mu_1,\mu_2)'_l\}_L,\{j_l,\left(\mu'_1,\mu'_2\right)_l\}_L\Big)$$

it is not possible for ξ to guarantee that

$$h\Big(s_{\xi,A},\{i_l,(\mu_1,\mu_2)_l\}_L,\left\{j'_l,\left(\mu'_1,\mu'_2\right)'_l\right\}_L\Big) = h\Big(s_A,\{i'_l,(\mu_1,\mu_2)'_l\}_L,\{j_l,\left(\mu'_1,\mu'_2\right)_l\}_L\Big)$$

and

$$h\Big(s_{\xi,B},\{i'_l,(\mu_1,\mu_2)'_l\}_L,\{j_l,\left(\mu'_1,\mu'_2\right)_l\}_L\Big) = h\Big(s_B,\{i_l,(\mu_1,\mu_2)_l\}_L,\left\{j'_l,\left(\mu'_1,\mu'_2\right)'_l\right\}_L\Big)$$

for all l.

One may question if it is possible for ξ to learn at least some bits of S. If h is not carefully designed, here is how ξ can learn even almost all bits of S. Assume for example that we choose:

$$h\Big(s,\{i_l,(\mu_1,\mu_2)_l\}_L,\left\{j'_l,\left(\mu'_1,\mu'_2\right)'_l\right\}_L\Big) = f\Big(s,\left(\oplus_{l=1}^{L}\left(i_l\oplus j'_l\right)\right),\left(\oplus_{l=1}^{L}\Big((\mu_1,\mu_2)_l\oplus\left(\mu'_1,\mu'_2\right)'_l\Big)\right)\Big)$$

fridaya@uef.fi

ξ can then sacrifice the last instance $j'_L, \left(\mu'_1, \mu'_2\right)'_L$ to control that $c'_B = v_B$ without knowing s_B, by resolving the equations:

$$\begin{aligned}&\left(\oplus_{l=1}^{L} i_l\right) \oplus \left(\oplus_{l=1}^{L-1} j'_l\right) \oplus j'_L = \left(\oplus_{l=1}^{L} i'_l\right) \oplus \left(\oplus_{l=1}^{L} j_l\right)\\ &\left(\oplus_{l=1}^{L} (\mu_1, \mu_2)_l\right) \oplus \left(\oplus_{l=1}^{L-1} \left(\mu'_1, \mu'_2\right)'_l\right) \oplus \left(\mu'_1, \mu'_2\right)'_L\\ &= \left(\oplus_{l=1}^{L} (\mu_1, \mu_2)'_l\right) \oplus \left(\oplus_{l=1}^{L} \left(\mu'_1, \mu'_2\right)_l\right)\end{aligned}$$

The interaction flows perform as shown in Fig. 4 below.

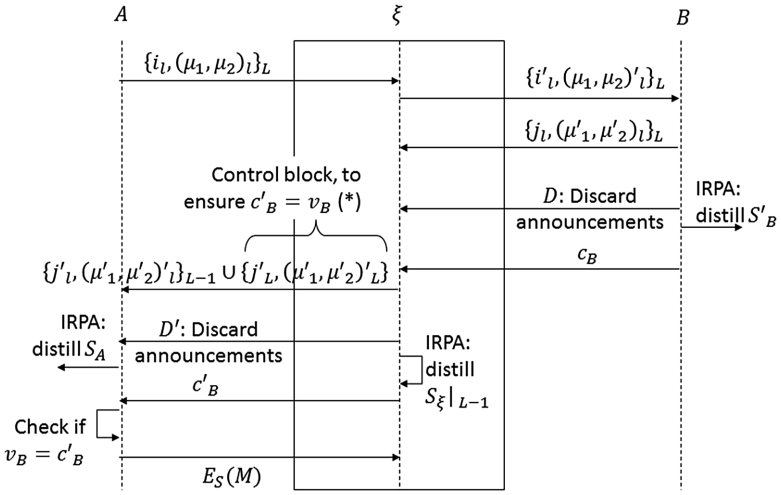


Fig. 4. Diagram on residual MITM attack

By proceeding as described above, ξ can gain the amount of knowledge of S given by performing correctly all instances except the last one, which will give

$$I(S_\xi; S) \geq \frac{L-1}{L} H(S)$$

This attack is made possible because of the existence of transformation over $\left\{j'_l, \left(\mu'_1, \mu'_2\right)'_l\right\}_L$ and/or $\left\{i'_l, (\mu_1, \mu_2)'_l\right\}_L$ that can let h invariant independently of s.

The design of h should then guarantee:

Property 3:

There does not exist any non-trivial transformation T applying on the space $\left\{i'_l, (\mu_1, \mu_2)'_l, j'_l, \left(\mu'_1, \mu'_2\right)'_l\right\}_L$ *such that*

fridaya@uef.fi

$$\forall s, \quad h\Big(s, \big\{i_l, (\mu_1, \mu_2)_l\big\}_L, T(\Big\{j'_l, \big(\mu'_1, \mu'_2\big)'_l\Big\}_L)\Big)$$
$$= h\Big(s, T\big(\big\{i'_l, (\mu_1, \mu_2)'_l\big\}_L\big), \big\{j_l, \big(\mu'_1, \mu'_2\big)_l\big\}_L\Big)$$

or

$$\forall s, \quad h\Big(s, T\big(\big\{i'_l, (\mu_1, \mu_2)'_l\big\}_L\big), \big\{j_l, \big(\mu'_1, \mu'_2\big)_l\big\}_L\Big)$$
$$= h\Big(s, \big\{i_l, (\mu_1, \mu_2)_l\big\}_L, T(\Big\{j'_l, \big(\mu'_1, \mu'_2\big)'_l\Big\}_L)\Big)$$

We propose to use:

$$h_{a,b}(x) \triangleq ((ax+b) \bmod p) \bmod 2^{\frac{H(s_A)}{2}}$$

$$h\Big(s, \big\{i_l, (\mu_1, \mu_2)_l\big\}_L, \big\{j_l, \big(\mu'_1, \mu'_2\big)_l\big\}_L\Big)$$
$$\triangleq \big(\oplus_{l=1}^{L}\big(h_{a_l,b_l}\big(i_l + s\overline{i_l}\big)\big)\big) \oplus \Big(\oplus_{l=1}^{L}\Big(h_{a'_l,b'_l}\big(j_l + s\overline{j_l}\big)\Big)\Big)$$
$$\oplus \big(\oplus_{l=1}^{L}\big(h_{c_l,d_l}\big(\mu_{1_l} + s\overline{\mu_{1_l}}\big)\big)\big) \oplus \Big(\oplus_{l=1}^{L}\Big(h_{c'_l,d'_l}\big(\mu'_{1\,l} + s\overline{\mu'_{1\,l}}\big)\Big)\Big)$$
$$\oplus \big(\oplus_{l=1}^{L}\big(h_{e_l,f_l}\big(\mu_{2_l} + s\overline{\mu_{2_l}}\big)\big)\big) \oplus \Big(\oplus_{l=1}^{L}\Big(h_{e'_l,f'_l}\big(\mu'_{2\,l} + s\overline{\mu'_{2\,l}}\big)\Big)\Big)$$

where p is a prime superior to $2^{\frac{H(s)}{2}}$, and $\{a_l, b_l, c_l, d_l, e_l, f_l, a'_l, b'_l, c'_l, d'_l, e'_l, f'_l\}$ are random non-zero integers, fixed as parameters of the protocol (not negotiated); they can typically be the successive outputs of a congruential generator to avoid to store all them in memory. The permutation μ_{1_l} is written above with its canonical binary form to be able to compute it with arithmetic. $i_l + s\overline{i_l}$ is a non-linear form on i_l also involving s for each $i_l, \mu_{1_l}, \mu_{2_l}, j_l, \mu'_{1_l}, \mu'_{2_l}$. $\{h_{a,b}(x)\}$ is a well known universal hashing class of functions, thus benefiting from the good probabilistic properties of such class.

The modularity structure of h in terms of using blocks of the form $\big(h_{a_l,b_l}\big(i_l + s\overline{i_l}\big)\big)$ and composing them by simple $\oplus$ is chosen on purpose; in case public information $\big\{i_l, (\mu_1, \mu_2)_l\big\}$ are reused from a round to another, the associated block can be reused without re-computation of the arithmetical operations. This reuse technique will be illustrated in optimization techniques for $\mathcal{P}$, presented in further work.

2.4 Other Considerations

Eventually, note that the computation of h is obviously highly sensitive to errors in the transmission of the $\big\{i_l, (\mu_1, \mu_2)_l, j_l, \big(\mu'_1, \mu'_2\big)_l\big\}_L$, and therefore it is recommended to associate each $\big\{i_l, (\mu_1, \mu_2)_l\big\}$ with an error control checksum.

The size parameters $H(s), L_M, L$ should also be fixed as parameters of the protocol and not negotiated, in order to avoid that they become also vulnerable to MITM attacks.

fridaya@uef.fi

3 Conclusion

We have proposed an authentication scheme resisting to unlimited MITM attackers. This scheme is working in conjunction with a key exchange protocol and is only made possible by the very special non-reversibility property of Deep Random Secrecy. The scheme requires a prior shared authentication secret that is kept fully independent from the shared encryption key generated between the partners as the output of the protocol. The shared authentication secret has to be renewed after each round of the protocol for a full secrecy against unlimited opponent, and the renewal process is part of the protocol.

This protocol is then suitable to replace protocols like TLS in scenarios where opponents are assumed unlimited (typically because equipped with quantum computers), but requires a prior registration phase between the partners to initialize the shared authentication secret. Such registration is quite realistic in practice (think about creating an online account for instance). However, one could question the poor level of performance (bandwidth, CPU) of the protocol to replace a popular protocol like TLS.

The main performance concern is of course the quantity of information needed to generate an output shared key S. That quantity is superior to $|S|Rn \log n$ (where R is a constant depending on IRPA method used in $\mathcal{P}^*$). In general, such a key exchange protocol based on Deep Random Secrecy and resistant to unlimited MITM cannot exceed a bandwidth performance of $|S|/C$, where $C \geq 1$ is the Cryptologic Limit introduced in [9]. The CPU performance question can be addressed by searching fastest possible functions h that verifies properties 1, 2 and 3. It can also be questioned if one can lower the requirement of Property 3 together with maintaining the security model; Property seems too much demanding at first sight compared to the objective because many transformations T could be harmless to the protocol.

References

1. Shannon, C.E.: Communication theory of secrecy systems. Bell Syst. Tech. J. **28**, 656–715 (1949)
2. Brassard, G., Salvail, L.: Secret-key reconciliation by public discussion. In: Cryptology – Proceedings of EUROCRYPT '93, Lecture Notes in Comput. Science, vol. 765, pp. 410–423 (1994)
3. Bennet, C.H., Brassard, G.: Quantum cryptography and its application to provable secure key expansion, public-key distribution and coin-tossing. In: Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India, pp. 175–179, December 1984
4. Bennet, C.H., Brassard, G., Robert, J.-M.: Privacy amplification by public discussion. SIAM J. Comput. **17**(2), 210–229 (1988)
5. Maurer, U.M.: Secret key agreement by public discussion from common information. IEEE Trans. Inf. Theory **39**(3), 733–742 (1993)
6. Cachin, C., Maurer, U.M.: Unconditional security against memory-bounded adversaries. In: Proceeding of CRYPTO '97, Lecture Notes in Computer Science, Springer (1997)
7. Jaynes, E.T.: Prior probabilities. IEEE Trans. Syst. Sci. Cybernet. **4**(3), 227–241 (1968)
8. Cox, R.T.: Probability, frequency and reasonable expectation. Am. J. Phys. **17**, 1–13 (1946)

fridaya@uef.fi

9. de Valroger, T.: Perfect secrecy under deep random assumption. Arxiv.org (http://arxiv.org/abs/1507.08258 full version)
10. de Laplace, P.S.: A philosophical essay on Probabilities, Paris 1814
11. Wegman, M.N., Carter, L.: New hash functions and their use in authentication and set equality (1980)
12. Beth, T., Desmedt, Y.: Identification tokens or: solving the chess grand-master problem. In: Advances in Cryptology – CRYPTO'90 Proceedings, pp. 169–176. Springer (1991)

fridaya@uef.fi

Dynamic Password Protocol for User Authentication

H. Channabasava(✉) and S. Kanthimathi

PES Institute of Technology, Bangalore, India
channachannu.97@gmail.com, kanthimathi@pes.edu

Abstract. The Internet security is no more a choice, but a priority. Even with the availability of best security tools and protocols, it is inevitable to avoid unauthorized access and cyber threats. The existing user authentication which is based on username and password fails to work against Cyber-attacks because of its static nature. A novel approach for user authentication called "Dynamic Password Protocol" is proposed, where passwords used for authentication consist of two parts one is a static password which remains static and dynamic password which changes regularly based on the user-defined parameters. These parameters refer to distinct objects which are unique to every user and are accessible by users only at the time of login, parameters defined may be a clock-based mechanism or based on the geo-location of the machine from which the user is logging in. The parameters set by one user are unknown to the other users. Hence the proposed protocol strongly protects the Internet against replay attacks, phishing attacks and other cyber-related threats by means of generating dynamic passwords that are valid for a particular period of time.

Keywords: Dynamic password · Parameter based authentication · Cyber security · User authentication · Cyber attacks

1 Introduction

Securing the data over the digital network has been big concern for the organizations, with people accessing and storing data over the digital network have increased dramatically over a short period but at the same phase it has been difficult for them to secure it from unauthorized access which exploits the user's privacy, with a drastic increase in data generated over the internet security is no more a concern but a priority. One of the most useful and important mechanisms being implemented in securing the data over the network is user authentication which is responsible for authentication and authorization of users in most of the digital network. Present Security Infrastructure uses a username and static password for user authentication where password and username remain static in the database. This kind of security postures is commonly adopted by all interconnected systems over the Internet. Since the passwords are static in nature, vulnerability and repeatability are the two most common problems faced by static passwords, which means they are quite easy to steal and once they are stolen they can be reused. It has been easy for the cybercriminals to bypass these security tools and have the unauthorized access to the data with the help of huge amount of computational

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 597–611, 2019.
https://doi.org/10.1007/978-3-030-22868-2_43

fridaya@uef.fi

power and other password cracking tools which are available over the internet. To address these problems, organizations are adapting to better security tools to protect themselves from cyber threats and data loss but at the same time, cybercriminals are still finding their way for bypassing these security tools. Few of security tools being implemented by the platform of the digital network for the user authentication are Multi-Factor Authentication [6], One Time Password (OTP). These tools provide a lot more secure and safer way of authenticating users over the digital network, but each of these tools has their own respective drawbacks, in both the cases user have to wait for the server to respond with some 4–6 digits code which is randomly generated at the time of login and sent to the user's mobile phone. They are much more reliable and provide a better security for the digital data than the single factor authentication. Since the passwords are generated at the time of login and sent to the user's mobile number, to some extent the authentication is dynamic in nature [7]. Password policies being implemented makes it mandatory for the passwords to be minimum of 8 characters, must include some special characters and much more and moreover, these policies lead to the possibility of forgetting the password. Forgetting passwords is a common problem among most of the users and it is often users tend to keep the password simple to make it easier to remember such as their phone number, date of birth, person name which in turn leads to poor quality of passwords and makes it easier for unauthorized access to their data. User authentication based on Dynamic passwords and Graphical passwords [2] also plays a major role in authentication process, dynamic nature of password can be based on various parameters which can be related to users or it can be random in nature and graphical passwords [2] refers to the set of points in particular sequence on an image or an graphical interface typically electronic screen which is used to authenticate users.

In Dynamic Password Protocol (DPP) user authentication is completely dynamic in nature, passwords are not generated at the time of login but rather they keep on changing regularly based on the user predefined parameters and there is no involvement of external device in the authentication process which is required in the case of One Time Password and Multi factor Authentication. These parameters refer to unique distinct objects predefined by the user during the registration process. These parameters may be a clock-based mechanism which is the system time of machine from which the user is logging-in and geo-location of the user login in terms of coordinates. Passwords in Dynamic Password Protocol is composed of two parts one is a static part which refers to the static characters and the other one is a dynamic part; this dynamic part refers to the characters which keep changing regularly based on predefined parameters. The combination and occurrence of these two parts in the actual password is defined by the user, like the position of dynamic characters, the objects used to represent the dynamic characters in the actual password. The static part is encrypted and is stored in the database separately since the dynamic part changes regularly the parameter on which it changes, objects used to represent the dynamic characters and the position of the dynamic part in the actual password is encrypted and then stored in the database. Since the user is required to remember only the parameter and the static part of the actual password, it is easy for the user to remember password and possibility of forgetting the password is ruled out in dynamic password protocol. Dynamic Password Protocol improves security by using the parameter to generate dynamic passwords that

fridaya@uef.fi

change on each authentication attempt thus preventing repeatability and passwords are dynamic in nature they cannot be predicted nor useless if stolen, which in turn reduce the vulnerability.

2 Related Work

User Authentication is a most important layer of the network security which authenticates and authorizes the user to access the data, user authentication cannot completely rely on single-factor authentication like username and password which proposes many security threats and vulnerabilities. Research work is going on to implement the dynamic passwords in the authentication process, few related paperwork is discussed below.

A novel approach called Knowledge-Based Dynamic Password is presented by the author Basharzad [3] which is based on user's knowledge on depicted picture, user authentication is done based on his/her preselected points on a picture. Possibilities of forgetting passwords are ruled out in this case, user has to remember only the point on a picture which is easier to remember. This is similar to Dynamic password Protocol where user has to remember only the parameter and position of the dynamic part.

Dynamic Password Policy Generation written by the authors of [4, 6] proposes the idea of generating dynamic password policies for the user to create password. In most of the security tools, users are requested to use the passwords which meet the standards set by the password policies and these policies remains the same for all the users. But in case of Dynamic Password Policy Generation, policies are generated in real-time from server and are unique for every user. Policies used for one user will not be known by the other user, which is similar to parameters defined by the user in Dynamic Password Protocol. This dynamic nature of policies and parameters makes it difficult for the hackers to predict the passwords.

In [5] author named Nawaf Aljohani presented a mechanism called Robust Password System Based on Dynamic Factors which works similar to the Dynamic Password Protocol but it consists of a set of different passwords which are used to authenticate the user in a particular period of time I.e. one password is used during the early morning and the other password in evening, in this way set of passwords are predefined by the user during the registration process and are used at different interval of time. In case of Dynamic Password Protocol password changes randomly based on parameters but not on the given interval of time.

3 Server Side Implementation of Dynamic Password

Parameters predefined in Dynamic Password Protocol are responsible for the dynamic nature of the password which makes it unique from the present user authentication methods. The first parameter taken into consideration is a clock-based mechanism, which refers to the system time of the machine from where the user is logging in, which is represented in the form of HHMM where HH means hours and MM means minutes and these values are taken as a dynamic part of the password. Unique objects are used

fridaya@uef.fi

to represent these dynamic characters which are predefined by the user itself and it is left to the user to choose any of the letters among [a–z] such as 'x' to represent those dynamic characters, since time is considered as parameter dynamic part will be represented as 'xxxx' which resembles 'HHMM' time format. User has to specify the position of the dynamic part in the actual password along with the static part, consider the static part to be some characters 'Botnet' and now combining both the parts we are left with 10 characters.

Total number of combination to place these dynamic characters into the static part will be Factorial of 10, which will be 3628800 combination left to the user to choose.

This is one of advantage we have in DPP, the possibilities of two users having same passwords are less and even if in those cases when one of the password is compromised, the password will not be a valid password without the dynamic part and its unique position. In this way DPP provides security to the user information.

Consider user has given static part as 'Botnet' and dynamic part as 'xxxx', the actual password obtained from the combination of these two would be 'Botxxnetxx', 'Botnetxxxx', 'Boxtxnextx', 'Bxoxtxnet' and so on, user can place these dynamic characters anywhere in between the static part which gives rise to new password for every available combination. Both the server and client's machine are synchronized to the local time zone so that there is on mismatch in the time format between client's system and server's time, in case of servers based in India Both the machines are synchronized to IST time zone. If the user is logging-in and system time is 12:32 AM/PM then the dynamic part will be '1232' and actual password after considering the pattern 'Botxxnetxx' would be 'Bot12net32' for next minute when time changes to 12:33 dynamic part will be '1233' which in turn generates new password 'Bot12net33' which is totally different from the password generated earlier and for next hour password will be 'Bot13net32' and so on, password keeps on changing with respect to the system time. The user has to use the different password for every login attempt, which is why dynamic password protocol is less vulnerable to any of the phishing [9], replay attacks and totally secure against brute force attacks. Dynamic password protocol has a major role in a web application, mobile and cloud storage where authentication and authorization of the user are crucial. A clear picture of working is given below:

Parameter defined = 'System Time'
Parameter Representation = 'x'
Static Part be = 'Botnet', Dynamic Part be = 'xxxx'
Actual Password will be = 'Botxxnetxx'
When the system time is 12:32 AM/PM
Then 'xxxx' would be '1232'
The actual password will be 'Bot12net32'
For the next hour, it will be 'Bot13net32'
Again, it would change to 'Bot14net43'
And then to 'Bot15net32' and so on……
Both the Passwords are stored in the database separately, which is shown below.
Actual Password = 'Botxxnetxx'
Static part is stored after encrypting = encrypted (Static part)
Dynamic part = 'xxxx'

fridaya@uef.fi

Dynamic part is stored based on its occurrence in actual password, it checks for the parameter in the actual password when it find the parameter place bit '1' else place bit '0' in the dynamic password, given that actual password is 'Botxxnetxx' then the dynamic password would be Dynamic Password = '0001100011' this dynamic password is stored along with the encrypted static password in the database separate column. There are many possibilities for placing the dynamic part in the actual password. Since the password is different for every minute they cannot be compromised by using brute force attacks or any other attacks.

3.1 Registration Phase

The user is requested to define the character to represent the dynamic password and also to define the position of the dynamic characters. Here in the shown example, 'x' is defined for representation of parameter and also the user has defined the position of these characters in the actual password. Instead of remembering the whole password in case of present user authentication, users have to remember only the static part i.e. 'Botnet' along with the position of the dynamic password. This reduces a lot of pressure on users and makes it easy for remembering the passwords. After the registration process is complete users are redirected to the login page. This is shown in Fig. 1.

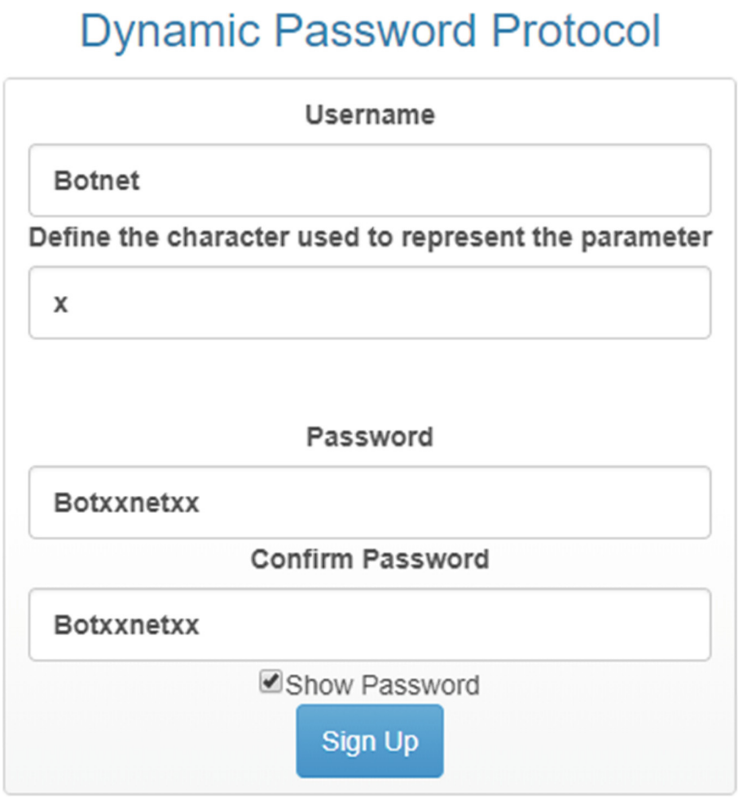


Fig. 1. Registration phase

fridaya@uef.fi

When the user gets the Login response from the server, the user authenticates himself to the server with username and password which is the composition of both the dynamic and static part. The server then verity's the user with the database if the user exists in database then it returns with the parameter.

After the registration process is completed, users are redirected to the login page. Bcrypt function is used for encryption of the static part, which uses Blowfish encryption algorithm. This algorithm generates random salt for every hash value which makes it impossible for the hash collision to take place.

Algorithm for Registration Phase

```
1. Username ← Input name       //Botnet
2. Parameter ← Time (HH:MM)   //parameter being discussed
3. Parameter representation ←  'x'
4. Password ← Botxxnetxx
5. //Combination of both dynamic and static part
6. For each I in Password
   a. If (I == x)
      (1) Append '1' to dynamic part
   b. Else
      (1) Append '1' to static part
      (2) Append '0'
   c. End If
7. End For each
8. //now dynamic part = 0001100011
9. //and static part = Botnet
10. Hash = Encrypt(static part)    //Encrypt() returns hash value of static part
11. Parameter = dynamic part
12. Insert into Database(Username, Hash, Parameter)
13. End
```

Algorithm for Encrypt ()

```
14. Password  =  static part
15. Cost = 10
16. Bcrypt (Cost, Salt, Password)
    d. State ← EksBlowfishSetup (Cost, Salt, Password) //Blowfish algorithm
    e. Ctext ← "OrpheanBeholderScryDoubt"
    f. Repeat (64)
    g. Ctext  ← EncryptECB (State, Ctext)
17. Return Concentration (Cost, Salt, Ctext)
```

fridaya@uef.fi

3.2 Login Phase

When the user gets the Login response from the server shown in Fig. 2, the user authenticates himself to the server with username and password which is the composition of both the dynamic and static part. The server then verifies the user with the database if the user exists in database then it returns with the parameter being defined by the user, with the help of parameter obtained server then extract the static part and dynamic part from the given password. For example, if the predefined parameter is system time, when the system time is 12:24 AM/PM then the user enters the password as 'Bot12net24' based on parameter defined server extracts the dynamic part from the database which would be

Parameter = '0001100011'.

Based on the parameter obtained server then extracts the static and dynamic part from the given password, where the static part would be 'Botnet' and dynamic part would be '1224'.

First the server verifies the dynamic part of the given password with the parameter value, if it is valid only then it proceeds with the further authentication process which verifies the static part with the encrypted static password stored in the database. If the dynamic part is invalid the server would respond with invalid credentials response.

If the user attempts to login again in the next minute with same password i.e. 'Bot12net24' server responds with invalid login credentials, because the dynamic part of the password has been changed to '1225' which means the actual password for the next minute will be 'Bot12net25'.

In this way, parameters are responsible for the dynamic nature of the passwords and also generate a new set of passwords for every minute. In this way, Dynamic Password Protocol makes it less vulnerable to the replay attacks and phishing attacks, also provides better encryption to the user information.

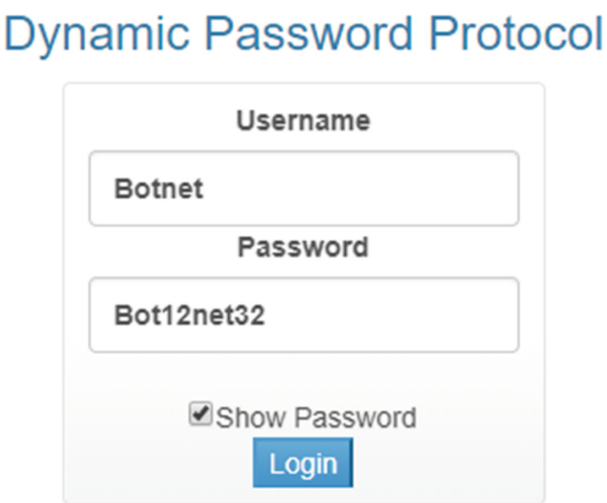


Fig. 2. Login phase

fridaya@uef.fi

Algorithm for Login Phase

- Username ←Input username //Botnet
- Password_input ← Inputs password //Bot12net32
- Verify (Username)
    - If user exists, return True
    - Else exit
- If (True)
    - Get (Password, Parameter)
    - For each I in parameter
    - If (I == 1)
        - Append dynamic = Password_input
    - Else
        - Append static = Password_input
    - End If
    - End For each
    - Time = Date(HH:MM) //Gets system in HHMM format
    - If (Time == dynamic)
        - Password verify (static, Password)
        - // Verifies the Hash value of static with that of the Password
        - If (True)
            - Permit Access
        - Else
            - Exit()
    - Else
        - Exit()
    - End If
- Else
    - Exit()
- End

3.3 Internal Architecture of Registration and Login Phase

Below attached figures gives the internal architecture of Registration and Login phase, which represents the internal operation performed by the server. Figure 3 represents the Registration Phase which performs internal operation to extract the dynamic and static part from the input password based on the parameter chosen by the user and its representation. Figure 4 represents the Login Phase which represents the unique feature of the proposed protocol, authentication is done at two stages i.e. first the server verifies the dynamic part and if it is true then it proceeds with the authentication of the static part else the server responds with the invalid credentials.

fridaya@uef.fi

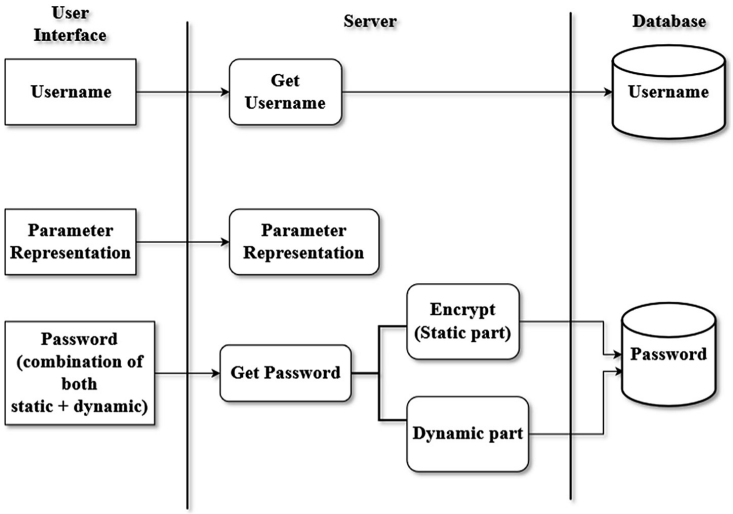


Fig. 3. Architecture of registration phase

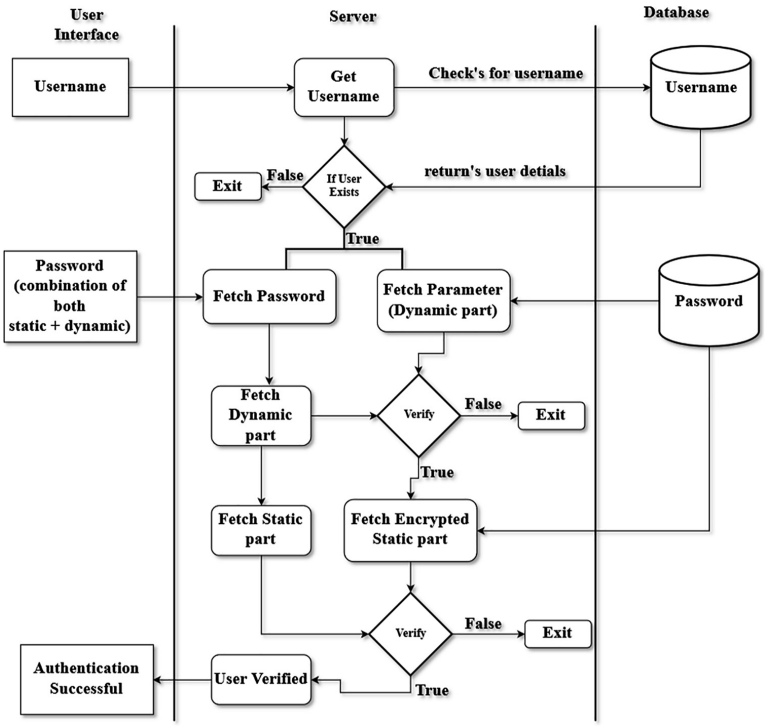


Fig. 4. Architecture of login phase

fridaya@uef.fi

4 Advantages of DPP Over Existing Protocol

4.1 Authentication at Multi Stages

In proposed protocol authentication is done at two stages, server verify both the static and dynamic part at two different stages, when users send a login request to the server, server first verify whether the username is valid or not if there is no account with the username then it respond with invalid credentials. If username exists then the server fetches parameter from database associated with the username and extracts the static part and dynamic part from the given password, then the server verify the dynamic part with the update value of the parameter predefined by the user and this is first stage of authentication. If the dynamic part does not matches with the updated value of predefined parameter value then server responds with invalid request. If it is valid then further authentication is carried by the server to validate static part of the given password with the encrypted static part from the database, this is second stage of authentication. Proposed protocol is similar to user using two factor authentication like OTP. In this way the proposed protocol provides better authentication infrastructure then the existing protocols.

4.2 Multi Factor Authentication and One Time Password

In One Time Password (OTP) [15], the user has to wait for the server to respond to the user with the randomly generated 4–6 digits to the registered user's phone number. In case of the server is down or the user might not have the access to his mobile phone, OTP cannot provide an efficient and effective way for user authentication and also it makes it mandatory for the user to carry their mobile phone with them all the time. In case of Dynamic Password Protocol, no external device is involved in the authentication process and also part of the password is automatically generated based on predefined parameters.

Multi-Factor Authentication [15] which refers to Two Factor Authentication ends up in consuming a lot of time in login process by asking the user to go through two authentication process i.e. at first user is requested to login with username and password, after successful attempt user is requested to use the random code sent to their mobile phone which is similar to OTP else the user is asked to answer secret question which is predefined by the user. This time consumption is reduced in the Dynamic Password Protocol by making the authentication process simpler and faster with the help of both static and dynamic passwords, first the protocol verifies with static part of the password similar to logging with username/password and then the protocol checks for the dynamic part of the password which is same as sending some randomly generated code to phone numbers or asking for predefined questions at the same time; thus, reducing the processing time for user authentication.

5 Mathematical Analysis of DPP Against Various Threats

In the proposed protocol, Bcrypt function is used to encrypt the static part, which is based on Blowfish Algorithm.

fridaya@uef.fi

Let us consider the specification of Blow-fish algorithm:

- It is preferred to have a cost value of 10 I.e. 10 iterations of hashing.
- Salts are added to the given value to get the hash value.
- Salts are randomly generated for every hash.
- Returns 24 byte of hash value I.e. 192 bits.
- It takes around 47 ms to generate hash for a given value on a CPU with 8x Nvidia Gtx 1080 processor.
- According to 8x Nvidia Gtx 1080 processor Hash Cat Benchmark 2017 [11]
- It would take 47 ms to compute hashes by bcrypt function.
- It takes around 105700 hashes/sec to decrypt the hash value by the 8x Nvidia 1080 processor.

5.1 Birthday Attack

This attack is based on the possibility of matching the hash value with the total combination of available, given size of the hash value [12, 13]. Attacker generates all possible hash values from the encryption algorithm which was used to encrypt the user's information, when he gets hash values of all the passwords, attacker matches these hash values with an available combination of hash values and retrieves the actual password if the hash matches.

In the proposed protocol passwords are encrypted using bcrypt function.

Here the size of single hash is 192 bits, then the total number of possible hash combination would be 2^ (N/2) where N is the total number of bits that each hash consists of, which is around total number of combination, given by Eq. (1).

$$\mathrm{NC} = 2\hat{}(192/2) = 7.9228\mathrm{E} + 28 \tag{1}$$

Total time taken, which is around

$$\mathrm{TT} = 2.28\mathrm{E} + 21 \text{ years} \tag{2}$$

Even with latest GPU processor it would reduce to 2.2309E+21 years.

When generating hashes requires huge amount of time, then amount of time required to compare and retrieve password from the obtained hash values will take much longer time, which is infeasible for an attacker to perform such kind of attack.

So Birthday attacks are ruled out in the proposed Dynamic password protocol which uses bcrypt encryption algorithm to store the user information.

Let's consider the Worst case scenario, with the help of Supercomputer or Botnets, if the attacker is able to retrieve the password from the hash value it would be still invalid because of passwords in DPP are combination of both static and dynamic part, both static and dynamic part are required for authenticating the user.

5.2 Brute Force Attack

As one of password policies in DPP, users are advised to have minimum 6 characters for static part of the actual password and including 4 dynamic part characters to it, actual password will be of length 11 characters.

fridaya@uef.fi

In worst case scenario, consider the total length of actual password be 8 characters then consider character set to be 95 character which is combination of 26 uppercase, 26 lowercase letters, 10 numeric, 33 special characters [13, 14].

Password entropy,

$$E = \left(95^8\right) = 52\,\text{bits} \tag{3}$$

From the above Eq. (3), crack time using all the possible combinations of 8 character length word given by Eq. (4).

$$\text{Time} = 2\text{^}E/(\text{Number of guess}) \tag{4}$$

Assume, attacker will crack some of them in last, first or in between of total combinations of available word list

$$\text{Time} = 2\text{^}E/(\text{Number of guess}) * 0.5$$

Average cracking time,

$$\text{Time} = 2\text{^}E/105700 * 0.5 = 675 \text{ years}$$

Even when the attacker uses some word list or the passwords list it could come down to some 100 years which is still infeasible way for cracking.

If the attacker is able to brute force into the system and retrieve the password [10], it will be invalid for the next login attempt. It's because the dynamic part of the password changes regularly which in turn changes the actual password which is completely different then the password used earlier.

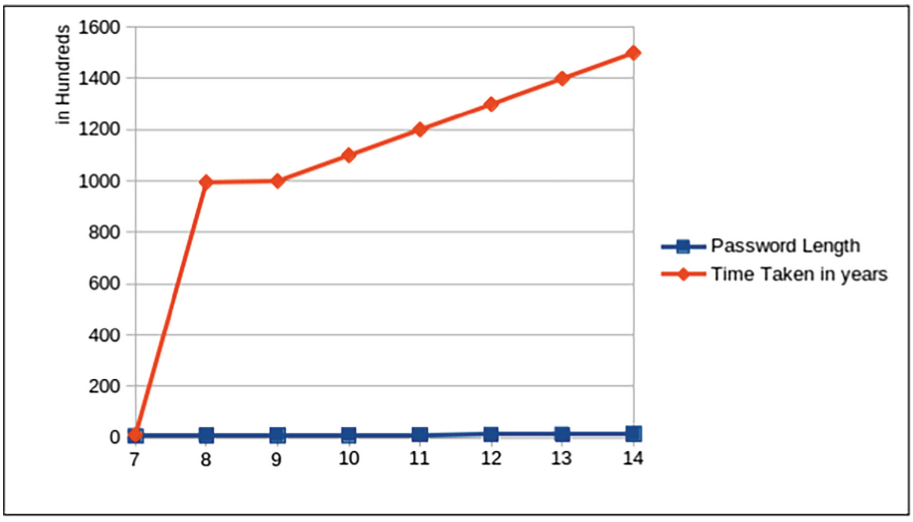


Fig. 5. Time taken by Brute force attacks vs. password length

fridaya@uef.fi

The above graph, Fig. 5 shows the time taken to brute force the passwords with the help of available word list and password list with the minimum of 7 characters of password length to maximum of 14 characters.

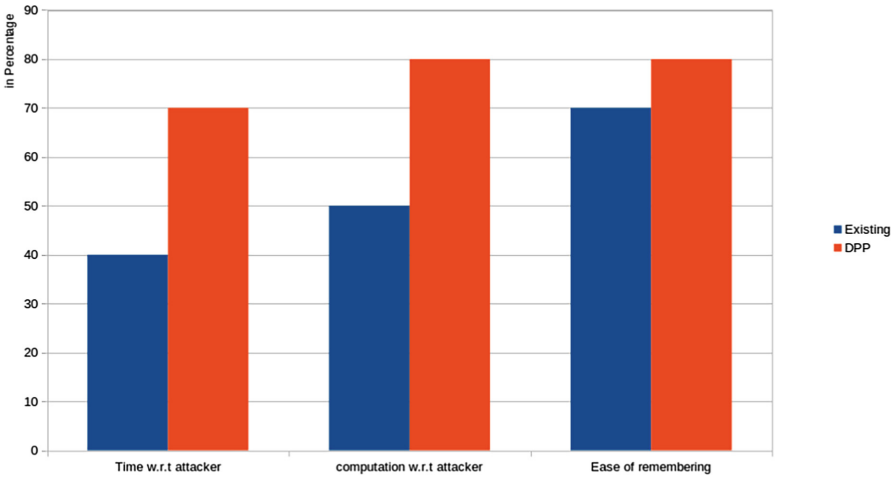


Fig. 6. Comparison among existing and proposed protocol

When proposed protocol was compared against existing protocol [1, 8] with respect to few parameters like Time complexity in brute forcing the passwords, computational complexity in generating hashes and then retrieving the information from them and then finally the ease of remembering the passwords, DPP makes it easy for users to remember the password shown in Fig. 6. Consider the password is of length 10 characters, in case of existing protocols is difficult for the average user to remember all the 10 character but in proposed protocol user has to remember only the 6 characters i.e. only the static 6 static characters and remaining 4 characters will be of dynamic part.

5.3 Addressing Password Attacks Over DPP

There is a wide range of methods available for bypassing the user authentication process and dynamic password protocol stands against these attacks, few of them are discussed below:

- Replay Attack

Consider the attacker is able to intercept the communication between the client and server [9] and if the user is subjected to this attack an attacker has the password which the user has used to log in, let the password which user has given be 'bot12net24' and hacker has this password. When the hacker uses this password for the next login attempt, the server responds him with the invalid credentials every time he tries to log in because the password is not the same as the one the actual user used earlier, it has been changed and the earlier password is no more useful. This approach of Dynamic Password Protocol makes it completely secure against these attacks.

fridaya@uef.fi

- User is compromised

The most common problem faced by the static passwords is the difficulty of remembering them i.e. the possibility of forgetting password is huge among most of the user. The user ends up in writing down those passwords on a paper or they end up in using a poor password like their phone numbers, date of birth and this password are easy to gather and are easily available. Dynamic Password Protocol makes it easier to remember and to have a stronger password at the same time since users are requested to remember only the static part and parameter being defined, it is effective for the user to remember password thus reducing the possibility of forgetting the passwords.

6 Future Enhancement

Proposed paper is limited to the clock-based mechanism as a parameter and can extend to the geo-location (in terms of coordinates) of the user being used as a parameter. Since the position of the dynamic part in the actual password remains static research work is going on to make it dynamic and send the its position to the user through the secure channel only at the time of login and few more parameters are being added to the protocol to provide users the opportunity to choose among many parameters.

7 Conclusion

In Dynamic Password Protocol, passwords are verified at two levels, first the dynamic part of password is verified and if it is valid then further authentication process is carried out by verifying the static part of the password otherwise if the dynamic part is invalid at the first place then the user is not authorized to access the data even without proceeding to further authentication process. Considering all the aspects of Dynamic Password Protocol it has the potential to provide a safer and secure way for user authentication by avoiding the vulnerabilities of static password and making it easier for users to remember the password, at the same time have a stronger password. In the proposed Protocol user need not carry any kind of external device to authenticate, at the same time it can rule out most of the drawbacks from the existing protocols. Static password being vulnerable to most of the cyber threats whole internet cannot completely rely on it and user data is the most valuable part of the internet, single-factor authentication like static username and password cannot withstand against the evolving cyber threats. With the implementation of Dynamic Passwords, user authentication can be made even stronger and safer process by reducing the vulnerability. Hence Dynamic passwords have a major role in user authentication and maintaining privacy of user's data over all the platforms across the internet.

fridaya@uef.fi

References

1. Yang, S., Ji, S.: A dynamic password policy generation system. IEEE Trans. Inf. Forensics Secur. **13**(3), 545–558 (2018)
2. Chien, W.: Random graphic user password authentication scheme in mobile devices. In: Proceedings of the 2017 IEEE International Conference on Applied System Innovation IEEE-ICASI 2017
3. Basharzad, S.N., Fazeli, M.: Knowledge based dynamic password. In: IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), pp. 0367–0372, December 22 QG (2017)
4. Aljohani, N.: Robust password system based on dynamic factors. In: 2016 6th International Conference on Information Communication and Management (2016)
5. Anani, W., Ouda, A.: The importance of human dynamics in the future user authentication. In: 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE) (2017)
6. Biswas, S., Biswas, S.: Password security system with 2-way authentication. In: 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCIN) (2017)
7. Kalaivani, S.: Multi factor authentication scheme for shadow attacks in social network. In: 2017 International Conference on Technical Advancements in Computers and Communications (2017)
8. Pramana, M.I.W.: Keystroke dynamics for authentication using dynamic time warping. In: 2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE) (2017)
9. Tian, Z.: An automated brute force method based on webpage static analysis. In: 2018 10th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA) (2018)
10. Dieckow, A.: A white paper by Andreas Dieckow Principal Product Manager InterSystems Corporation. https://www.intersystems.com/wp-content/uploads/assets/Using_Two-Factor_Authentication_WP.pdf
11. Software: Hashcat v3.00-beta-145-g069634a, Nvidia driver 367.18, 8x Nvidia GTX 1080 Hashcat Benchmarks (2017)
12. Birthday Attack Homepage. https://wikipedia.org
13. Stamp, M.T.: Information Security Principles and Practice, 2nd edn. Wiley, Hoboken (2011)
14. Brute Force Attack Homepage. https://wikipedia.org
15. Two Factor Authentication Homepage. https://wikipedia.org

fridaya@uef.fi

Side-Channel Attacks on the Mobile Phones: Applicability and Improvements

Roman Mostovoy[1], Pavel Borisenko[1], Daria Sleptsova[1], Alla Levina[1,2(✉)], and Igor Zikratiov[3]

[1] ITMO University, Saint-Petersburg, Russia
alla_levina@mail.ru
[2] Saint Petersburg Electrotechnical University "LETI", Saint-Petersburg, Russia
[3] Bonch-Bruevich Saint Petersburg State University of Telecommunications, Saint-Petersburg, Russia

Abstract. On one hand, side-channel attacks are a powerful approach to cryptoanalysis, on the other hand, these attacks have a very specific attacker model. As a result, a practical application of SCA is sometimes disputable. The threat level in each case has to be evaluated given individual features of a certain target system. Endpoint client applications like mobile phone applications are especially vulnerable because of an uncomplicated access to the device. More attention has to be paid to their security since they're more accessible to an attacker and usually carry a large amount of sensitive information.

Current research represents the informativity assessment of the side-channel signals obtained from mobile phones. Low-priced accessible equipment was used in order to decrease requirements to an attacker level and consequently increase the applicability of an attack. Logically, it results in a lack of information that can be obtained from such traces, but there are ways to achieve a good interpretation of the signal. For staying within the consumer-segment equipment, the algorithmic approach increasing traces informativity was implemented. The algorithm is Normalized Inter-Class Variance and it allows to distinctly improve the efficiency of trace analysis by reducing the number of points and increasing of the quality of reference traces.

Keywords: Side-channel attacks · Mobile phones · NICV algorithm

1 Introduction

A new class of attack is threatening modern devices and crypto algorithms, which is called side-channel attacks (SCA). Side-channel attacks target physical parameters of a system and not mathematical basis behind an algorithm as it would be in the case of a classic attack. In order to use any crypto algorithm, it needs to be implemented on a device, like a PC or a dedicated hardware module, and the hardware level brings a whole new threat surface. Hardware consumes power, emits electromagnetic signals and leaks information in the acoustic range depending on the operations being processed. Cryptographic algorithms consist of different operations that require different power and time

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 612–621, 2019.
https://doi.org/10.1007/978-3-030-22868-2_44

fridaya@uef.fi

to be processed. By analysing leaked information an attacker can conclude what part of the algorithm was run or even what key was used on the device.

The first ever side-channel attack was done in 1956, when MI5 placed a microphone near the Hagelin enciphering machine, recording click sounds. First cryptographic use of side-channel was done by P. Kocher, where he used timing attacks to break RSA, DH, DSS by measuring time of execution of different code branches and operations [3]. More complicated attacks followed around 1999 – it was discovered that power consumption leaks information about DES and AES using both simple [17] and differential analysis [4]. Modern side-channel attacks use many non-invasive ways to measure leakage, for example, electromagnetic signal [7, 8] or acoustic cryptanalysis [9, 10].

A side-channel attack usually requires a target device to be in the physical proximity to an attacker (the only exception could be timing attacks or software-based SCA). In light of this fact, the most exposed devices are ones that are always with the victim, that are movable and mobile such as smartphones. Those devices are often left unattended and can be easily approached by an attacker. Accessories pose another threat: rechargeable power banks or protective cases may convey malicious recorders of electromagnetic emanations or power consumption.

Ever since smartphones grew to perform many functions, other than calling, and to substitute many other devices, information that they possess has been growing in confidentiality and sensitivity. They contain not only personal or sensitive information of the owner, but also access to remote servers, private computers and critical management systems. All this information is concentrated on one portable device reachable for an attacker, which makes it an attractive target.

When it comes to choosing a method for the attack, there are two important options. Side-channel attacks can be divided into two groups - active and passive attacks [2]. Active attack means direct intervention to the operation of a device. For example, it is performed by decapsulating the chip or by inducing faults in the device using power glitches. Information about fault combined with the input and output can recover the plaintext. Passive attacks are more complicated, but they are less visible to the user, and the device remains untouched. This work are concentrating on the latter, more sophisticated type of SCA because they pose a real danger to smartphone users. They are able to steal information unnoticed, sometimes even without direct interaction with the victim device.

This paper is organized as follows: the second section provides the theoretical background to side-channel attacks with stress on mobile devices and the current state of the art in the field of passive attack on mobile devices. Section 3 presents a practical example of an attack on a mobile device. Section 4 introduces the NICV method, which is further elaborated to aid the attack. Conclusions and further work can be found in the last section.

2 Theoretical Background

Any computation on a physical device will produce a physical signal, giving away information about its physical state at every point in time. For cryptographic algorithms leaking such information may be crucial. Many algorithms have different logic or

fridaya@uef.fi

pieces of code that depend on a secret key. They have distinguishable patterns of physical characteristics, which attacker can use to recover secret data. This kind of attacks is called side-channel attacks, and lately, it has been used to break many cryptographic algorithms and devices running cryptographic code.

A side-channel attack can be denoted by three main components: target, side channel and attacker. Depending on their interactions with one another, an attack may be passive or active. Let us refer to a definition conducted in a recent survey [1] in order to elaborate it further. Figure 1 shows the overall structure of an attack.

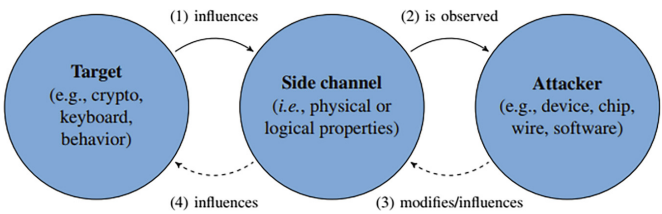


Fig. 1. General notion of passive (?) and active (?) side-channel attacks

A passive attack is depicted on the steps 1 and 2. The target device is in operation and therefore it results in physical emanation or changes, giving away secret information through a side channel. An attacker can observe these physical properties of the device, and connect the physical effect with the actual programmatic cause of it.

An active attack extends further and includes steps 3 and 4 as well. An attacker interacts with the device via a side channel, e.g., via an external interface or environmental conditions. This gives more control over the computation flow, a possibility to circumvent some security controls. By observing the resulting properties of the device, the attacker is able to deduce the secret information.

Steps 3 and 4 in a smartphone environment are not impossible but are unrealistic. Therefore, passive attacks and a "classic" side channel are most interesting to study. Mobile side-channel attacks are a very broad field that includes various channels for getting information [5, 6]. For example, it covers cache attacks, fingerprinting and classic power and EM analysis. This investigation uses a classic side channel, which is power consumption. For the moment there is one work done in the power dimension. Yan et al. [11] were able to identify running apps from their power consumption. There are several works on EM emanations [12, 14], mostly targeting elliptic curve cryptography [12, 13, 15].

There are several obstacles for conducting a successful attack on a smartphone:

1. First of all, there is an unwanted signal coming from the graphic user interface (GUI). Any application needs it to interact with the user, but working with GUI elements leads to active power consumption from the phone screen. Its signal mixes up with the signal directly related to cryptographic operations.
2. Other applications distract users and try to get their attention, or just an external source gathers users' attention. Multi-tasking is a usual thing for smartphone users. It leads to noisy parasitic signal and a lot of garbage in the signal.

fridaya@uef.fi

As a way to improve signal processing and overcome these obstacles, method called NICV was employed. Normalized Inter-Class Variance (NICV) was presented in 2013 in [16]. This method allows detection of interesting time samples, without the need of a profiling stage on a clone device. It allows to speedup up analysis stage of an attack and to compress SCA traces. The main characteristics of NICV are:

- no clone device is required for NICV method, so there's no requirement for profiling stage, and the same set of traces can be analyzed using the method;
- it requires only public information such as plaintext or ciphertext;
- the method doesn't depend on a leakage model, it doesn't behave as an analysis tool but helps to accelerate the analysis;
- it can be used for evaluating the accuracy of different leakage models and choosing the one that fits the best.

In continuation to previous work which includes implementation and testing of NICV for analyzing the signal received by side-channel attacks along with other methods [9], the current research is extended into the area of smartphones and use NICV method to improve the attack.

3 SCA on Mobile Phones

Multiple restrictions are put on side-channel attacks by real-life applications, so a special application was created that contains a "sandbox" of cryptographic primitives. The first refinement minimizes GUI usage, because it is not generally used required by cryptographic operations. Secondly, it contains an easily expandable set of cryptographic primitives, and a managed set of secret encryption keys. In addition, the application is capable of executing required cryptographic operations at a high frequency for a swift acquisition of necessary amount of data that is required for an attack.

The following list is the complete description of application's features:

- Support of DES, 3DES, AES, RSA encryption;
- Alterable values of secret keys and initial vectors;
- Support of detailed logging along with time-stamping, as it is crucial to keep the data and the attack tool synchronised;
- Delay management for cryptographic operations with high-frequency. Production application contains similar functionality to help data gathering;
- Automatic generation of random keys and plaintexts;
- Supported encryption schemes: CBC and ECB;
- Configurable number of equally starting signatures.

Coming from the premise that no expensive equipment is available for an attacker to conduct an attack, two widely accessible tools were used to record traces - a jack to jack cable and a sound card connected externally.

As part of this study, it was decided to focus on an attacker model with low-level hardware capabilities. A similar approach was adopted to the software under study. This section covers an attack on two distinct models of mobile phones (Alcatel POP3

fridaya@uef.fi

and Xperia M2). All computational operations are performed in a limited well-controlled environment and further analysis of the results is provided.

A *clean* run is recorded using the sandbox application, that is, most of the device activity is reduced, background processes are turned off and the screen is switched off. The reason behind this decision is the character of the pseudo-random behaviour of user input, i.e., registered signals are highly dependent on the actions of the mobile phone user. The cryptographic operations performed in the "silent mode" have generally clearer traces, as shown in Figs. 2 and 3.

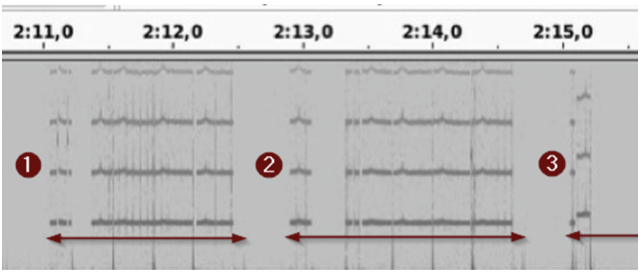


Fig. 2. Trace of Alcatel POP3 running only encryption: numbers denote each encryption operation

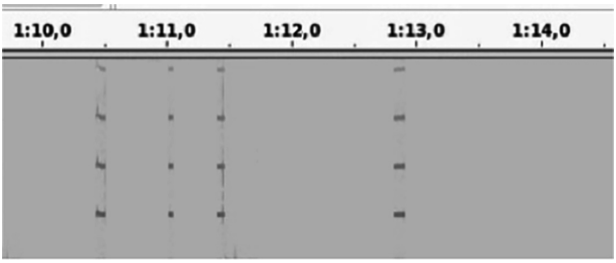


Fig. 3. Trace of Xperia M2 running only encryption: no visible leakage on the jack port

At this stage, it is clear that differences in phones' hardware result in a different signal profile: on the Xperia M2 trace cryptographic operations cannot be reliably distinguished.

On the next step a comparison was made of the induced signal on the connector from two phones using the same application, but with different side processes. Wi-Fi, Bluetooth and other common modules were turned on, and the user was interacting with the device. And again, there is a clear difference in data clarity. Traces on Figs. 4 and 5 show a completely indistinguishable noise flow for Alcatel and an almost empty (in terms of information level) trace for Xperia.

Based on the gathered recordings and their analysis, it was decided to proceed the analysis of traces only from Alcatel POP3 due to the lack of informational component in the traces collected from the Xperia M2.

fridaya@uef.fi

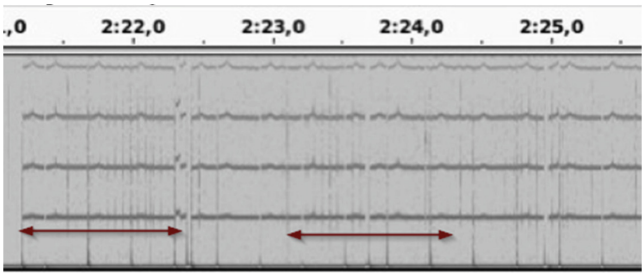


Fig. 4. Trace for Alcatel POP3 in the normal mode with user interaction, software-recognisable pattern

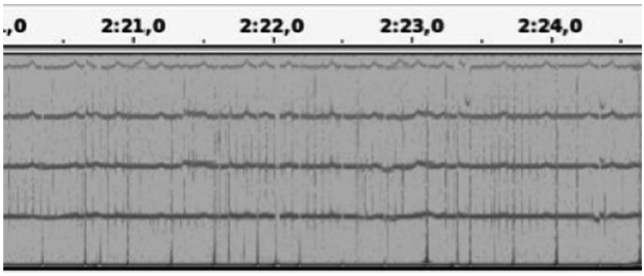


Fig. 5. Trace for Xperia M2 in the normal mode with user interaction, no discoverable leakage

Next parsing the raw data to the form of vectors was required to allow the neural network to work with it. In the mentioned form each vector represents a single signal recorded during a particular operation.

If the synchronisation is fully achieved for the trace with recognisable encryption operations, i.e., knowing the first N trace signatures, the attacker can predict the next starting and ending points of the signature. The precision of this mechanism is highly dependent on the sampling of the sound card and the periods of time necessary to perform cryptographic operations on the mobile device.

A special parsing software was developed in Python in order to recognise differences between signatures The parser receives a signal recorded during cryptographic operations, and a corresponding time-stamped file that was produced by the sandbox application. The purpose of the parser is to recognise all cryptographic operations from the acquired signal, avoiding noise, and forward them as separate vectors for further processing in the neural network.

For this goal the parser consists of two equivalent methods:

- *Convolution function*: In this method the analysing function uses the entire trace and the signatures inside as functions. As was stated earlier there is a simplification to the process, which consists of setting equal several N signatures at the stage of traces recording. The parser looks at all hypotheses of the first signature and calculates the resulting convolution function between the hypothesis and the whole

fridaya@uef.fi

trace. The peaks of the resulting functions outline which hypotheses are most alike and how many signatures corresponding to this hypothesis exist in the trace. After the first signature is found, the parser analyses the trace and determines the starting and ending points of other signatures, basing on the starting point of the signature file and the timestamp. Finally, each signature is specified in the trace and can be loaded as a separate vector. The results of this method are presented on the Fig. 6 (left).

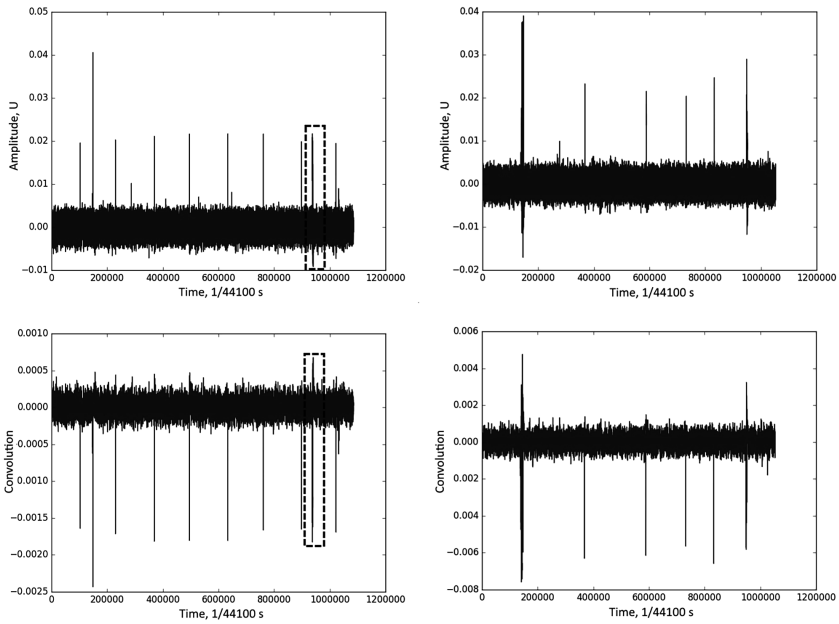


Fig. 6. Signatures detection with convolution function (on the left) and convolution function with neighbourhoods (on the right)

- *Convolution function with neighbourhood*: Second method has many similarities to the previous one. Here the analyser uses information from the timestamp file (how much time is left in between and during cryptographic operations) as a "lattice" for signature detection. Convolution function is computed between each hypothesis of the first signature and the corresponding hypothesis of other signatures that are discovered only on the basis of time: it is assumed that the correct starting point of the first signature is known, and calculating the duration from the timestamp file, it is possible to assume where each other signature is on the trace. Thus, by analysing the maxima of convolution calculations the parser can recognise which hypothesis of the first signature is the best. After that, it checks the starting and ending points of other signatures and loads them. The results of this method are presented in Fig. 6 (right).

fridaya@uef.fi

After the analysis was done it is possible to conclude two facts:

- Signatures are not detailed enough due to the resolution and low sampling rate of the sound card, which is significant for further analysis;
- In the second example, it is clear that some signatures were found incorrectly. The first method on the same trace produced a similar outcome. The roots are the same as in the first point: such inaccuracy is a result of the poor resolution of signatures.

The use of more expensive and functionally improved equipment for data collection would mitigate both problems from above.

4 NICV Application

As a method for detecting vulnerable data points in encryption process traces, NICV algorithm has been chosen. It is based on calculating Normalized Intra-Class Variance for each trace within a given sample in order to determine the points, most suitable for subsequent attacks [16].

Encryption key detection when analysing large amounts of data can take from several hours to a day. The main application of NICV is to search for the most vulnerable pieces of data in the traces to speed up the key search.

For The calculation of NICV is based on the following formula from [16]:

$$\mathrm{NICV} = \frac{\mathrm{Var}\ [\mathbb{E}[Y|X]]}{\mathrm{Var}[Y]}$$

In the above expression, Y is a leakage value received by the attacker while discovering encryption process data for plaintext or cipher text containing byte X. Var designates variance function, and $\mathbb{E}$ stands for mean.

Calculation of NICV function consists of several stages:

1. Determining all possible values of the byte X.
2. Classification of traces based on the value of the byte X (256 classes).
3. Mean calculation of the leakage Y for traces within each constructed class.
4. Calculating leakage variance of value Y for all collected traces - Var (Y).
5. Calculating leakage variance of value Y for each individual class using the formula:
6. Element-by-element addition of the corresponding coordinates of 256 vectors obtained on step 5.
7. Element-by-element division of vector coordinates obtained on step 6 into corresponding coordinates of the vector Var (Y).

NICV reflects a maximum of possible correlations, computable from X with Y [16]. In practice, a NICV root is used, which is due to a decrease in the effect of computational noise. NICV values are in the range [0; 1].

The recorded points of the encryption process for which the NICV takes values above the specified threshold are identified as the most vulnerable. During the experiments for this work, the NICV was calculated for records containing 10000 points.

fridaya@uef.fi

Due to the aforementioned issues with signal quality recorded using low-cost equipment, NICV method didn't show a significant difference between points with respect to their vulnerability. The reason for such effect is that a large volume of noise affects the traces. Because of this, the variance of each particular byte is generally higher than thresholds defined. It leads to weak data compression and finally the number of points was decreased in each trace record from 10000 to only ∼9000 (10%).

5 Conclusion

Information infrastructure of today is full of mobile phones and wearable devices that become more and more popular. Considering the importance of the information users store on them, the issue of their safety is critical. This paper shows that there are opportunities for attacks on such devices with no expensive equipment and, more important, in a passive way - i.e., without a threat of exposure for attackers.

However, the attack should be accurately tuned for each specific device. Two targets – Alcatel POP3 and Xperia M2 – showed completely different leakage. The latter didn't have any leaking signal on the jack port, unlike the Alcatel phone.

It is shown that attackers are able to use cheap tools and public optimization methods like NICV in order to uncover more sensitive information from target devices, in less time. We've reduced the number of points by 10% and the quality of reference traces. It allowed to perform successful attacks with accessible equipment that is sold in every tech store.

References

1. Francis, J.G.F.: The QR transformation I. Comput. J. **4**, 265–271 (1961)
2. Standaert, F.X.: Introduction to side-channel attacks. In: Secure Integrated Circuits and Systems, pp. 27–42 (2010)
3. Kocher, P.: Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems. In: CRYPTO 1996, pp. 104–113 (1996)
4. Kocher, P., Jaffe, J., Jun, B.: Differential power analysis. In: CRYPTO 1999, pp. 388–397
5. Spreitzer, R., Moonsamy, V., Korak, T., Mangard, S.: Systematic classification of side-channel attacks: a case study for mobile devices. IEEE Commun. Surv. Tutorials **20**(1), 465–488 (2018)
6. Levina, A., Borisenko, P., Mostovoy, R.: SCA as mobile security threat. In: Proceedings of the 20th Conference of Open Innovations Association FRUCT - 2017, pp. 236–241 (2017)
7. NACSIM 5000 Tempest Fundamentals (Report): National Security Agency, February 1982
8. Genkin, D., Pipman, I., Tromer, E.: Get your hands off my laptop: physical side-channel key-extraction attacks on PCs. J. Crypt. Eng. **5**(2), 95–112 (2015)
9. Backes, M., Dürmuth, M., Gerling, S., Pinkal, M., Sporleder, C.: Acoustic side-channel attacks on printers. In: Proceedings of the 19th USENIX Conference on Security (USENIX Security'10), p. 20. USENIX Association, Berkeley, CA, USA
10. Genkin, D., Shamir, A., Tromer, E.: Acoustic cryptanalysis. J. Cryptol. **30**(2), 392–443 (2017). https://doi.org/10.1007/s00145-015-9224-2

fridaya@uef.fi

11. Yan, L., Guo, Y., Chen, X., Mei, H.: A study on power side channels on mobile devices. In: Symposium of Internetware – Internetware 2015, pp. 30–38. ACM (2015)
12. Gebotys, C.H., Ho, S., Tiu, C.C.: EM analysis of Rijndael and ECC on a wireless java-based PDA. In: Cryptographic Hardware and Embedded Systems – CHES 2005, ser. LNCS, vol. 3659, pp. 250–264. Springer (2005)
13. Goller, G., Sigl, G.: Side channel attacks on smartphones and embedded devices using standard radio equipment. In: Constructive Side-Channel Analysis and Secure Design – COSADE 2015, ser. LNCS, vol. 9064, pp. pp. 255–270. Springer (2015)
14. Ometov, A., Orsino, A., Andreev, S., Levina, A., Borisenko, P., Mostovoy, R.: Mobile social networking under side-channel attacks: practical security challenges. IEEE Access **5**, 2591–2601 (2017)
15. Genkin, D., Pachmanov, L., Pipman, I., Tromer, E., Yarom, Y.: ECDSA key extraction from mobile devices via nonintrusive physical side channels. In: Conference on Computer and Communications Security – CCS 2016, pp. 1626–1638. ACM (2016)
16. Bhasin, S., Danger, J.-L., Guilley, S., Najm, Z.: NICV: normalized inter-class variance for detection of side-channel leakage. In: IEEE International Symposium on Electromagnetic Compatibility, vol. 3, pp. 310–313. Tokyo (2013)
17. Bechtsoudis, A., Sklavos, N.: Side channel attacks cryptanalysis against block ciphers based on FPGA devices. In: Proceedings of IEEE Computer Society Annual Symposium on VLSI (IEEE ISVLSI'10), Kefalonia, Greece, 5–7 July 2010

fridaya@uef.fi

Improving Learning Efficiency and Evaluation Fairness for Cyber Security Courses: A Case Study

Emin Çalışkan, Risto Vaarandi(✉), and Birgy Lorenz

TalTech University, Tallinn, Estonia
{emin.caliskan, risto.vaarandi, birgy.lorenz}@taltech.ee

Abstract. During the last decade, active learning and e-learning methods have become increasingly popular, and a number of universities have redesigned their courses by replacing lectures with online content. Since traditional lecture-based teaching methods have been criticized in recent research papers, we have studied how well a lecture-based approach suits for teaching a highly technical cyber security course. Our study is based on the analysis of survey data collected from past students who have taken the same cyber security course during the last seven years, and our findings provide insights for redesigning similar network and cyber security courses.

Keywords: Cyber security education · Teaching methods · E-learning

1 Introduction

During the last two decades, cyber-attacks have grown both in size and sophistication. Today's cyber threats involve a wide variety of actors with different goals, for example, state actors conducting APT-based cyber espionage, cyber criminals executing DDoS attacks for extortion, etc. The rapidly evolving threat landscape has introduced a requirement for dedicated cybersecurity personnel in many organizations. Also, it has motivated universities to create both undergraduate and graduate programs for educating cybersecurity specialists and meeting the demand of the labour market. Cybersecurity study programs usually cover a wide spectrum of areas and include different courses, for example, courses on penetration testing and ethical hacking, legal aspects of cyber defence, monitoring technologies (such as intrusion detection), cryptography, cyberpsychology, big data analysis and machine learning (such as advanced algorithms for log analysis). Taking such an interdisciplinary study program involves a considerable effort since complex topics from different fields have to be studied in parallel. Since there is often little overlap between different courses, memorizing their content is more challenging than with closely related courses from intradisciplinary study programs. Therefore, it is essential to find ways for improving learning efficiency during cyber defence courses.

During the last decade, active learning and e-learning techniques have become increasingly popular. Active learning and e-learning have obvious benefits - for

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 622–638, 2019.
https://doi.org/10.1007/978-3-030-22868-2_45

fridaya@uef.fi

example, replacing lectures with video material allow for teaching considerably larger audiences [3]. According to some research papers, the use of these techniques results in considerable improvement of student results [8, 17]. Also, a number of universities have redesigned some of their courses by using online materials for replacing all lectures or for significantly shortening the lecturing time [3, 4, 13, 15]. With such setting, students visit the classes for in-depth discussions, lab experiments, group work on more complex problems, etc. However, according to some sources, such techniques alone do not always lead to improvements in student results [6, 14]. Another common approach is so-called blended learning where active learning and e-learning are combined with lectures [2, 11, 16]. With this approach, classroom response systems are often used for facilitating active learning during lectures [1, 9].

One non-trivial issue which arises during the evaluation of cyber security students is the evaluation fairness and cheating during graded tests and exams. On one hand, graded tests and exams of many cybersecurity courses involve highly technical questions where the use of a computer seems natural, and recent research papers have recommended this examination format for computer science courses [5, 7]. For example, developing a correct network IDS signature during the exam is easier if the student can test the signature on a virtual machine which runs the IDS software. On the other hand, cybersecurity students become familiar with many hacking and data exfiltration techniques during their studies. This knowledge makes it considerably easier to utilize computers for illegal information sharing during examinations and for practising other forms of cheating (for example, receiving illegal help from an outsider). In order to address this issue, the right balance has to be found between computer-based tests and other forms of examinations which do not involve computers or other devices with network connectivity.

In this paper, we present a case study for the Cyber Defense Monitoring Solutions course which is part of the Cyber Security MSc program at TalTech University (Estonia). The course has been taught since 2010 in lecture-based fashion with some active learning elements. In this paper, we are focusing on several research questions:

- How efficient have past blended learning based course iterations been for teaching a complex cyber security course?
- Is the current student evaluation methodology fair and efficient against plagiarism?
- Given the recent criticism of lecture-based teaching, should the course continue with blended learning teaching methodology, or is a radical redesign necessary which would replace all lectures with e-learning materials?

For answering those questions, we have measured a long-term knowledge retention of past cyber security students from years 2011–2017. We have also conducted a survey and interviews with students from 2011–2017, in order to study their opinions on course content and learning efficiency, evaluation fairness, cheating and plagiarism. Based on the results of the survey and interviews, we are planning to introduce several improvements to the course. Finally, our case study can serve as an example for improving similar blended learning based cyber security courses.

fridaya@uef.fi

2 Related Work

To better understand the learning efficiency for the cybersecurity courses, previous work on educational objectives and measuring success plays an important role. One of the first notable publications in this area is a taxonomy of educational objectives published by Bloom, which divide the functions of thought into 6 categories [18]. Almost half a century later, Anderson et al. revised Bloom's work and updated the keywords of learning objectives. The figure below depicts these objectives with a comparison [19] (see Fig. 1).

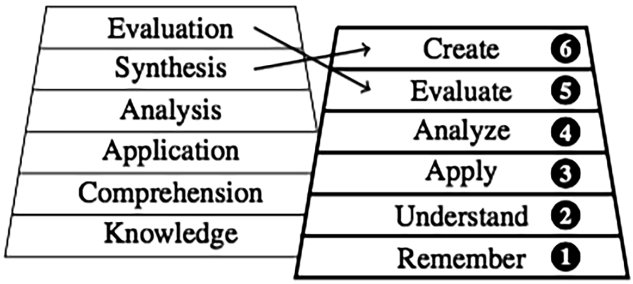


Fig. 1. The comparison of learning objectives: Bloom's Taxonomy (on the left) vs Anderson's Taxonomy (on the right).

In their study, Moses and Petullo mapped Anderson's learning objectives to the domain of computer security, with an aim to maximise learning effectiveness at the undergraduate level [20]. They have applied some key design elements of a secure architecture such as confidentiality, integrity, availability and encryption to the 6 levels of Anderson's objectives list. Finally, they studied a group of students to observe the correlation between challenge completion rates and these learning objectives in a cybersecurity exercise. Although the study underlines the importance of categorisation and the level of capability for each security element, it did not elaborate the efficient learning mechanisms for that respect.

Vasserman et al. followed a different approach in their study and assessed student outcomes from a variety of pedagogical approaches with an aim to measure student interest and self-efficacy in relation to classes and research involving cybersecurity [21]. In that study, a quantitative assessment tool for cybersecurity courses developed and piloted in a test run. The study focused on the output and effects on the students rather than the efficiency of particular courses exist in the curriculum.

Another approach which focuses on instructional models has been developed by Merrill which is known as the First Principles of Instruction [22]. According to his model (see Table 1), the most effective learning processes or environments are problem-centered and involve the learner in the five distinct phases.

Merrill uses the word problem to include a wide range of activities, with the most important characteristics being that the activity is a whole task, rather than small parts of a task and that the task is representative of those the students will encounter in the real world. His approach focuses on the practical applications of the learning objective, which is in line with cybersecurity training and education goals.

fridaya@uef.fi

Table 1. Learning phases according to the first principles of instruction model by David Merrill

Phase	Definition
Problem	Engaging in solving real-world problems
Activation	Activation of existing knowledge
Demonstration	Demonstration of the new knowledge to the learner
Application	Application of the new knowledge by the learner
Integration	Integration of the new knowledge into the learner's world

Adam et al. have used a mobile application, "Quiz It", which acts as a Classroom Response System (CRS) to enhance active learning methods and assess students' understanding during the lectures [1]. The application enabled students to respond to various questions and stay active during the lectures.

Bajwa has discussed the effectiveness of the Flipped Learning method, in which students follow video lessons outside the classroom and the instructor uses the class time for using active learning techniques as a more effective mode of instruction [3]. The author suggests that using Flipped Learning methods are particularly advantageous in terms of time, cost and infrastructure effectiveness. Furthermore, there are promising results for improving learning outcomes as well.

Brandsteidl et al. present an experience report in their study which discusses replacing traditional classroom lectures with lecture videos, with a special focus on Object-Oriented Modeling courses [4]. The main difference of this approach comparing to a typical video recording of the lectures is that instructors record the courses using their own computers and recording slideshows with audio. They have identified the main advantage of this approach is the flexibility it brings to students as well as instructors. Time-effectiveness is another important point for both parties.

Fairness of the exams is yet another important topic to discuss, especially for computerized exams. In his paper, Grotz discusses a web-based exercise platform for providing a fair computer-based exam for a large group of students when the number of computers is limited [5]. For a coding exam, he found that by mixing the difficulty and order of the questions, it was possible to achieve this goal, although further research is advised.

In their paper, Akakura et al. analyze differences in knowledge acquisition resulting from how an e-learning system operates, based on the findings they have observed via a blended learning system they have implemented, with video recordings of lectures and a variety of functions like practice exercises [2]. They have found interesting results highlighting the importance of exercises; students who have not completed the online training materials participated and achieved successful results with the practical exercises. According to these results, it is fair to say that students have a tendency for practical examples, especially in e-learning environments.

In another study, Guy et al. focus on students' performance comparison of traditional face-to-face lectures and web-based tutorials [6]. Their study spanned 4 years between 2008 and 2012 with a total number of 375 students. Supporting other research, they have also discovered that web-based tutorials were as efficient as traditional lectures.

fridaya@uef.fi

3 Course Improvement Study

3.1 Description of the Course

A Cyber Security MSc program of TalTech University was introduced in 2009. The incentive for implementing this program came from public, private, and academic sectors. As of 2018, the program offers three specializations: cybersecurity, digital forensics, and cryptography. The Cyber Defense Monitoring Solutions course has been part of the Cyber Security MSc program TalTech University since 2009. Since 2010, the course has been taught in each fall semester.

The course is focused on commonly employed security monitoring concepts and technologies, using open source security monitoring solutions for mastering practical skills of students. Since the first iteration, the course has had the same structure with the following eight modules:

- Introduction to concepts of event logging, syslog protocol and rsyslog [24] software suite
- Packet filtering with Linux Netfilter [25] firewall
- Introduction to regular expression language for searching data from event logs
- Monitoring event logs with Swatch [26] and the use of regular expressions for complex log management tasks
- Advanced event logging with syslog-ng [27]
- Introduction to concepts of event correlation and Simple Event Correlator (SEC) [28]
- Advanced event correlation with SEC
- Intrusion detection and prevention with Snort [29]

Starting from the first iteration, the course has been taught in a blended learning fashion during 16 weeks of the semester, where each module is presented during a lecture, with the lecture being followed by a lab on same topics in the following week. During the lab, group work is encouraged and several lab assignments assume cooperation between students since student collaboration has been identified as an important active learning technique [10]. So far, no specialized e-learning techniques have been employed for replacing some of the lectures (such as pre-recording lecture material and making it available online). For accomplishing labs and homework assignments, the students have been provided with a Linux virtual machine image at the beginning of the course, with the image containing all necessary software for the course. In order to facilitate distance learning, course slides have been composed in a highly detailed fashion and can be used for studying course topics even without being physically present at lectures.

During the course, students have to complete two homework assignments, both yielding up to 20 points. For facilitating active learning through student collaboration, students are expected to form groups of 2–3 persons for tackling homework assignments. During the exam, each student receives three individual assignments which have to be completed within three hours, and the exam yields up to 60 points. Therefore, the student can collect up to 100 points from homework and exam, and his/her grade is determined by the collected score according to the rules in Table 2.

fridaya@uef.fi

Table 2. Rules for determining the grade of the student

Passing Grade Formula
If score > 90, then grade = 5 (A or "excellent")
If score > 80 AND score <= 90, then grade = 4 (B or "very good")
If score > 70 AND score <= 80, then grade = 3 (C or "good")
If score > 60 AND score <= 70, then grade = 2 (D or "satisfactory")
If score > 50 AND score <= 60, then grade = 1 (E or "pass")

In 2011–2017, the average yearly grade of students has fluctuated in between 2.9 and 4.0, with the average grade for all students being 3.38. In 2010–2015, exams were conducted in a computer class and students were permitted to use a classroom computer or a personal laptop during the exam. However, in 2014–2015 several anonymous complaints were received by the lecturer about cheating cases during the exam, with some students receiving outside help over the Internet from third parties. For preventing such cheating scenarios, students have been not allowed to use computers since 2016 and have to surrender all devices with network connectivity before the exam. The exam is strictly paper-and-pencil based, although the students are allowed to use hardcopy materials during the exam (such as lecture slides printed on paper).

3.2 Collecting Student Feedback About the Course

In recent research papers, various active learning and e-learning techniques have been recommended for replacing traditional lectures [3, 4, 13, 15]. On the other hand, blended learning which combines lectures with active learning and e-learning remains a widely used approach [2, 11, 16]. This raises the question: should the course continue with the current blended learning approach, or should lectures be replaced with e-learning materials?

It should be noted that any redesign of teaching and grading methods should take into account the needs and preferences of the audience - for example, a number recent studies have collected student opinions and measured student satisfaction [10, 11, 14, 15]. In this paper, we have adopted a similar approach and conducted an extensive electronic survey and a focus group interview for past students from 2011–2017. Our aim was not only to collect opinions for improving the course but also to find which parts of teaching and grading system are efficient enough and should not be adjusted. Therefore, our study can also serve as an example for improving similar blended learning courses that employ lectures, so that advantages of the course are not lost during the redesign process.

During our study, we sought answers to the following questions:

- How efficient have past iterations of the course been in terms of knowledge retention, and how well do past students remember the course content?

fridaya@uef.fi

- What is the opinion of students on cheating and plagiarism, and is paper-based exam regarded as an efficient countermeasure against cheating? Also, is the evaluation process regarded fair by students?
- What is the opinion of students on the course quality and how can the course be improved? Should the course continue with blended-learning teaching methodology or should the lectures be discontinued for the favour of online materials?

For collecting the opinions of past students from 2011–2017, we contacted all 192 students who have taken the course. Initial invitation to participate in the study was issued via e-mail, with a reminder being sent about one week later. We employed e-mail addresses that were used by students at the time of taking the course (unfortunately, some of the e-mail addresses were no longer valid). For providing their feedback, two options were offered for different student groups - an electronic survey (offered for students from 2011–2017) and focus group interview (offered for students from 2017 only). Both the electronic survey and focus group interview were organized in March 2018.

The electronic survey consisted of two parts:

- Part 1: A test of 12 questions with multiple choice answers that were designed to measure the knowledge of past students
- Part 2: A questionnaire of 20 questions about the course quality, evaluation fairness, plagiarism and cheating, and propositions for improving the course content

Since the students from 2017 took the exam in January 2018 and their knowledge and impressions from the course were fresh in March 2018, we decided to invite them to focus group interview, asking them to take the electronic survey only if participation in the interview is not possible. Students from 2011–2016 were invited to take part in the electronic survey only. During the focus group interview, students were asked to answer the same questions as in part 2 of an electronic survey, in order to establish their opinion on course quality, plagiarism, etc. In addition, the students were asked a number of questions about ways for improving the course content and evaluation fairness, fighting plagiarism and cheating, and the employment of various active learning and e-learning techniques. For facilitating the discussion, students were demonstrated different e-learning environments and styles of teaching, asking them which styles and techniques seem most natural and efficient.

From 192 students who have attended the course and taken the exam in 2011–2017, 51 students (26.6%) provided their feedback, with 46 students taking the electronic survey and 5 students participating in the focus group interview. The overall participation rate can be regarded relatively high, considering the fact that participation was voluntary and a number of student e-mail addresses were no longer valid. Note that sample size 51 from the population size 192 means that for confidence levels 90% and 95%, margins of error are relatively high (9.89% and 11.79% respectively). For achieving 95% confidence level and 5% margin of error, sample size 129 would be required (67.2% of all students) which was clearly not realistic for the seven-year time frame of our case study. However, despite a smaller sample size and a higher margin of error the survey nevertheless revealed interesting trends and patterns which will be covered in the following subsection.

fridaya@uef.fi

3.3 Analysis of Student Feedback

Table 3 provides an overview of students who have participated in the study by the year the course was taken (since no-one from the participants failed to pass the course, the table has no column for grade 0).

Table 3. Participation data

Year	# of students who took the course	# of students who took part in the survey or interview	Participation rate	# of students with grade 5 (A)	# of students with grade 4 (B)	# of students with grade 3 (C)	# of students with grade 2 (D)	# of students with grade 1 (E)
2011	18	4	22.2%	2	1	1	–	–
2012	27	4	14.8%	3	1	–	–	–
2013	27	8	29.6%	6	1	–	1	–
2014	29	7	24.1%	4	2	1	–	–
2015	35	12	34.3%	3	4	4	1	–
2016	33	10	30.3%	6	3	–	–	1
2017	23	6	26.1%	2	2	–	1	1

Table 3 provides several interesting insights into participant data. Firstly, we expected the earlier years have a significantly lower participation rate since students who have taken the course 5–7 years ago might no longer consider it relevant. However, apart from the year 2012, participation rates are comparable for all years. Secondly, as can be seen from Table 3, the average grade of participants was 4.16. On the other hand, the average grade for all students was 3.38 which indicates that the performance of participants during the course tended to be higher than the average.

In the electronic survey, part 1 was a test of 12 questions which was designed to measure the knowledge of past students about the course content. All questions had 4 predefined answers with exactly one answer being correct. No time limit was set for completing the test and respondents were asked not to use help from other persons, although the use of printed and Internet-based materials was allowed. The difficulty level of questions ranged from fairly basic to highly complex. Table 4 provides an overview of test results for 46 respondents.

Table 4. Test results for all respondents

# of correct answers	# of respondents	The average course grade of respondents	Average # of years since the course was taken
12	1	5.0	7.0
11	5	3.8	4.2
10	11	4.6	3.8
9	8	4.5	3.5
8	3	4.0	4.3
7	9	3.8	4.1
6	2	4.0	3.0
5	6	3.7	2.8
4	1	5.0	6.0

fridaya@uef.fi

As data in Table 4 indicate, the test was non-trivial, since only one respondent managed to pass it with full score. Nevertheless, results also indicate that respondents remember course topics reasonably well - 25 respondents (54.3%) managed to provide right answers to at least 75% of questions, while 39 respondents (84.7%) answered at least half of the questions correctly. When analyzing the data, we also expected to see better test results for students who have taken the course more recently and passed the course with a higher grade. However, as data in Table 4 indicate, the test result is not correlated with the grade or time since the course was taken. Part 2 of the electronic survey also contained a question about the current job of the past student and the areas of expertise that the job assumes. The question provided 8 predefined areas, with 4 predefined areas covering major security monitoring fields which are taught during the course. When preprocessing the survey data, we converted answers to this question to a real number from the range 0..1, where 0 indicates that the current job of the respondent does not assume any expertise on security monitoring, while 1 means that the respondent's job relies on expertise on all 4 security monitoring fields from the survey.

Table 5 below depicts test results for respondents that have been grouped by expertise their current jobs assume. According to Table 5, the average test result of respondents increases with the knowledge requirements of their jobs and is highest for respondents who need to be familiar with all major security monitoring fields that the course covers (expertise level 1). Surprisingly, almost half of respondents with jobs that do not assume the course knowledge (expertise level 0) performed a lot better than expected during the test, achieving approximately the same average result as respondents with expertise level 0.5.

In order to find out the opinions of respondents about the value of the course, we asked them to evaluate five statements in the scale 1..5, where 1 denotes "I disagree" and 5 denotes "I agree". Table 6 lists the feedback to five statements, where each cell is listing the number of respondents who provided the given answer.

Table 5. Test results for groups with different expertise.

Expertise level of the respondent	# of respondents	The average percentage of correct answers
0 (does not have to be familiar with any of the 4 security monitoring fields)	22	70.5%
0.25 (has to be familiar with 1 security monitoring field out of 4)	9	56.5%
0.5 (has to be familiar with 2 security monitoring fields out of 4)	6	69.4%
0.75 (has to be familiar with 3 security monitoring fields out of 4)	4	75.0%
1 (has to be familiar with all 4 security monitoring fields)	5	81.7%

fridaya@uef.fi

Table 6. The value of the course.

Question	I disagree	I mostly disagree	I neither agree nor disagree	I mostly agree	I agree
The course is a valuable addition to the cyber security MSc program	–	3	1	7	40
The course content has been useful during my professional career	2	3	11	10	25
When attending the course, fellow students regarded me an expert on course material and often asked for my advice	7	6	16	16	6
The course was too complex for me and I never understood most of the course material	39	10	2	–	–
When attending the course, I did not like it and felt that it needs to be redesigned from scratch	36	11	2	1	1

As can be seen from Table 6, 47 respondents (92.2%) think that the course is a valuable addition to the entire MSc program, and only 2 respondents (3.9%) did not like the course and think that it needs a redesign. Furthermore, 49 respondents (96.1%) did not consider the course too difficult for them, while 22 respondents (43.1%) were regarded either knowledgeable or highly knowledgeable by other students. Also, although 22 respondents (43.1%) have jobs which do not assume any security monitoring knowledge, 35 respondents (68.6%) have nevertheless found the course useful during their professional careers. Together with data from Tables 4 and 5, these results indicate that the current version of the course is well designed and useful in the opinion of respondents, resulting in a fairly high knowledge retention. Since our survey also allowed to provide free-form recommendations about the content of the course, we investigated feedback from 2 respondents who were not satisfied with the course. One of the respondents recommended to reduce the number of modules in the course and suggested to focus on just 2–3 topics like firewalling and log collection. However, we argue that omitting other essential topics such as network intrusion detection would not provide an adequate view of security monitoring to students. The second respondent recommended to include demos from production systems in the course.

In the survey, we also asked for feedback about evaluation fairness, plagiarism and cheating during the course. Our first question concerned student satisfaction with their final grade and whether the grade reflected their real knowledge on course topics. Respondents were asked to answer on the scale 1..5, where 1 denotes “my grade was too low and I deserved a higher grade, 3 denotes “my grade was entirely adequate”, and 5 denotes “my grade was too high and I deserved a lower grade”. Table 7 depicts feedback about evaluation fairness (since only 46 respondents out of 51 took the test, test result averages are not listed in the table).

fridaya@uef.fi

Table 7. Evaluation fairness.

Evaluation fairness	# of respondents	Average grade of respondents
1 (I deserved a higher grade)	1	5.0
2	7	2.1
3 (My grade was adequate)	35	4.5
4	7	4.3
5 (I deserved a lower grade)	1	5.0

According to the Table 7, 8 respondents (15.7%) were unhappy with their grades. However, the only respondent who was strongly stating his/her grade should have been higher actually completed the course with grade 5 which suggests a mistake when filling the survey form. While the average grade of other dissatisfied respondents was 2.1, other respondents groups had significantly higher grades, with 35 respondents (68.6%) feeling their grade was entirely fair. It is interesting to note that 8 respondents (15.7%) regarded their grades too high. Since the survey also allowed for providing feedback about homework and exam, we investigated the opinions of dissatisfied students. Several students were regarding the exam and homework as too complex. Also, some students proposed to increase the number of homework and their weight to the final grade, while also simplifying the homework.

In the survey, we also collected student opinions on plagiarism and cheating and asked respondents to assess the fraction of students that were involved in academic misconduct. Only 7 respondents (13.7%) believed that there were no cases of plagiarism and cheating when they took the course. In contrast, 22 respondents (43.1%) and 27 respondents (27.5%) had the opinion that 10% and 20% of students were involved in plagiarism and cheating, respectively. Also, 8 respondents (15.7%) believed at least 30% of students have plagiarized or cheated during the course. These results indicate that plagiarism and cheating are issues which need attention and evaluation methods for countering them. In order to establish the views and personal experiences of respondents with these illegal practices, our survey contained several questions which are listed in Table 8 (cells of the table provide the number of respondents for the given answer).

Table 8. Respondent opinions on plagiarism and cheating.

Questions	I disagree	I mostly disagree	I neither agree nor disagree	I mostly agree	I agree
I think it is important that the university fights plagiarism and cheating, since all grades should reflect the real knowledge of the student	–	3	9	15	24
When doing the homework, I personally witnessed cases of plagiarism (for example, homework being copied and pasted from someone else's work)	24	11	12	1	3

(*continued*)

fridaya@uef.fi

Table 8. (*continued*)

Questions	I disagree	I mostly disagree	I neither agree nor disagree	I mostly agree	I agree
When taking the exam, I personally witnessed cases of cheating	36	3	6	4	2
It was hard to cheat during the exam	2	7	12	13	17
Conducting exam on computers with Internet connection makes cheating easier (for example, receiving illegal help from other persons)	2	7	17	13	12

As can be seen from Table 8, the majority of respondents (39 students or 76.5%) think that plagiarism and cheating are not acceptable and the university should fight these illegal practices, and only 3 respondents (5.9%) disagree with that. One disagreeing respondent provided a comment in the survey, arguing that cheating during exams should be accepted since students might never face challenges that are similar to exam assignments during their professional careers. Although 44 respondents (86.3%) believed that at least 10% of students were involved in plagiarism or cheating, only 6 respondents (11.8%) witnessed them personally during the exam, while 4 respondents (7.8%) witnessed them while working on homework assignments. This suggests that the number of actual cases of academic misconduct is smaller than many respondents believe. This view is partly supported by the fact that according to only 9 respondents (17.6%), cheating was easy or relatively easy during the exam.

In order to counter academic misconduct during the exam, we have introduced a paper-and-pencil based exam format since 2016. According to Table 8, almost half of the respondents (25 students) believe that computer-based exam makes cheating easier. Interestingly enough, this opinion is stronger among respondents from 2016–2017 - if the average opinion for years 2011–2015 ranged from 3.12 to 3.50 (on the scale of 1..5), for years 2016 and 2017 the average opinions were 3.70 and 4.83, respectively. Moreover, during the focus group interview with 5 students from 2017, students strongly suggested continuing with the paper-based exam as it allows to prevent cheating efficiently. When the new examination format was announced in the fall semester of 2016, several students expressed an apprehension that the exam might be more difficult to pass, resulting in noticeably lower grades. However, actual examination results do not support this concern - in 2014 and 2015, average grades from computer-based exam were 3.34 and 3.31, while average grades for 2016 and 2017 were 3.12 and 3.35, respectively. Other researchers have made a similar observation that results from computer-based and paper-based exams will not vary [12]. Also, since homework assignments have been designed for a computer-based setting and they contribute a significant part of the final grade, we argue that the evaluation is done in a fairly balanced way.

During the focus group interview, a subtle issue concerning evaluation fairness was raised by participants (the same issue was also brought up by one respondent who took

fridaya@uef.fi

the electronic survey). The homework assignments are tackled by groups of 2–3 students, in order to facilitate active learning through student collaboration. However, if a member of a group does not participate in the group work, he/she would nevertheless get points for the homework solution, eventually receiving a grade which does not correspond to his/her real contribution. While homework that is plagiarized from other student groups are easy to detect due to their similarity to other submissions, the focus group members suggested introducing a method for preventing such form of plagiarism within student groups.

For collecting student feedback on how to improve the learning efficiency and evaluation fairness during future iterations of the course, the electronic survey contained several questions which allowed to provide free-form answers:

- What recommendations would you give about the course content?
- What recommendations would you give about homework and exam?
- Is there any other information or feedback you would like to provide?

In addition, we also asked a number of questions from the focus group, in order to establish their opinion about the current quality level of the course, the use of various e-learning techniques, evaluation methods, improvements on course content, etc. During the focus group interview and electronic survey, respondents identified several aspects they liked about the course:

- Course slides are very detailed and facilitate independent learning
- Course virtual machine is well prepared and does not need any additional tuning from students. The virtual machine has been designed to work on all operating system platforms, and can be used outside the university (e.g., on personal laptops or home computers)
- During the course, tutors provided detailed and timely feedback about lab, homework, and exam assignments
- Paper-and-pencil based exam works best for preventing academic misconduct during the exam, and this examination format should not be changed

When analyzing answers from electronic survey and focus group, we identified the following major suggestions for improving the course:

- Although the current homework and lab assignments are efficient for preparing the students for the exam, the labs should contain additional graded assignments, since it would motivate students to work continuously throughout the semester, and would thus improve knowledge retention. This proposition was made by the focus group of 5 students, and a similar feedback was also provided by 7 other students who took the electronic survey.
- When taking the electronic survey, several respondents recommended using additional graded tests with multiple-choice answers that would contribute to the final grade. These tests could be part of the exam or arranged during labs and lectures.
- During lectures, interaction with students should be increased. In particular, students from focus group recommended introducing more quizzes into each lecture where student feedback is instantly collected and displayed to the audience.

fridaya@uef.fi

- For most complex topics of the course, e-learning materials should be prepared (such as 5–10 min videos with quizzes). However, when we asked the focus group whether traditional lectures should be replaced with pre-recorded lecture material, students were not favouring this idea and recommended to keep the current lecture-based course format. According to students, the primary advantage of lectures is an opportunity to ask questions from the lecturer with instant feedback and start discussions on more complex lecture topics.
- The course should contain live demos from security monitoring production systems, in order to illustrate how the concepts and solutions taught during the coursework in practice.

3.4 Planned Improvements for the Course

In order to implement these major suggestions for future course iterations, we are planning to introduce a number of changes to the course. Firstly, we are planning to introduce more quizzes into lectures with instant feedback collection, in order to facilitate discussions during the lecture and measure student progress. Secondly, we will start preparing 5–10 min e-learning modules for the most complex topics of the course (such as the computational complexity of regular expressions and techniques for avoiding exponential complexity).

E-learning modules would be particularly useful for highly theoretical and hard to understand topics. From the implementation point of view, students will be able to access a dedicated website remotely and follow pre-recorded video material and interact with the content.

Instructors will record the lectures using;

- Desktop recording via software solutions like Camtasia, FinalCut Pro and Quicktime
- Talking head recording via webcams or dedicated cameras

Once the material is ready, e-learning solutions, such as Learning Management Systems (LMS) will provide easy to access, easy to use learning dashboard for each student. A popular and effective way of doing this is via LearnDash [23]. This enables instructors to quickly set up an environment with features like quizzing, content drip feed, lesson timers, certificates and badges. This will help students to stay focused and motivated as well; since they will have the flexibility to follow the course material to understand the content clearly, create their own time plans to complete the course, challenge themselves with quizzes and earn badges and certificates which will simulate a gamified environment to make it more interesting and fun.

We are also planning to introduce graded tests into labs, but this introduces one subtle issue. Currently, the course has been designed to facilitate both on-site and distance learning, and attending lectures and labs is not compulsory. All graded homework that yields up to 40 points are announced 4–5 weeks before the deadline, and they are submitted via e-mail without an oral defence. The only part of the course which requires the physical presence of the student is the final exam that lasts for 3 h and yields 60 points. Introducing compulsory graded tests into labs would, therefore, prevent distance learning students from attaining the maximum score of 100 points

fridaya@uef.fi

during the course. In order to solve this problem, we are planning to implement simple graded tests with multiple-choice answers and short time limit (e.g., 2–3 questions with a time limit of 2–3 min), so that tests would be voluntary and would add 4–5 extra points to 100 points from regular homework and exam. Therefore, distance learning students would not be penalized, since the final grade would still be calculated according to rules from Table 2 (e.g., if a distance learning student would collect 91 points and a regular student 105 points, they would both pass the course with grade 5). However, this evaluation scheme creates an additional issue - since homework for student groups and graded tests would yield 45 points in total, it is possible to pass the course with a positive grade 1 by collecting just 6 points from the exam (i.e., 10% of the exam point score). Also, during the focus group interview, the problem of homework plagiarism within student groups was identified which makes the aforementioned problem even worse. In order to achieve more balanced and fair evaluation, we are considering to change the weight of homework and individual exam to the final grade, so that the exam would contribute 75 points and homework 25 points. In addition, all homework submissions must provide a detailed description of the contribution of each member of the student group.

One of the major recommendations from past students was to include demos of live security monitoring systems from real-life organizations into the course agenda. Unfortunately, using live demos from external organizations (e.g., private companies and governmental institutions) would assume data sharing contracts and non-disclosure agreements, but external organizations are rarely willing to share their confidential security data with universities, especially for teaching purposes which involve wider audiences. However, we are planning to implement a security monitoring system for the university network, and using this system for some live demos might be a valuable addition to the course content.

4 Conclusion

In this paper, we have presented a case study of redesigning an MSc level cybersecurity course and investigated if lecture-based teaching methodology is suitable for teaching a highly technical cyber security course. In our case study, we have investigated how efficient past blended learning based course iterations have been and whether the blended learning based approach should be discontinued by replacing all lectures with online materials. We have also investigated whether the current student evaluation methods are fair and effective for preventing academic misconduct. For conducting the study, we have collected feedback from past students of 2011–2017 time frame. Based on the collected data, we are planning to redesign the future iterations of the course. Our findings can be useful for redesigning similar courses in other educational institutions.

fridaya@uef.fi

References

1. Adam, D., Kioutsiouki, D., Karakostas, A., Demetriadis, S.N.: Do your students get it? Quiz it! The android classroom response system. In: IEEE 14th International Conference on Advanced Learning Technologies, pp. 168–170. IEEE (2014)
2. Akakura, T., Kawamata, T., Kato, K.: Development of a blended learning system for engineering students studying intellectual property law, and an analysis of the relationship between system usage and the knowledge acquisition process. In: IEEE 6th International Conference on Teaching, Assessment, and Learning for Engineering (TALE), pp. 114–117. IEEE (2017)
3. Bajwa, W.U.: Flipping large classes on a shoestring budget. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 7006–7010. IEEE (2018)
4. Brandsteidl, M., Mayerhofer, T., Seidl, M., Huemer, C.: Replacing traditional classroom lectures with lecture videos: an experience report. In: Proceedings of the 8th edition of the Educators' Symposium, pp. 21–27. ACM (2012)
5. Grotz, N.: Fair exams in split groups: (Implementation of equitable computer-based exams for large groups in small test centers). In: The Global Engineering Education Conference (EDUCON), pp. 957–962. IEEE (2018)
6. Guy, R., Lownes-Jackson, M.: Web-based tutorials and traditional face-to-face lectures: a comparative analysis of student performance. In: Proceedings of the Informing Science and Information Technology Education Conference, Informing Science Institute (2013)
7. Hammer, S., Hobelsberger, M., Braun, G.: Using laboratory examination to assess computer programming competences: questionnaire-based evaluation of the impact on students. In: The Global Engineering Education Conference (EDUCON), pp. 355–363. IEEE (2018)
8. Hui, W., Hu, P.H., Clark, T.H., Tam, K.Y., Milton, J.: Technology-assisted learning: a longitudinal field study of knowledge category, learning effectiveness and satisfaction in language learning. J. Comput. Assist. Learn. **24**(3), 245–259 (2008)
9. Jackowska-Strumiłło, L., Nowakowski, J., Strumiłło, P., Tomczak, P.: Interactive question-based learning methodology and clickers: fundamentals of computer science course case study. In: The 6th International Conference on Human System Interaction (HSI), pp. 439–442. IEEE (2013)
10. LoPresto, M.C., Slater, T.F.: A new comparison of active learning strategies to traditional lectures for teaching college astronomy. J. Astron. Earth Sci. Educ. **3**(1), 59–76 (2016)
11. Miller, C.J., McNear, J., Metz, M.J.: A comparison of traditional and engaging lecture methods in a large, professional-level course. Adv. Physiol. Educ. **37**(4), 347–355 (2013)
12. Özalp-Yaman, Ş., Çağiltay, N.E.: Paper-based versus computer-based testing in engineering education. In: Education Engineering (EDUCON), pp. 1631–1637. IEEE (2010)
13. Schmidt, B.: Improving motivation and learning outcome in a flipped classroom environment. In: 2014 International Conference on Interactive Collaborative Learning (ICL), pp. 689–690. IEEE (2014)
14. Seeling, P., Eickholt, J.: Levels of active learning in programming skill acquisition: from lecture to active learning rooms. In: 2017 IEEE Frontiers in Education Conference (FIE), pp. 1–5. IEEE (2017)
15. Shryock, K.J.: Engaging students inside the classroom to increase learning. In: Frontiers in Education Conference (FIE), pp. 1–7. IEEE (2015)
16. Winterstein, T., Greiner, F., Schlaak, H.F., Pullich, L.: A blended-learning concept for basic lectures in electrical engineering: a practical report. In: 2012 International Conference on Education and e-Learning Innovations (ICEELI), pp. 1–4. IEEE (2012)

fridaya@uef.fi

17. Zhang, D.: Interactive multimedia-based e-learning: a study of effectiveness. Am. J. Distance Educ. **19**(3), 149–162 (2005)
18. Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R.: Taxonomy of educational objectives: the classification of educational goals: handbook I: cognitive domain (No. 373.19 C734t). New York, USA (1956)
19. Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., Wittrock, M.C.: A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, abridged edition. White Plains, New York (2001)
20. Moses, K.V., Petullo, W.M.: Teaching computer security (2014)
21. Vasserman, E.Y., Bell, R.S., Sayre, E.C.: Developing and Piloting a Quantitative Assessment Tool for Cybersecurity Courses. American Society for Engineering Education (2015)
22. Merrill, M.D.: First principles of instruction. Educ. Technol. Res. Dev. **50**(3), 43–59 (2002)
23. LearnDash WordPress LMS Plugin. https://www.learndash.com. Accessed 12 Oct 2018
24. Rsyslog home page. https://www.rsyslog.com. Accessed 25 Nov 2018
25. Netfilter home page. https://www.netfilter.org. Accessed 25 Nov 2018
26. Hansen, S.E., Atkins, E.T.: Automated system monitoring and notification with Swatch. In: USENIX 7th System Administration Conference, pp. 145–152 (1993)
27. Syslog-ng home page. https://www.syslog-ng.com. Accessed 25 Nov 2018
28. Vaarandi, R., Blumbergs, B., Çalışkan, E.: Simple event correlator – best practices for creating scalable configurations. In: 2015 IEEE CogSIMA Conference, pp. 96–100. IEEE (2015)
29. Snort home page. https://www.snort.org. Accessed 25 Nov 2018

fridaya@uef.fi

Multi-Confirmations and DNS Graph Mining for Malicious Domain Detection

Hau Tran[1(✉)], Chuong Dang[1], Hieu Nguyen[1], Phuong Vo[1(✉)], and Tu Vu[2]

[1] Data Scientist, FPT Telecom, Ho Chi Minh, Vietnam
{HauTX2,ChuongDM,HieuNT124,PhuongVTH}@fpt.com.vn
[2] EVP FPT-Telecom, Ho Chi Minh, Vietnam
tuva@fpt.com.vn

Abstract. The Internet criminals use domain name service (DNS) as cornerstone for variety of malicious activities. Typically, attackers register domains name to locate their *C&C* (command and control servers) for spam mails, ddos attacks, etc. There are thousands of cyber attacks happening every day and the number of those increase with exponential growth, which becomes an urgent matter for cyber security. To tackle this issue, several recent techniques have been proposed to identify malicious domains. Most of them use DNS data which contains much of global information cannot be controlled by attackers. Hence, using labeled domains to identify malicious domains relying on their association by analyzing history log is reasonable. Besides, the graph is usually used to present the relationship between objects and many powerful algorithms have been successfully implemented on graph in literature. So, a method with modified version of the Belief-Propagation algorithm to detect malicious domains based on a graph of domains and hosts IPs were proposed in our previous paper. In addition, graph inference is also applied with many different methods with very impressive results. However, these methods cannot handle all domains on the graph because the domains blacklist does not cover all graph components. This can be considered a bottle neck of the methods proposed. To solve this problem, we propose a new method by extending the previous work. This paper proposes a new approach which employs the **law of total probability** based on relationship between domains and clients to calculate malicious score of domains. This paper not only uses new score to label for domain but also combines with the previous score. The scores from the two methods are combined via a weighted sum to arrive at a score for each domain. The study shows high performance in malicious domain prediction on test case with accuracy, precision and recall as: **97.6%**, **99.2%** and **93.1%**. This technique can discover huge potential malicious domains by mining on small knowledge data (black and white list domain) with highly precision. It indicates that the proposal technique is strongly effective in detecting malicious domains.

Keywords: Malware detection · Domain resolution · Graph mining · Belief propagation · Cyber security · Graph interface · Social network

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 639–653, 2019.
https://doi.org/10.1007/978-3-030-22868-2_46

fridaya@uef.fi

1 Introduction

The World Wide Web is the platform for a myriad of beneficial Internet activities. Unfortunately, it is also vulnerable to various Internet crimes, such as commercial scams, financial fraud etc. While these schemes may differ in motivations, the common thread among them is the requirement that unsuspecting users visit malicious sites. These visits can be driven by emails, Web search results or links from other Web pages or some tricks requiring the users to click in the desired Uniform Resource Locator (URL). Clearly, if one could inform users beforehand that particular URL is dangerous to visit, much of this problem could be alleviated. Many methods are raised to solve this problem as developing blacklisting services encapsulated in tool bars, appliances and search engines for warning users about these dangers. These blacklists are in turn constructed by a range of techniques including manual reporting, and Web crawlers with site analysis heuristics. However, many malicious sites are not in blacklist either because they are too new, not evaluated, or evaluated incorrectly. Facing with this problem, some client-side systems analyze the content or behavior of a website as it is visited.

As a contribution to solve this problem, we propose a method to detect malicious domains by combining graph mining with a statistical approach. In our previous paper, we renovate on the **Belief-Propagation algorithm** to perform graph mining on the graph of domains and host IPs. Despite the model's high precision in predicting malicious domains, a shortcoming of the model is its inability to detect malicious domains in graph components without labelled malicious domains. This is a significant issue because there exists numerous subgraphs with no domains knowledge which are effectively removed from the graph-mining algorithm. Extending our previous paper, this work resolves this issue by adding clients information to calculate the malicious probability of domains with a new statistical method. A combination of the two scores given by the previous and new method will then give a strong confirmation as to whether a domain is malicious, while not limited to only predict malicious domains in graph components with labeled malicious domains. Indeed, it is encouraging to see that: in testing the method's performance, while there is a mild decrease in precision, the number of domains the model can predict increases by more than 10 times with the corresponding performance index as: **97.6%** accuracy, precision and recall rates: **99.2%** and **93.1%**.

The rest of this paper is organized as follows: Sect. 2 describes related works. Section 3 explains about graph interface, graph mining technique and statistic approach. In Sect. 4, evaluation of the proposed method is explained and comparison results of this proposed method in many scenarios. In final section, we give some conclusions and range effect of the method in real work.

2 Literature Review

The Internet increase with exponential growth and cyber attack are also various range. So, malicious domains detection is highly raising with many techniques

fridaya@uef.fi

in the literature, utilizing different types of data and analysis techniques. Here we only discuss the representative works that are most relevant to our approach.

Manos Antonakakis proposed a study [2] which is one of the first studies used passive DNS data to identify malicious domains. His method assigned reputation scores of unknown domains given three features: network addresses, DNS zones, and malicious clues. Fast-Flux has been used by botnets to build malicious content distribution [3,4]. Domain Flux malware can generate a large number of pseudo random domain names in order to escape domain blacklists [5]. EXPOSURE [7] recognizes malicious domains such as malware, phishing site, by a classification algorithm based on DNS traffic features (including query time, response content, TTL, and domain). [5] took advantage of the information of upper authoritative domain server to mine malicious domain based on the distribution. [8] developed a methodology to detect such "domain fluxes" in DNS traffic by looking for patterns inherent to domain names that are generated algorithmically. The graph inference was introduced in [9] and the authors proposed a message method to find out malicious domains with high very accuracy. In the paper [10] details the design and results of long-running experiments that identify maliciousness in popular websites in a vulnerability and exploit-independent manner. [11] explained the principles behind the belief propagation (BP) algorithm, which is an efficient way to solve inference problems based on passing local messages. In [12] designed a method to scoring the malicious probability of domains. He employed passive DNS data to build a domain resolution graph. Based on the graph, any path link from unknown domains to malicious domains, a mathematical formula was constructed to scoring malicious probability for unknown domains. That method is very flexible to facing with small labeled data, it only care if there are existed a path from a domain to malicious then the malicious score would be computed. Author in [13] used the co-occurrence relation of two different domain names to find unknown black domain names and extend the blacklist. If a domain name frequently co-occurs with a known black name, they assume that the domain name is also black.

Compare with the aforementioned approaches [2–7], the detection method proposed in this study has a flexible private detection. The associations between nodes in DNS graph is built through DNS data and relies on the association of all them to compute the reputation score to verify each node in graph is legitimate or not. As introduced in [9,11,12], this approach is also based on domain resolution graph but considering about Markov Random Filed (MRF) and employ a developing version of belief propagation algorithm to get malicious score of domains. Beside that law of total probability is also used to expending blacklist domain.

3 The Proposal Method

The paper is implemented on passive DNS log which contains rich information of various aspects. However, to achieve our target, we only use three fields: domain name, host answer IP and client IP address. Domains and answer ips are used for

fridaya@uef.fi

graph mining while domains and client-IPs are to calculate malicious probability of domain by **law of total probability** method. Because the probability of malicious domains will be computed rely on features of client set who query to domains, we call that method is **Multi-Confirmations algorithm** and for convenience, it is written in brief form as **MC algorithm**.

3.1 Graph Building

This study uses bipartite graphs to present relationship between objects and we present two independent methods and combine at the end. Therefore, this study build two graphs to support for mining malicious domains: The first graph presents for relationship between domains and answer host IPs, this graph is used for graph mining with the **BP algorithm** to compute the reputation score of each node on the graph. The second graph presents the relationship between domains and client IPs and is used to calculate malicious probability of each domain by applying the **MC algorithm**. There are two graphs with two independent methods. Consequently, we have two scores for the domains that need labeling. Hence, we need a method to combine the two scores, that method is presented in final score estimation section. And now, the graph building is presented.

As mentioned above, this study retains all properties of the graph mining model (the **BP algorithm**) as a part of the new proposal method. Therefore, this study strongly focuses on the extending part and the combination method at the end. Thus, this study just briefly introduces the DNS graph (graph of domains and host answer IPs) and for more details the reader can refer to the paper [1]. In paper [1], to do graph mining model, we just focus on analyzing **A-records** (A record maps a domain name to the IPv4 address) in DNS data and only use two columns which are formed as (d, IP). The BP graph is built in bipartite form where one side is the domains and the other one is the answer host IPs. A sketch of graph interface is presented as Fig. 1.

The second graph is used for **MC algorithm** is described in detail as follow. First, we note that the second graph is named as the MC graph in the below contents for convenience. Like the DNS graph, the MC graph is also a bipartite graph, it presents the relationship between domains and client IPs. To build the MC graph, we do not do any filtering. Instead, all records are loaded and duplication of two columns domain names and client IPs are removed. An example of graph visualizing with only the connection between domain names and client IPs is showed as Fig. 2. It is a bipartite graph with one side corresponding to domains and the other side to client IPs. An edge is formed from a domain to a client IP if there is exist any request from client to domain.

3.2 Algorithm Development

Developing from the previous paper, this paper not only uses relationship between domains-host answer IPs but also the relationship between domains

fridaya@uef.fi

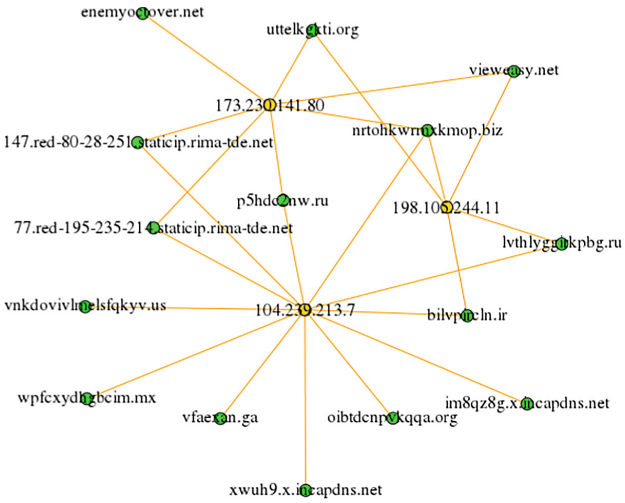


Fig. 1. The BP graph

- client *IPs*. The malicious scores are computed for the domains by two methods and finally we estimate weights for each score to classify domain labels. In this study, we are reproducing all properties of the belief propagation algorithm in the paper [1] to get first score for domains and using law of total probability rely on the relationship between domains - client to calculate second score. In the following sections, we call **Multi-Confirmations method** as mention to the second method, abbreviated as **MC method**.

3.2.1 The BP Algorithm

Also known as *sum-product message passing*, the Belief-Propagation algorithm is a popular graph-inference algorithm designed to transfer information between objects on an undirected graph obeying *Markov properties*. In this graph, each vertex corresponds to a random variable following Markov properties. In other words, a node on the graph only depends on properties of its adjacent nodes. For malicious domain mining, some nodes of graph are labeled as malicious or white by using the labelled dataset. Meanwhile, adjacent nodes interact with each other via an inference function in order to update each other's belief. Then, the reputation scores of all nodes are computed using the BP algorithm. This study reproduces the BP algorithm and also reapply the proposals as raised by the authors of the paper [1]. For a detailed description of the BP algorithm and the improvement made to the method, please refer to the paper [1].

fridaya@uef.fi

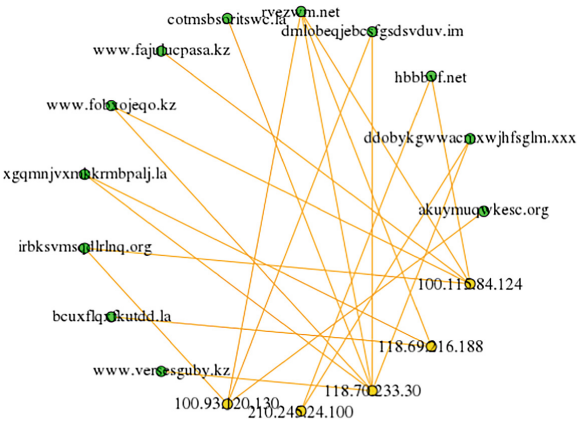


Fig. 2. The MC graph

3.2.2 Muti-Confirmations Algorithm

As the name of algorithm, our algorithm will make decision by information of objects relationship. In this study, the relationship between domain and client will be considered to make decision about labels of domains.

3.2.2.1 Relationship Inference: As described in the **Graph building** section, this section we describe the MC graph along with functional for convenience during the method introduced below. Unlike the BP graph which uses only the **A-records**, the MC graph makes use of all types of records in the DNS log. To present the relationship between domains and clients, the bipartite graph is formed of two groups of vertices: the domain collection $D = \{d_1, d_2, ..., d_n\}$ and the client collection $C = \{c_1, c_2, ..., c_m\}$. Then the edge collection is defined as $E = \{e_{ij} \quad | \quad i \in (1, n) \quad \& \quad j \in (1, m)\}$. An edge e_{ij} is established if there exists a request from client c_j to domain d_i. In the end, we obtain the undirected graph $G_{dc} = (V, E)$ where V is the vertex collection of domains and clients and E is the edge collection presenting the relationship of domains and clients. As with the above BP model, this graph is used for statistical inference to determine malicious domains, the details of which is presented in the next section.

3.2.2.2 Statistical Inference: We observe that infected clients or attackers tend to send requests to malicious domains, for example in order to communicate with a bot master or perform command and control ($C\&C$) etc. Therefore, a domain frequently requested by clients who have queried to other malicious domains is more likely to be itself malicious. This insight motivates our idea of computing the malicious probability of a domain given the set of clients who request to that domain. Using **the law of total probability**, we devise the method to compute this probability as follows:

(s1) First of all, we need to calculate probability of a request to domain d_k is sent by client c_i^k. To calculate this probability, client set and request set are

fridaya@uef.fi

defined. The client set $C_k = \{c_1^k, c_2^k, ..., c_n^k\}$ is set of all of clients who queried to domain d_k and another one set $R_k = \{r_1^k, r_2^k, ..., r_n^k\}$ is set of number of request sent to domain d_k by client $c_i^k \in C_k$. And then the probability $\mathbb{P}(c_i^k)$ of a query which sent by client c_i^k to domain d_k is defined as

$$\mathbb{P}(c_i^k) = \frac{r_i^k}{\sum_{t=1}^{n} r_t^k} \tag{1}$$

(s2) Secondly, the malicious probability of a domain d_k if we know that the domain d_k is requested by client c_i^k. This probability will be computed based on the collected knowledge domain. So, let $c_i^k \in C_k$ the condition probability $\mathbb{P}(d_k|c_i^k)$ is computed as.

$$\mathbb{P}(d_k|c_i^k) = \frac{N_m}{N_m + N_l} \tag{2}$$

Where N_m is the number queries of client c_i^k towards malicious domains and N_l is number queries of client c_i towards legitimate domains.

(s3) Finally, we apply law of total probability to calculate malicious probability of domains. Let d_k is domain queried by client set $C = \{c_1^k, c_2^k, ..., c_n^k\}$, then malicious probability of domain d_k is calculated as

$$\mathbb{P}(d_k) = \sum_{i=1}^{n} \mathbb{P}(c_i^k)\mathbb{P}(d_k|c_i^k) \tag{3}$$

as defined above $\mathbb{P}(d_k|c_i^k)$ is probability of domain d_k is malicious domain if it is requested by client c_i^k and $\mathbb{P}(c_i^k)$ is probability of a query to domain d_k is sent by client c_i^k.

3.3 Final Score Estimation

As mentioned, there are two scores given by two independent methods to assess domain's labels which we need combine for final score. For convenience, denote the score by the BP algorithm and the MC method as S_1 and S_2 respectively. We recall that in the BP algorithm, some domains are not labelled and so have a missing S_1 score, either because they are not in the domains - host IPs graph or they are in a graph component with no known labeled domains (and so cannot be inferred). The both scores are in range $[0, 1]$, for those with missing S_1 score, we can assume they have $S_1 = 0.5$. In the end, for the final score, we calculate as follow function:

$$\text{Final Score} = \lambda S_1 + (1 - \lambda)S_2 \tag{4}$$

The λ parameter is evaluated from the precisions of the two methods by the following function:

$$\lambda = \frac{P_{BP}}{P_{BP} + P_{MC}} \tag{5}$$

Where P_{BP} is precision of the BP algorithm and P_{MC} is precision of the MC method. The λ parameter is in range $[0, 1]$ and it presents the dependent of final score to the precision of algorithm which higher precision.

fridaya@uef.fi

4 Experiments

This section details the train and test datasets along with results evaluation. Differently from the previous method, the final result is combined from BP method and MC method. Currently we can not integrate the two methods explicitly, thus this study presents the results of two methods separately and the combined result is given at the end.

4.1 Train and Test Datasets

In our dataset, there are one million white domains which have been collected from top ranking on www.alexa.com and the malicious domains have been collected from four sources:

(1) urlblacklist.com
(2) malwaredomains.lehigh.edu
(3) osint.bambenekconsulting.com/feeds/dga-feed.txt
(4) malwares generation and some malicious domains collected from the authors of the papers with related work.

The final domains set then contains 31, 436 unique malicious domains in our DNS data. The dataset is divided into two parts, one for training model and the other for testing.

4.2 Innovations

The graph mining method is dealing with a problem which sub-graphs without knowledge. The domains on these sub-graphs will be not scoring then many domains are missing class if we just apply graph mining based on domains and host answers. Therefore, beside improvements the BP algorithm to apply graph mining, we also do many study to solve the above problem (many domains can not classify). We see that rely on relationship between domains and clients most domains can not be labeled by graph mining will be labeled. The detail of graph mining problems, improvements and problems solved are presenting in detail as follows:

The graph mining problems and the BP improvements will be reminded first as:

1. Although the graph mining apply the BP method give us the outstanding performance results in the paper [1] but there exists some issues that need solving. The DNS graph is built from DNS log with **A-records** and return codes are **NOERROR**. It means that many DNS records as DGA domains (domain generation algorithm) are not labeled because most of them are marked as **NXDOMAIN** records. In addition, not all nodes in the DNS graph itself can connect together to form large components. Therefore the DNS graph consists of a lot of sub-graphs many of which we don't have any

fridaya@uef.fi

knowledge about regarding malicious and legitimate domains. These subgraphs have no effect on predicting domains, thus all domains in these subgraphs are effectively removed before the graph mining algorithm.
2. The general of BP algorithm don't care about vertices without knowledge during each interactive but when we study about BP algorithm to apply for our problem, we see that the message of unknown label node will be not meaning because in our context unknown nodes have malicious probability as 0.5 then the message does not make change malicious probability of itself neighbours.
3. Initially, the known labeled nodes start with the highest malicious probability for malicious domains and the lowest probability for white domains. These probabilities should not change after each iteration of the BP algorithm since the labels are known and should not change regardless of the neighbor nodes' beliefs. Therefore, the malicious probability of the labeled nodes is kept the same as when initialized.

Of the three issues, the latter two have been solved by modifications to the BP algorithm as in the paper [1]. As for scoring domains which we cannot label by graph mining, we propose a method based on the relationship between domains and clients. A popular domain is known by many clients most of whom are normal clients, on the other hand, a malicious domain is not popular and is queried to by clients if they are infested malware or attackers. Hence, it is reasonable to label a domain if we know about properties of clients who query to the domain. The malicious probability of domains are calculated using clients properties, we call the method as multi-confirmations (MC) method. Although there are some domains which come out of knowledge label, this mean that client set of the domains don't have any query to knowledge domain, then these client do not have knowledge label to confirm about the label of domain. But during labeling processing we don't care about return code or any conditions, that make the MC method can scoring for more wide range of domains.

We just explain about some adjustments of the BP method and an idea to solve the problems not yet solved by the BP method. Next, we present the results of malicious domain detection in three case: the result of BP method, results of MC method and finally we will present the result of the combine version of BP & MC method.

4.3 The Results

This section summarizes the results of malicious domain detection in three contexts of model and then explains why we decide to extend our graph model with the MC model. As we said in the paper [1] "DNS graph mining for malicious domain detection", the DNS graph mining based on domain names and host answers achieved significant results with precision and recall over 95% and some comparisons with results of graph mining methods in our context. That method is applying for daily prediction and send alerts to our customers. Here we show a sample daily predict performance again to compare with the method use in this paper and supporting for explain the proposal method.

fridaya@uef.fi

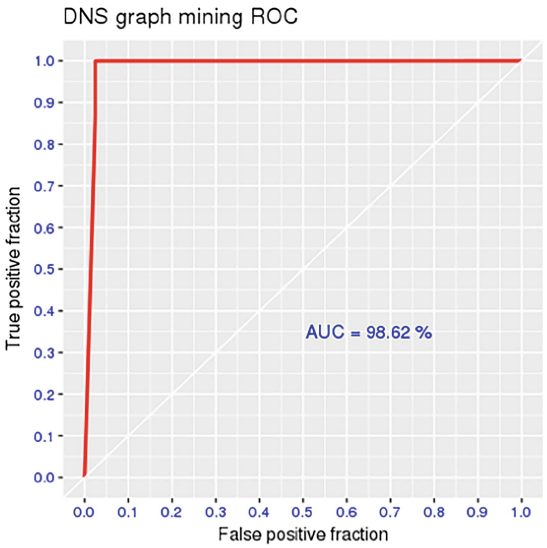


Fig. 3. ROC curve of DNS graph minig model (BP model)

From Fig. 3, the BP ROC shows significant AUC: it is over 98%, which is perfect to show the efficiency of the BP model in determining malicious domains. But in this paper, we want to show two more tables - confusion matrix and the performance indices of the BP method to compare with the proposal method. Through this comparison, we justify our decision to extend our previous method.

Table 1. The BP confusion matrix

		Predicted	
		Malicious	white
Actual	Malicious	778	17
	White	26	14443

Table 2. The BP Performance indexes

Performance ratios	
Precision	97.9%
Recall	96.7%
Accuracy	99.5%

From the performance indexes in Table 2 of the BP method, we can see the significant effect of this method when applied for malicious domain detection.

fridaya@uef.fi

However, at graph building step, the data is filtered with many conditions e.g. it must be $A - record$, answers ip are in ip form and it is different to 127.0.0.1, etc. So, these filters make the model miss many domains in prediction. The fact that there are so many malicious domain names not predicted by the BP method compels us to research new ways to help us handle more malicious domain names while keeping the performance indices competitive. After considering many features in dns data, we see that there is a certain relationship between the label of the domain and the customers who request to that domain. It is easy to see that if a domain is visited by many clients then the malicious probability will be less than the domains that are accessed by few clients. If a customer accessed to malicious domain names in the past then the probability of malicious access to malicious domains would be higher.

Based on that idea, we develop a method call that **Multi-confirmations** (MC) rely on law of total probability. This method can handle all most domain names in dns data. The results of MC method is showed as follows:

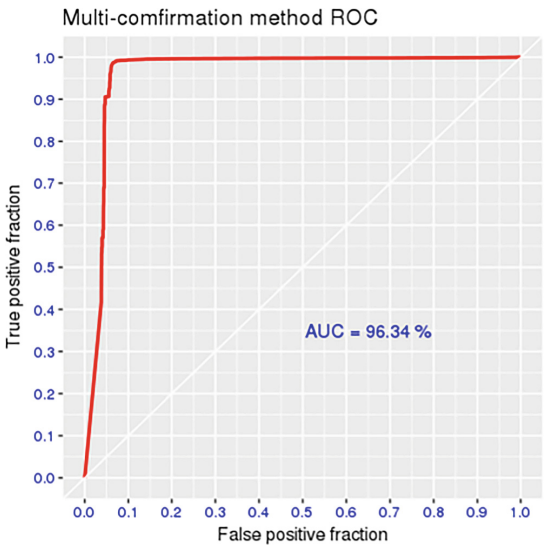


Fig. 4. ROC curve of Multi-Confirmations (MC) algorithm.

From Fig. 4, we can see that the AUC is 96.34% for the MC method which decreases slightly compared to that of the BP method. In our context, it is highly acceptable because the new method, while having a slightly lower precision, can recover significantly many more domains. In our experiment, the number of domains that the MC algorithm can predict increases by 10 times compared to the BP method. The following tables can demonstrate the differences between the MC method and the BP method: Comparing Tables 1 and 3, we can see that the BP method discovers only 778 malicious domains from the test dataset

fridaya@uef.fi

whereas the MC method can predict 9, 405 malicious domains. Additionally, the Table 4 shows us the precision has not changed much. In order to take advantage of both methods, we combine both results from these two methods to produce the final forecast. Indeed, the accuracy of the forecast increases and so does the number of domains which can be predicted, as shown in the following metric evaluations.

Table 3. The MC confusion matrix

		Predicted	
		Malicious	White
Actual	Malicious	9405	662
	White	237	16923

Table 4. The MC Performance indexes

Performance ratios	
Precision	97.5%
Recall	93.4%
Accuracy	96.9%

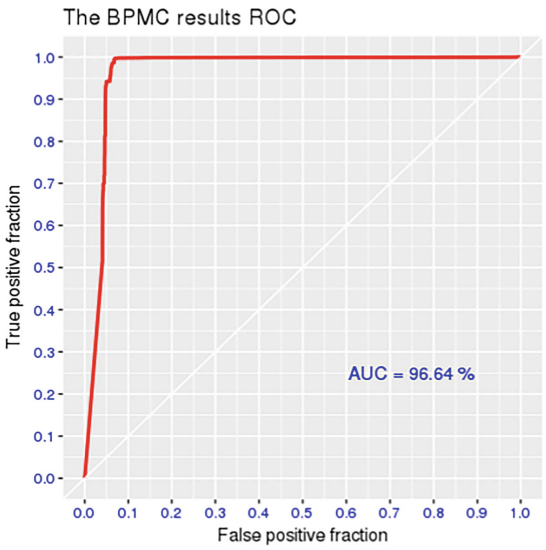


Fig. 5. ROC curve of BPMC algorithm.

fridaya@uef.fi

The Fig. 5 shows a small increase in the AUC of the combined version compared to the AUC of the MC method. As you see, in the Table 5 show increasing of number of malicious domain which predictable and another thing which we can easy to see that the number false positive is small. A reason for the small number of false positive is because we think that when switching algorithms we can predict many malicious domains and these domains will be used for warning to our customer on the system. Therefore, the number of false alerts should be reduced to the lowest possible. To do that we choose the threshold to classify domain label is **0.55**. The domain with final score higher the threshold will be labeled as malicious domain and otherwise for white domain. Apply this threshold to final test phase, we get **99.2%** precision and it's showed in Table 6.

Table 5. The BPMC confusion matrix

		Predicted	
		Malicious	White
Actual	Malicious	9431	698
	White	75	22406

Table 6. The BPMC Performance indexes

Performance ratios	
Precision	99.2%
Recall	93.1%
Accuracy	97.6%

From the above results of malicious domain detection with three phases, we see that our proposal method can solve two problems, the new method can labelling for many domain which can not predict by the previous model (the BP method), another one is also said above that we can change the threshold to increasing the precision and using the malicious domains which are labelled by the proposal method to expanding our blacklist domain.

5 Conclusion

This study proposes a new method for mining malicious domain based on two relationship $domains - clients$ to do multi-confirmations algorithm and combine with the BP algorithm to hold the accuracy of the final result. The method is scalable and effective for malware detection, botnet detection, etc. It has the following implications:

1. A new approach is proposed to solve issues not handled by the BP algorithm-based graph mining in the previous version [1].

fridaya@uef.fi

2. The approach achieved significant results. It is good technique for some classification problems which builds on relationship between objects. Applied on real-world DNS traffic, the algorithm has achieved a precision rate of **99.2%** with a recall rate of **93.1%**. This result shows that the approach is effective for malicious domain detection in practice.
3. The method detects not only malicious domains but also C&C server or the victim clients as well. It is also easy to do statistical analysis of malicious infection in different aspects, which is more insightful.
4. This method does not tie to specific DNS technique. Therefore, in general it can be applied for detecting various types of malicious sources such as: DGA, malware, botnet, etc.
5. This study gives a new method to expand range of malicious domain. While our primary focus is on malicious domain prediction, since our approach is based on the relationships of **(domains, host answer ips)** and **(domains, ip clients)**, the method can flexibly be extended to areas such as botnet detection, *C&C* servers, etc.

In summary, the experiment and positive results with this work have demonstrated one promising solution to the problem of detecting malicious domain. Currently, this study just focuses on malicious domains. In the future, we can extend our algorithm with smoother combination of the BP algorithm and the MC algorithm and expand our target for many objects like **C&C servers**, **botnet**, **social malicious botnet**, etc. and so forth, so as to enhance the relationship among nodes and assess the detection effect after these improvements.

Acknowledgment. This research is sponsored by FPT Telecom cyber security project. Many thanks to all members who have had the pleasure to work during this and related projects, each of them have important contribution.

References

1. Tran, H., Nguyen, A., Vo, P., Vu, T.: DNS graph mining for malicious domain detection. In: 2017 IEEE International Conference on Big Data (Big Data), IEEE 2017, Boston, MA, USA, 11–14 December 2017, Proceedings (2017). https://doi.org/10.1109/BigData.2017.8258515
2. Antonakakis, M., Perdisci, R., Dagon, D., Lee, W., Feamster, N.: Building a dynamic reputation system for DNS. In: USENIX Security 10 Proceedings of the 19th USENIX Conference on Security, Proceedings, p. 18. USENIX Security Symposium (2010). ISBN 888-7-6666-5555-4
3. Nazario, J., Holz, T.: As the net churns: fast-flux botnet observations. In: Proceedings of the 3rd International Conference on Malicious and Unwanted Software (MALWARE 2008), IEEE Alexandria, VA, USA, October 2008, Proceedings, pp. 24–31. Fairfax, October 2008. https://doi.org/10.1109/MALWARE.2008.4690854, ISBN 978-1-4244-3288-2
4. Yadav, S., Reddy, A.K.K., Reddy, A.L.N., Ranjan, S.: Detecting algorithmically generated malicious domain names. In: Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, Proceedings, Melbourne, Australia, 1–3 November (2010). https://doi.org/10.1145/1879141.1879148

fridaya@uef.fi

5. Antonakakis, M., Perdisci, R., Nadji, Y., Vasiloglou, N., Abu-Nimeh, S., Lee, W., Dagon, D.: From throw-away traffic to bots: detecting the rise of DGA-based malware. In: The 21th USENIX Security Symposium, Bellevue, WA, USA, 8–10 August 2012, Proceedings (2012)
6. Bilge, L., Kirda, E., Kruegel, C., Balduzzi, M.: EXPOSURE: finding malicious domains using passive DNS analysis. In: The Network and Distributed System Security Symposium, San Diego, California, USA, 6th–9th February 2011, Proceedings (2011)
7. Cova, M., Leita, C., Thonnard, O., Keromytis, A.D., Dacier, M.: An analysis of rogue AV campaigns. In: 13th International Symposium Recent Advances in Intrusion Detection, RAID 2010, Ottawa, Ontario, Canada, 15–17 September 2010, Proceedings. Lecture Notes in Computer Science, vol. 6307 (2010)
8. Crawford, H., Aycock, J.: Kwyjibo: automatic domain name generation. Softw. Pract. Exp. **38**(14), 1561–1567 (2008). https://doi.org/10.1002/spe.885
9. Manadhata, P.K., Yadav, S., Rao, P., Horne, W.: Detecting malicious domains via graph inference. In: 19th European Symposium on Research in Computer Security, Wroclaw, Poland, 7–11 September 2014. Proceedings (2014). https://doi.org/10.1145/2666652.2666659, ISNM 978-1-4503-3153-1
10. Royal, P.: Quantifying maliciousness in Alexa top-ranked domains, Proceedings, December 2012
11. Yedidia, J.S., Freeman, W.T., Weiss, Y.: Understanding belief propagation and its generalizations. In: Lake-Meyer, G., Nebel, B. (eds.) Exploring Artificial Intelligence in the New Millennium, Proceedings, San Francisco, Calif, USA, p. 239. Morgan Kaufmann Publishers, San Francisco (2003)
12. Khalil, I., Yu, T., Guan, B.: Discovering malicious domains through passive DNS data graph analysis. In: ASIA CCS 2016 Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, Proceedings, New York, NY, USA, pp. 663–674 (2016). https://doi.org/10.1145/2897845.2897877, ISBN 978-1-4503-4233-9
13. Sato, K., Ishibashi, K., Toyono, T., Hasegawa, H., Yoshino, H.: Extending black domain name list by using co-occurrence relation between DNS queries. IEICE Trans. Commun. **E95–B**(3), 794–802 (2012). https://doi.org/10.1587/transcom.E95.B.794

fridaya@uef.fi

On Non-commutative Cryptography with Cubical Multivariate Maps of Predictable Density

V. Ustimenko(✉) and M. Klisowski

University of Maria Curie Sklodowska, 20036 Lublin, Poland
{vasyl, mklisow}@hektor.umcs.lublin.pl

Abstract. Multivariate cryptosystems are divided into public rules for which tools of encryption are open for users and systems of El Gamal type for which encryption function is not given in public and for its generation opponent has to solve discrete logarithm problem in affine Cremona group. Infinite families of transformations of free module K^n over finite commutative ring K such that the degrees of their members are not growing with iteration are called stable families of transformations. Such families are needed for practical implementations of multivariate cryptosystems of El Gamal type. New explicit constructions of such families and families of stable groups and semigroups of transformations of free modules are given. New methods of creation of cryptosystems which use stable transformation groups and semigroups, and homomorphisms between them are suggested. The security of these schemes is based on a complexity of decomposition problem for element of affine Cremona semigroup into product of given generators. Proposed schemes can be used for the exchange of semigroup messages in a form of elements of a free module and for a secure delivery of multivariate maps which could be encryption tools and instruments for digital signatures.

Keywords: Multivariate cryptography · Stable transformation groups and semigroups · Decomposition problem of nonlinear multivariate map into given generators · Wild and tame families of transformations · Tame homomorphisms · Key exchange protocols · Cryptosystems · Algebraic graphs

1 On Post Quantum, Multivariate Cryptography and Non-commutative Cryptography

Let us consider basic algebraic objects of multivariate cryptography, which are important for the choice of appropriate pairs of maps f, f^{-1} in public key algorithm or in asymmetric algorithm with protected encryption rule. Post Quantum Cryptography (PQC) serves for the research of asymmetrical cryptographical algorithms which can be potentially resistant against attacks based on the use of quantum computer. Multivariate cryptography is one of the oldest directions of PQC. It uses as security tools a nonlinear polynomial transformations f of kind: $x_1 \to f_1(x_1, x_2, \ldots, x_n)$, $x_2 \to f_2(x_1, x_2, \ldots, x_n)$, $\ldots$, $x_n \to f_n(x_1, x_2, \ldots, x_n)$ acting on the affine space K^n, where fi, $i = 1$,

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 654–674, 2019.
https://doi.org/10.1007/978-3-030-22868-2_47

fridaya@uef.fi

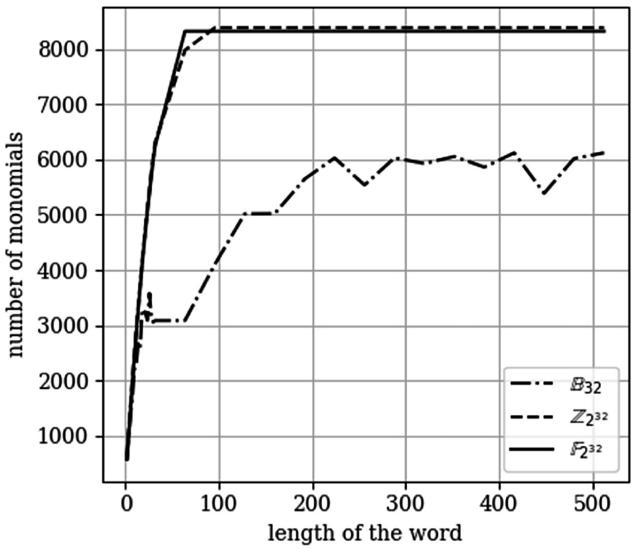


Fig. 1. Number of monomial terms of the cubic map induced by the graph ($\boldsymbol{n} = 128$) (graph $\boldsymbol{D}(\boldsymbol{n}, \boldsymbol{K}), \boldsymbol{K} = \boldsymbol{B}(32), \boldsymbol{Z}_{2^{32}}, \boldsymbol{F}_{2^{32}}$), case I.

2, …, n. from $K[x_1, x_2, \ldots, x_n]$ are multivariate polynomials given in a standard form, i.e. via a list of monomials in a chosen order (see [1]).

The most popular form is the usage of a very special map f in a *public keymode*. First examples were based on families of quadratic bijective transformation f_n (see [1–3]), such choice implies rather fast encryption process.

The paper is devoted to other aspects of Multivariate cryptography. Everybody knows that Diffie Hellman key exchange protocol can be formally considered in general case of any finite group or semigroup G. In the case of group corresponding El Gamal cryptosystem can be introduced. Notice that security of this algorithm depends not only on abstract group G but on the way of its generation in computer memory.

For instance if $G = Z_p^*$ is multiplicative group of large prime field then discrete logarithm problem (DLP) is difficult one and guarantees the security of the protocol, if the same abstract group is given as additive group of Z_{p-1} protocol is insecure because DLP will be given by linear equation.

If G is non-commutative group correspondents can use conjugations of elements involved in protocol, some algorithms of this kind were suggested in [4–7], where group G is given with the usage of generators and relations. Security of such algorithms is connected with Conjugacy Search Problem (CSP) and Power Conjugacy Search Problem (PCSP), which combine CSP and Discrete Logarithm Problem and their generalisations.

This direction belongs to **Non-commutative cryptography** which is active area of cryptology, where the cryptographic primitives and systems are based on algebraic structures like groups, semigroups and non-commutative rings (see [8–16]). This direction of security research has very rapid development [17, 18] and further

fridaya@uef.fi

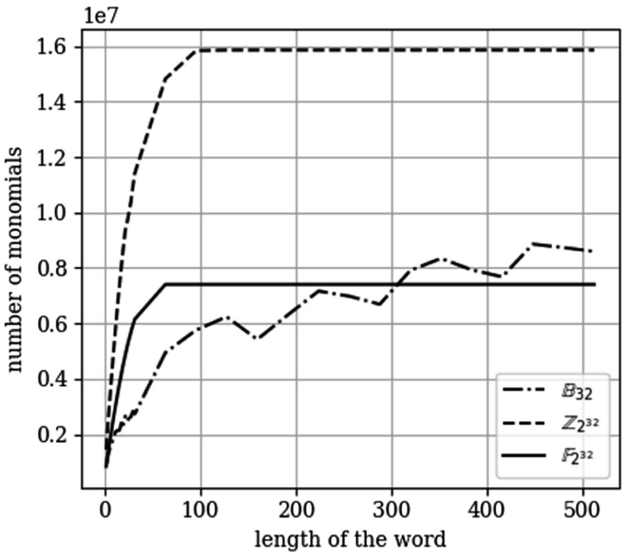


Fig. 2. Number of monomial terms of the cubic map induced by the graph ($\boldsymbol{n} = 128$) (graph $\boldsymbol{D}(\boldsymbol{n}, \boldsymbol{K}), \boldsymbol{K} = \boldsymbol{B}(32), \boldsymbol{Z}_{2^{32}}, \boldsymbol{F}_{2^{32}}$), case II

references in these publications). One of the earliest applications of a non-commutative algebraic structure for cryptographic purposes was the use of braid groups to develop cryptographic protocols. Later several other non-commutative structures like Tompson groups and Grigorchuk groups have been been identified as potential candidates for cryptographic post quantum applications. The standard way of presentations of groups and semigroups is the usage of generators and relations (Combinatorial Group Theory), Semigroup based cryptography consist of general cryptographical schemes defined in terms of wide classes of semigroups and their implementations for chosen semigroup families (so called platform semigroups).

In [19, 20] the author considers some modifications of Diffie-Hellman protocol when G is given as subgroup of affine Cremona semigroup $S(K^n)$ over finite commutative ring K of all polynomial transformations. He assumes that each element is given in its standard form of Multivariate Cryptography. To use semigroup operation one has to compute the composition of transformations. This is an attempt to combine methods of Non Commutative Cryptography and Multivariate Cryptography.

At first glance the idea of such Diffie-Hellman protocol in affine Cremona semigroup looks as unrealistic one because of composition of two maps of degree r and s taken in "general position" will be a transformation of degrees. So in majority of cases $deg(F) = d$, $d > 1$ implies the exponential growth of function $d(r) = deg(F^r)$. To guarantee the polynomiality of algorithm the assumption of stability of involved transformations can be used, i.e. degrees of elements from cyclic group generated by chosen element have to be bounded by independent constant. It is a motivation of the search for stable subgroups and semigroups of affine Cremona groups.

fridaya@uef.fi

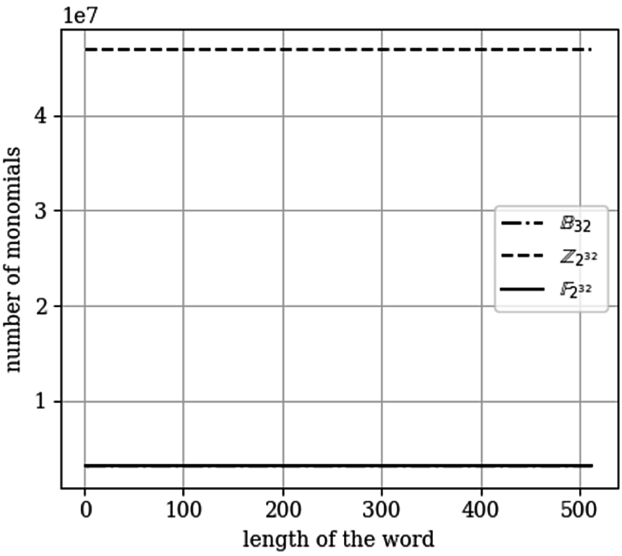


Fig. 3. Number of monomial terms of the cubic map induced by the graph $(\boldsymbol{n} = 128)$ (graph $\boldsymbol{D}(\boldsymbol{n}, \boldsymbol{K}), \boldsymbol{K} = \boldsymbol{B}(32), \boldsymbol{Z}_{2^{32}}, \boldsymbol{F}_{2^{32}}$), case III

Some general cryptograpical schemes of usage of stable sub-semi-groups and subgroups of affine Cremona group over finite commutative ring are given in [21].

In particular author presents new protocols which are not generalisations of Diffie-Hellman algorithm, their security rests on the decomposition of else these schemes use the concept of *tame homomorphism* for semigroups and the complexity of *word problem* for Cremona semigroup about the decomposition of a given polynomial transformation g into given generators.

Current paper is devoted to implementations of this schemes in terms of platforms which are stable subgroups defined via families of linguistic graphs with well-defined projective limit and the homomorphisms between the members of these increasing sequences. We use projective limits of well-known graphs $D(n, K)$ [22, 23] and $A(n, K)$ (see [24] and further references) defined *over* arbitrary finite commutative rings are implemented. Computer simulations demonstrate an interested effect of density stabilisation of generated cubical maps.

The time execution tables for algorithms of generation of maps and numbers of monomial terms are given in Appendix. The method of generation allows to construct for each bijective transformation of the *free module over K* its inverse map. Multivariate *and* nature of collision maps allows to use these algorithms for the safe exchange of multivariate transformations. Various *deformation rules* can be used for this purpose (see [21]).

fridaya@uef.fi

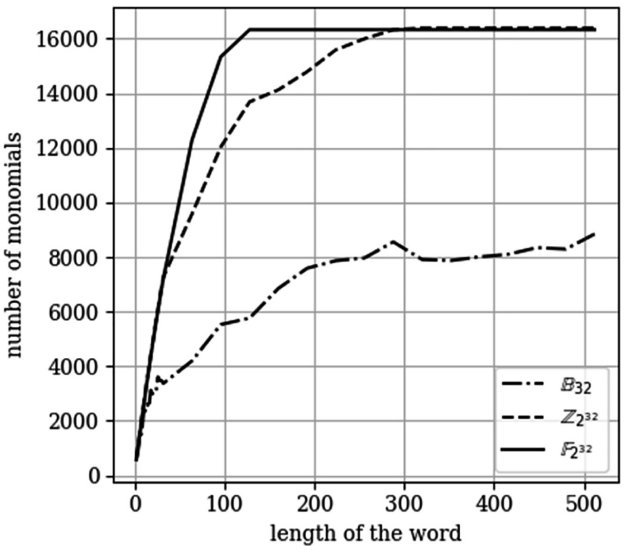


Fig. 4. Number of monomial terms of the cubic map induced by the graph ($\boldsymbol{n} = 128$) (graph $\boldsymbol{A}(\boldsymbol{n}, \boldsymbol{K})$, $\boldsymbol{K} = \boldsymbol{B}(32)$, $\boldsymbol{Z}_{2^{32}}$, $\boldsymbol{F}_{2^{32}}$), case I

2 Tame Families and Concept of Stability

Let us consider basic algebraic objects of multivariate cryptography, which are important for the choice of appropriate pairs of maps f, f^{-1} in both cases of public key approach or idea of asymmetric algorithms with protected encryption rules. Let us consider the totality $SF_n(K)$ of all rules f of kind: $x_1 \to f_1(x_1, x_2, \ldots, x_n)$, $x_2 \to f_2(x_1, x_2, \ldots, x_n), \ldots, x_n \to f_n(x_1, x_2, \ldots, x_n)$ acting on the affine space K^n, where f_i, $i = 1, 2, \ldots, n$. for given parameter n and chosen commutative ring K with natural operation of composition. We refer to this semigroup as semigroup of formal transformation $SF_n(K)$ of free module K^n. In fact it is a totality of all endomorphisms of ring $K[x_1, x_2, \ldots, x_k]$ wth the operation of their superposition. Each rule f from $SF_n(K)$ induces transformation $t(f)$ which sends tuple $(p_1, p_2, \ldots, p_n)$ into $(f_1(p_1, p_2, \ldots, p_n), f_2(p_1, p_2, \ldots, p_n), \ldots, f_n(p_1, p_2, \ldots, p_n))$. Affine Cremona semigroup $C(K^n)$ is a totality of all transformations of kind $t(f)$. The canonical homomorphism $t \to t(f)$ maps infinite semigroup $SF_n(K)$ onto finite semigroup $S(K^n)$ in the case of finite commutative ring K.

We refer to pair (f, f') of elements $SF_n(K)$ such that ff' and $f'f$ are two copies of identical rule $x_i \to x_i$, $i = 1, 2, \ldots, n$. as pair of invertible elements. If (f, f') **is** such a pair, then product $t(f)\, t(f')$ is an identity map. Let us consider the subgroup $CF_n(K)$ of all invertible elements of $SF_n(K)$ (group of formal maps). It means f is an element of $CF_n(K)$ if and only if there is f' such that ff' and $f'\, f$ are identity maps. It is clear that the image of a restriction of t on $CF_n(K)$ is affine Cremona group $C_n(K)$ of all transformations of K^n onto K^n for which there exists a polynomial inverse.

fridaya@uef.fi

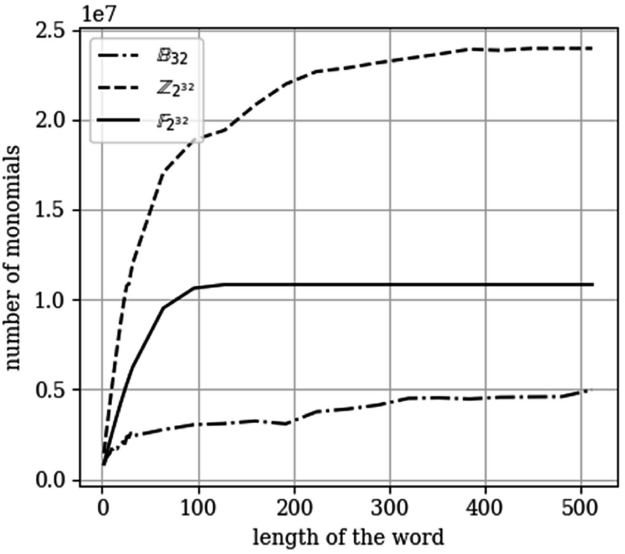


Fig. 5. Number of monomial terms of the cubic map induced by the graph ($\boldsymbol{n} = 128$) (graph $\boldsymbol{A}(\boldsymbol{n}, \boldsymbol{K})$, $\boldsymbol{K} = \boldsymbol{B}(32)$, $\boldsymbol{Z}_{2^{32}}$, $\boldsymbol{F}_{2^{32}}$), case II

We say that a family of subsemigroups S_n of $SF_n(K)$ (or $S(K^n)$) is stable of degree d if maximal degree of elements from S_n is an independent constant d, $d > 2$. If K is a finite commutative ring then stable semigroup has to be a finite set.

Let f_n from $SF_n(K)$ be a family of nonlinear maps of degree bounded by constant d. We say that f_n form tame family if there is a family g_n from $SF_n(K)$ of degree bounded by constant d' such that $f_n g_n = g_n f_n$ are identity maps. Let T_1 and T_2 be two elements from the group $AGLn(K)$ of all affine bijective transformations, i.e. elements of affine Cremona group of degree 1. Then we refer to $f'_n = T_1 f_n T_2$ as linear deformation of f_n. Obviously f'_n is also tame family of transformations, degrees of maps from this family is also bounded by d. The degrees of inverses of f'_n are bounded by d'. Let $G_n < SF_n(K)$ be a stable family of subgroups of degree d, $d \geq 2$, then nonlinear representatives f_n of G_n form tame family of maps.

3 On the Concept of Tame Homomorphism and Related Algorithms

Let $G = G_n$ be a family of stable subsemigroups of $SF_n(K)$ (or $S(K^n)$) and $L = L_m$, where m depends on n be a family of stable subsemigroups of $SF_m(R)$ (or $S(R^m)$, where K and R are commutative rings. There are tame homomorphisms $\varphi = \varphi_{\mathbf{n}}$ from G into L, i.e. value of φ in each point g from G_n is computable in polynomial time from n. Let us assume, that there are semigroups $B = B_n < G_n$ given by their generators $b_1, b_2, \ldots, b_r$. Let us assume that Alice has families of tame transformations π_1 of K^n

fridaya@uef.fi

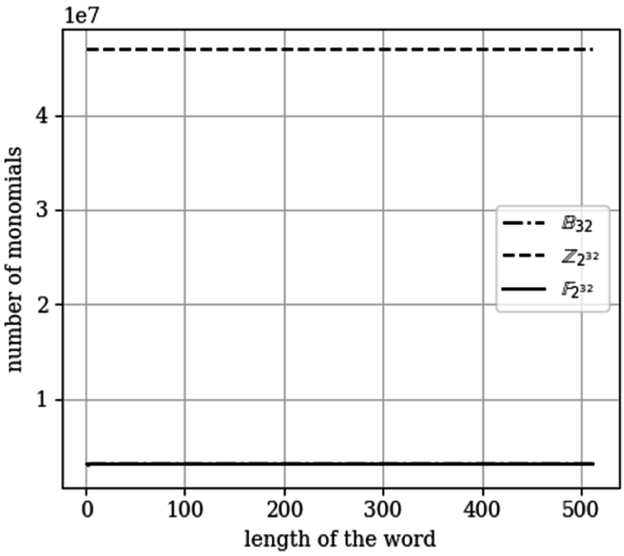


Fig. 6. Number of monomial terms of the cubic map induced by the graph ($\boldsymbol{n} = 128$) (graph $\boldsymbol{A}(\boldsymbol{n}, \boldsymbol{K})$, $\boldsymbol{K} = \boldsymbol{B}(32)$, $\boldsymbol{Z}_{2^{32}}$, $\boldsymbol{F}_{2^{32}}$), case III

and of π_2 of R^m. We assume that these data is known to Alice. She forms ($a_i = \pi_1 b_i \pi_1^{-1}$, $a_i' = \pi_2\varphi(b_i)\pi_2^{-1}$), $i = 1, 2, \ldots, r$ and sends them to Bob. The elements of these pairs are given in their standard forms for representatives of $SF_n(K)$ or $SF_m(R)$.

3.1 Key Exchange Protocol

The list of pairs known for Bob defines homomorphism Δ between sub-semi-groups $A = <a_1, a_2, \ldots, a_r>$ and $A' = <a_1', a_2', \ldots, a_r'>$ given by its values on generators $\Delta(a_i) = \Delta(a_i')$ for $i = 1, 2, \ldots, r$. Bob forms a via his choice of word $(a\, i_1)^{k1}(a\, i_2)^{k2} \ldots (a\, i_t)^{kt}$ in the alphabet of generators of A such that $a\, i_s \neq a\, i_{s+1}$ for $s = 1, 2, \ldots, t-1$. He sends a to Alice and keeps $\Delta(a) = a' = (a'\, i_1)^{k1}(a'\, i_2)^{k2} \ldots (a'\, i_t)^{kt}$ as collision element. Alice knows tame homomorphism φ and easily computes a' as $\pi_2\varphi\left(\pi_1^{-1} a \pi_1\right)\pi_2^{-1}$.

Complexity Remark.
Adversary has to solve *word problem* for sub-semi-group, i.e. find decomposition of a from A into generators a_1, $i = 1, 2, \ldots, t$. General algorithm to solve this problem in polynomial time in variable n is unknown as well as a procedure to get its solution in terms of quantum computations.

Remark 1. The condition of stability for semigroups G and L and the usage of tame transformations π_1 and π_2 allow us to estimate degrees of a and collision map. If maximal degrees of $\pi_1(n)$ and $\pi_1^{-1}(n)$ are l_1 and l_1' **a**nd degrees of $\pi_1(n)$ and $\pi_1^{-1}(n)$ are

fridaya@uef.fi

bounded by l_2 and l_2 and stable groups G and L are of degrees d and d', then degrees of a and $\Delta(a)$ are bounded by $l_1\, l'_1 d$ and $l_2 l'_2 d'$

Remark 2. One can use other natural conditions on π_1 and π_2. Let us assume that $G < G_1$ and $L < G_2$, where G_i are stable families of sub-semi-groups of degree t_1, $t_1 \leq d$ and, $t_2 t_2 \leq d'$ respectively. Let us consider normalisers N_1, and N_2 of G_1 and G_2 in affine Cremona semigroups $S(K^n)$ and $S(R^m)$. It means that N_1, and N_2, are totalities of transformations π from $C(K^n)$ and $C(R^m)$ such that $\pi_i G_i \pi_i^{-1}$, $i = 1, 2$ coincide with G_1 and G_2. We can take π_1 of kind $T_1 n_1$ and π_2 of kind $T_2 n_2$, where take π_1 of kind n_i are elements of N_i, $i = 1, 2$, transformations T_1 and T_2 are elements of groups $AGL_n(K)$ and $AGLm(R)$. Then degrees of a and $\Delta(a)$ are restricted by t_1 and t_2.

Notice that in the case $G = G_1$, $L = G_2$ degrees of a and $\Delta(a)$ are bounded by d and d'. We refer to presented above algorithm as *tahoma word* protocol. Term *tahoma* (name of shrift for word processing) stands for combination *tame homomorphism*.

The protocol exploits the complexity of the *word problem* for a semigroup of polynomial transformation of free module. In the case considered in *remark 2* we use term *stable tahoma word* protocol.

3.2 Inverse Tahoma Word Protocol

Let us modify above protocol 3.1 in the case of invertible elements $\varphi(a_i)$ with an assumption that their inverses are known to Alice. Instead of pairs (a_i, a'_i) Alice forms $(a_i, , a_i^*)$, where $a_i^* = (a'_i)^{-1}$, $i = 1, 2, \ldots, r$. Assume Bob gets from Alice the list of such pairs. Notice that A' is a group $< a_i^* | i = 1, 2, \ldots, r >$. So Bob is able to compute anti-isomorphism δ sending z from A into $\varphi(z)^{-1}$ because he knows values of δ on generators. Like in previous protocol Bob forms a via his choice of word $(a\, i_1)^{k1}(a\, i_2)^{k2} \cdots (a\, i_t)^{kt}$ in the alphabet of generators of A such that $a\, i_s \neq a\, i_{s+1}$ for $s = 1, 2, \ldots, t-1$ and sends a to Alice. Now he keeps $\delta(a) = (a * i_t)^{kt} (a * i_{t-1})^{kt-1} \ldots (a * i_2)^{k2} (a * i_1)^{k1}$ as an element of collision pair. Alice computes he part of collision pair $e = \delta(a)^{-1}$ as $\pi_2 \varphi(\pi_1^{-1} a \pi_1) \pi_2^{-1}$.

3.3 Inverse Tahoma Cryptosystem

Both above written protocols exploit the complexity of finding decomposition of a into product of given generating transformations. In case of 3.2 Alice and Bob can secure communicate because they are able to use mutually inverse transformations e and e^{-1} as encryption tools. Alice writes her message p and sends $e(p)$ to Bob, who decrypts via usage of $\delta(a)$. Bob can encrypt with $\delta(a)$ and Alice decrypts with e.

Remark.
In the case when transformation $e = e_m$ of free module R^m form stable family of degree d' adversary has to intercept $O(m^{d'})$ messages and conduct costly linearisation attack to restore e and $\delta(a)$ So correspondents can safely exchange $O(m^{d\prime-1})$ messages. Notice that any moment Alice and Bob can start a new session of inverse tahoma word protocol.

fridaya@uef.fi

4 On the Inverse Version of the Group Enveloped Symbolic Diffie-Hellman Key Exchange Protocol

Let $G = G_n$ be a family of stable subsemigroups of $SF_n(K)\,(or\,S(K^n))$ and $L = L_m$ where m depends on n, be a family of stable subsemigroups of $SF_m(R)\,(orS(R^m))$, where K and R are commutative rings. There are tame homomorphism $\varphi = \varphi_n$ from G into L, i.e. value of φ in each point g from G_n is computable in polynomial time from n Let us assume, that there are subgroups $A = A_m < L_m$ and $B = B_n < G_n$ such that $\varphi(b)a = a\varphi(b)$ for all elements a from A, b from B. Assume that two families of tame transformations $\pi = \pi(n)$ and $\mu = \mu(m)$ are chosen.

We assume that these data is known to Alice. She forms pairs $(c_i = \pi b_i \pi^{-1}$, $c_i^{-1} = \pi b_i^{-1}\pi^{-1})$ and $(d_i = \mu\varphi(b_i)\mu^{-1}$,$d_i^{-1} = \mu\varphi(b_i^{-1})\mu^{-1})$, $i = 1, 2, \ldots, r$ where elements b_i are from B_n, they, there inverses and images are given in their standard from $SF_n(K)$ or $SF_m(R)$ and sends them to Bob. Let Δ be a homomorphism from $<c_1, c_2, \ldots, c_r>$ to $<d_1, d_2, \ldots, d_r>$ sending c_i to d_i

We present briefly the protocol of symbolic computations introduced in [4] and define its inverse version. We refer to this protocols as *group enveloped Diffie-Hellman scheme* and *inverse group enveloped Diffi-Hellman scheme.*

4.1 Protocol

Alice takes positive integer k_A and a, a^{-1} from A_m and g' from the semigroup G. She computes $g_A = \mu a\varphi(g')a^{-1}\mu^{-1}$ and sends to Bob $g = \pi g'\pi^{-1}$ together with g_A. Bob chooses positive integer k_B and element c from R^m in $<c_1, c_2, \ldots, c_r>$ (via the choice of word in alphabet $c_1, c_2, \ldots, c_r$. He computes $g_B = cg^t c^{-1}$ with $t = k_B$ in standard form for elements of $SF_n(K)$ and sends it to Alice.Bob computes a map $\Delta(c)g_A^t\Delta(c^{-1})$ because he knows the decomposition of c and c^{-1} into their generators and keeps it as the collision map. Alice computes the collision map as $\mu a\varphi(\pi^{-1}g_B^s\pi)a^{-1}\mu^{-1}$ with $s = k_A$.

Remark.
Adversary has to consider group $C' = <c_i | i = 1, 2, \ldots, r>$ and solve *group enveloped discrete logarithm problem*, i.e. solve $yg^x y^{-1} = g_B$ where x is unknown integer parameter and y from C' (possible via solving of decomposition problem of g_B into semigroup generators $c_1, c_2, \ldots, c_r, g$ (decomposition of elements into Cremona semigroup generators, *word problem in affine Cremona semigroup*).

4.2 Inverse Protocol

Let us assume that Alice can generate g' such that $\varphi(g')$ is invertible and the inverse $\varphi(g')^{-1}$ is computable for her.

As in previous algorithm Alice takes positive integer k_A, and element a, a^{-1} from A_m and g' from the semigroup G. Now she computes $z = \varphi(g')^{-1}$ and $g_A = \mu aza^{-1}\mu^{-1}$ and sends to Bob $g = \pi g'\pi^{-1}$ together with g_A.

fridaya@uef.fi

As in previous algorithm Bob chooses positive integer k_B, and element c from $<c_1, c_2, \ldots, c_r>$ via the choice of word in the alphabet of generators. He computes $g_B = cg^t c^{-1}$ with $t = k_B$ in standard form of $SF_n(K)$ and sends it to Alice. Bob computes a map $e = \Delta(c)(g_A)^t \Delta(c)^{-1}$, $t = k_B$ because he knows the decomposition of c and c^{-1} into their generators and keeps e as his outcome of a collision. Alice computes the map e^{-1} as $\mu a \varphi(\pi^{-1}(g_B)^s \pi) a^{-1} \mu^{-1}$, $s = k_A$.

4.3 Inverse Group Enveloped Cryptosystem

Alice and Bob can communicate because of they have mutually inverse transformations e^{-1} and e. 1) Alice writes her message p and sends $e^{-1}(p)$ to Bob, who decrypts via usage of e. 2) Bob can encrypt with e and Alice decrypts with e^{-1}. The algorithm (4.3) was introduced in [20] (see [21] and further references) as *desynchronised symbolic El Gamal* algorithm.

5 Stable Groups of Cubical Maps and Protocol Based on Cremona Word Problem

The following family of stable groups is considered in [24]. Let K be a commutative ring. We define $A(n, K)$ as bipartite graph with the point set $P = K^n$ and line set $L = K^n$ (two copies of a Cartesian power of K are used). We will use brackets and parenthesis to distinguish tuples from P and L. So $(p) = (p_1, p_2, \ldots, p_n) \in P_n$ and $[l] = [l_1, l_2, \ldots, l_n] \in L_n$. The incidence relation $I = A(n, K)$ (or corresponding bipartite graph I) is given by condition $pI\,l$ if and only if the equations of the following kind hold.

$p_2 - l_2 = l_1 p_1$, $p_3 - l_3 = p_1 l_2$, $p_4 - l_4 = l_1 p_3$, $p_5 - l_3 = p_1 l_4$, …, $p_n - l_n = p_1 l_{n-1}$ for odd n and $p_n - l_n = l_1 p_{n-1}$ for even n.

Let us consider the case of finite commutative ring K, $|K| = m$. As it instantly follows from the definition the order of our bipartite graph $A(n, K)$ is $2m^n$. The graph is m-regular. In fact the neighbour of given point p is given by above equations, where parameters $p_1, p_2, \ldots, p_n$ are fixed elements of the ring and symbols $l_1, l_2, \ldots, l_n$ are variables. It is easy to see that the value for l_1 could be freely chosen. This choice uniformly establishes values for $l_2, l_3, \ldots, l_n$. So each point has precisely m neighbours. In a similar way we observe the neighbourhood of the line, which also contains m neighbours. We introduce the colour $\rho(p)$ of the point p and the colour $\rho(l)$ of line l as parameter p_1 and l_1 respectively.

Graphs $A(n, K)$ with colouring ρ belong to class of Γ *linguistic graphs* considered in [25], Paper [26] observes cryptographical applications of linguistic graphs, which were considered in [27] for slightly more general class of graphs with special colourings. Linguistic graph $\Gamma = \Gamma(K)$ is defined over commutative ring K is a bipartite graph with partition sets K^n and K^m have color set $L = K^s$ and $L = K^r$ respectively. Projection ρ of point $x = (x_1, x_2, \ldots, x_n)$, or line $y = [y_1, y_2, \ldots, y_m]$, on the tuple of their first s and r coordinates respectively defines colours of vertices.

fridaya@uef.fi

Each vertex has a unique neighbor of selected colour. So $n + r = m + s$. The incidence of linguistic graphs is given by a system of polynomial equation over ring K.

In the case of linguistic graph $\Gamma(K)$ with $s = r = 1$ the path consisting of its vertices $v_0, v_1, v_2, \ldots, v_k$ is uniquely defined by initial vertex v_0, and colours $\rho(v_i,), i = 1, 2, \ldots, k$ of other vertices from the path. We can consider graph $\Gamma(K[x_1, x_2, \ldots, x_n])$ defined by the same with Γ equations.

So the following symbolic computation can be defined. Take the *symbolic point* $x = (x_1, x_2, \ldots, x_n)$, where x_i are generic variables of $K[x_1, x_2, \ldots, x_n]$ and *symbolic key* which is a string of polynomials $f_1\prime, f_2\prime, \ldots, f_k$, from $K[x_1]$. Form the path of vertices $v_0, = x, v_1$ such that $v_1 I v_o$ and $\rho(v_1) = f_1(x_1)$, v_2 such that $v_2 I v_1$ and $\rho(v_2) = f_2(x_1), \ldots, v_k$ such that $v_k I v_{k-1}$ and $\rho(v_k) = f_k(x_1)$.

We use term *symbolic point to point computation* in the case of even k and talk about *symbolic point to line computation* in the case of odd k. We notice that the computation C of each coordinate of v_i via variables $x_1, x_2, \ldots, x_n$ and polynomials $f_1\prime, f_2\prime, \ldots, f_k$ needs only arithmetical operations of addition and multiplication. As it follows from the definition of linguistic graph final vertex v_k (point or line) has coordinates $(h_1(x_1), h_2(x_1, x_2), h_3(x_1, x_2, x_3), \ldots, h_n(x_1, x_2, \ldots, x_n))$, where $h_1(x_1) = f_k(x_1)$. Let us consider the map $H = \eta(C) : x_i \to h_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$ which corresponds to computation C. Assume that the equation $b = f_k(x_1)$ has exactly one solution. Then the map $H : x_i \to h_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$ is a bijective transformation. In the case of finite parameter k and finite densities of $f_i(x_1)$, $i = 1, 2, \ldots, n$ the map H also has finite density. If all parameters deg $(f_i(x_1))$ are finite then the map H has a linear degree in variable n. The idea of symbolic computation was proposed in [28].

Let us consider the totality $\sum = \sum(n, K)$ of point to point computations C with the symbolic key of kind $f_i(x_1) = x_1 + a_i$, $i = 1, 2, \ldots, t$, where parameter t is even. In case of a linguistic graphs with $r = s = 1$ we identify a computation C with the corresponding string $(a_1, a_2, \ldots, a_t)$.We assume that empty string is also an element of $\sum$. Natural product of two strings given by tuples $C_1 = (a_1, a_2, \ldots, a_t)$ and $C_2 = (b_1, b_2, \ldots, b_m)$ is a string $C = C1 \circ C2 = (a_1, a_2, \ldots, a_t, b_1 + a_t, b_2 + a_t, \ldots, b_m + a_t)$. This product transforms $\sum$ to a semigroup. The map η' sending C to $\eta(C)$ is a homomorphism of $\sum$ into affine Cremona group $C(K^n)$. In the case of linguistic graphs $A(n, K)$ we can prove that the totality $G(n, K) = \eta'(\sum(n, K))$ is a stable subgroup of degree 3 (see [25]).

We assume that $a_0 = 0$ and say that transformation $\eta(C)$ is irreducible if $a_i \neq a_{i+2}$, $i = 1, 2, \ldots, t-2$. If $a_1 \neq a_{t-1}$, and $a_2 \neq a_t$ we say that irreducible computation C and corresponding transformation $\eta(C)$ are standard elements. We have a natural homomorphism $G(n + 1, K)$ onto $G(n, K)$ induced by the homomorphism Δ from $A(n + 1, K)$ onto $A(n, K)$ sending point $(x_1, x_2, \ldots, x_n, x_{n+1})$ to $(x_1, x_2, \ldots, x_n)$ and line $[x_1, x_2, \ldots, x_n, x_{n+1}]$ to $[x_1, x_2, \ldots, x_n]$. It means that there is well defined projective limit $A(K)$ of graphs $A(n, K)$ and groups $G(K)$ of groups $G(n, K)$ when n is growing to infinity. As it stated in [25] case of $K = F_q$, $q > 2$ infinite graph $A(F_q)$ is a tree.

It means that group $G(F_q)$ is a group of walks of even length on q regular tree starting in zero point with natural addition of them. A standard computation C defines

fridaya@uef.fi

transformation $\eta(C)$ in each group $G(n, K)$, $n \geq 2$ and $G(K)$. An irreducible transformation $\eta(C)$ from $G(K)$ has an infinite order.

Some examples of tame homomorphism are considered in [21]. We are going to use the family of maps introduced below.

Let Δn, $m\prime$, $n > m$ be a canonical homomorphism of $A(n, K)$ onto $A(m, K)$ corresponding to procedure of deleting of coordinates with indexes $m + 1, m + 2, \ldots, n$. This map defines the canonical homomorphism $\mu(n, m)$ of group $G(n, K)$ onto $G(m, K)$.

Let R and K be finite extensions of a finite ring Q. Let us consider the diagram $S(R^m) > G(m, R) > G(m, Q) \leftarrow G(n, Q) < G(n, K) < S(K^n)$ with extreme nodes $S(R^m)$ and $S(K^n)$, where $n > m$ and arrow corresponds to canonical homomorphism $\mu = \mu(n, m), n > m$. Alice is going to use *stable tahoma word protocol* with $G = G(n, Q)$. She can use subgroups $H_1 = G_1$ and $H_2 = L_1$ of invertible elements instead of their whole normalisers $N_1 = N(G_1)$ of subgroup G_1 in affine Cremona group $S(K^n)$, and $N_2 = N(L_1)$ of subgroup in the Cremona group $S(R^m)$. She also works with affine subgroups $AGLn(K)$ and $AGLm(R)$. Alice forms $\pi_1 = T_1 n_1$, where $T_1 \epsilon AGL1n(K), n_1 \epsilon H_1$, and $\pi_2 = T_2 n_2$ where $T_2 \epsilon AGL1m(R), n_2 \epsilon H_2$. She takes subgroup where $G' = \pi_1 G \pi_1^{-1}$, $L' = \pi_2 L \pi_2^{-1}$ and homomorphism $\mu' : x \rightarrow \pi_2(\mu(\pi_1 x \pi_1^{-1}))\pi_2^{-1}$ between these semigroups.

Alice will work with G', L' and μ' instead of G, L, μ.The map μ' is tame for her because she knows the decomposition of μ' into conjugation with π_2, μ and the conjugation with π_1.

5.1 Example of Stable Tahoma Word Protocol

Alice takes subgroup B of $G\prime$ given by the generators $b1, b2, \ldots, br$ and sends these generators to Bob together with string $\mu'(bi)$, $i = 1, 2, \ldots, r$. So Bob knows just the restriction of μ' on B given by its values on generators. He does not know triple G', L' and μ' with its decomposition into 3 maps. So Bob selects $a1, a2, \ldots, as$ where ai are elements of the alphabet $\{b1, b2, \ldots, br\}$ He computes the composition b of selected powers of transformation ai from $S(K^n)$ and sends standard form of b to Alice. He keeps a composition b' of elements $\mu'(ai)$ according to selected order. Alice gets b' as $\mu'(b)$.

5.2 General Complexity Estimates for the Cryptosystem in the Case $K = R = Q$

Let us assume that Alice is going to use the homomorphism between $A(n, K)$ and $A(m, K)$ for $m < n$ and $m = O(n)$. We will count number arithmetical operations of commutative ring K which she need to generate an element of $g = G(n, K)$ which corresponds to symbolic computation with the key of length $O(1)$.

Counting steps of recurrent process of creation g via the symbolic automaton gives us $O(n)$ operations. Alice chooses affine transformations T and T^{-1}. Computation T^{-1} costs $O(n^2)$ elementary operations. It means that Alice can create $T^{-}gT^{-1}$ for $O(n^5)$ operations. Alice forms elements $b1, b2, \ldots, br$ from $G(n, K)$ together with their

fridaya@uef.fi

inverses and homomorphic images $\mu'(bi), i = 1, 2, \ldots, r$ from $G(m, K)$ in time $O(n)$. She takes T^- and T^{-1} from $AGLn(K)$ and forms $ai = T^- biT^{-1}$ and $a'i = T(bi^{-1})T^{-1}$ in time $O(n^5)$.

Bob receives the list of pairs ai, $a'i, i = 1, 2, \ldots, r$. He computes chosen word of kind $a = a_{i1}^{k1} a_{i2}^{k2} \ldots, a_{it}^{kt}$ for chosen finite parameter t and integers ki, $i = 1, 2, \ldots, t$ in time $O(n^{13})$ operations and sends it to Alice. Bob writes his message $p = (p_1, p_2, \ldots, p_m)$. To form ciphertext he applies to p transformation a'_{it} with multiplicity k_t, a'_{it-1} with multiplicity k_{t-1}, …, a'_{i1} with multiplicity k_1 and forms ciphertext c. It takes him $O(n^3)$ elementary operations. Alice computes cubical $b = aT$ with $O(n^4)$ operations. After she gets $d = T^{-1}b$ in time $O(n^4)$. Alice easily gets $\mu(d)$ and computes $e = T_1 d$ and $f = eT_1^{-1}$. She computes p as $f(c)$. The last step cost her $O(n^3)$ elementary ring operations.

5.3 Special Implementations

Let us consider an implementation of above cryptosystem in the case of a choice of maps T and T^1 as linear maps of finite density. Natural examples of such maps are monomial transformations or elements of direct sum of groups of kind $GLd(K)$, where d is a finite constant. Alice chooses affine transformations T and T^1. Computation gT^{-1} costs $O(n)$ elementary operations. It means that Alice can create $Tg\,T^{-1}$ for $O(n)$ operations. Alice forms elements $b1, b2, \ldots, br$ from $G(n, K)$ together with their inverses and homomorphic images $\mu(bi)$, $i = 1, 2, \ldots, r$ from $G(m, K)$ in time $O(n)$. She T and T^{-1} from $AGL_n(K)$, T_1 and T_1^{-1} and forms $ai = Tbi, T^{-1}$ and $a'i = T_1\mu(bi^{-1})\,T_1^{-1}$ in time $O(n)$. Bob receives the list of pairs ai, $a'i$, $i = 1, 2, \ldots, r$ of cubical elements of density $O(n)$. He computes chosen word of kind $a = a_{i1}^{k1} a_{i2}^{k2} \ldots, a_{it}^{kt}$ ^{t} for chosen finite parameter t and integers ki, $i = 1, 2, \ldots, t$ in time $O(n^5)$ operations and sends it to Alice. Bob writes his message $(p) = (p_1, p_2, \ldots, p_m)$.To form ciphertext he applies to p transformation a'_{it} with multiplicity k_t, a'_{it-1} with multiplicity k_{t-1}, …, a'_{i1} with multiplicity k_1 and forms ciphertext c. It takes him $O(n)$ elementary operations, Alice computes cubical $b = aT$ with $O(n^2)$ operations. After she gets $T_1 d = T^{-1}b$ in time $O(n^2)$. Alice easily gets $\mu(d)$ and computes $e = T_1 d$ and $f = eT_1^{-1}$. She computes p as $f(c)$. The last step cost her $O(n^2)$ elementary ring operations.

6 Graphs *a(N,Q)* and *D(N,Q),* Digital Condensed Matters Physics Effect

We can substitute graph $A(n, K)$ for other linguistic graph L of type $(1, 1, n-1)$ defined over the commutative ring K and rewrite the content of Sect. 5. We use graphs $A(n, K)$ and well known linguistic graph $D(n, K)$ of this type to implement all algorithm of previous section. Graphs $D(n, K)$ are bipartite with set of vertices $V = PUL, |P \cap L = 0|$. A subset of the vertices P is called the set of *points* and another subset L is called the set of *lines.* Let P and L be two copies of Cartesian power K^n,

fridaya@uef.fi

where $n \geq 2$ is an integer. Two types of brackets are used in order to distinguish points from lines. It has a set of vertices (collection of points and lines), which are n-dimensional vectors over $K : (p) = (p_1, p_2, p_3, p_4, \ldots, p_i, p_{i+1}, p_{i+2}, p_{i+3}, \ldots, p_n)$, $[l] = [l_1, l_2, l_3, l_4, \ldots, l_i, l_{i+1}, l_{i+2}, l_{i+3}, \ldots, l_n]$. The point $(\boldsymbol{p})$ is incident with the line $[\boldsymbol{l}]$, if the following relations between their coordinates hold: $l_2 - p_2 = l_1 p_1$, $l_3 - p_3 = l_2 p_1$, $l_4 - p_4 = l_1 p_2, l_i - p_i = l_1 p_{i-2}$, $l_{i+1} - p_{i+1} = l_{i-1} p_1$, : $l_{i+2} - p_{i+2} = l_i p_1 : l_{i+3} - p_{i+3} = l_1 p_{i+1}$, where $i \geq 5$. Connected component of edge-transitive graph $D(n, q)$ is denoted by $CD(n, q)$ [22]. Notice that all connected components of the natural projective limit $D(q)$ of graphs $D(n, q), n \to \infty$ infinite graph $D(q)$ are q -regular trees.

Let us denote as $G'(n, K)$ the group of elements of kind $g = \eta'(C)$ *of irreducible computation* computation $C = (a_1, a_2, \ldots, a_t)$ in the case of graphs $D(n, K)$.

We present time of generation (in ms) of element g from $G(n, K)$ and $G'(n, K)$ in the cases of graphs $A(n, K)$ and $D(n, K)$ and number $M(g)$ of monomial terms for g.

We refer to parameter t as *length of word.* We can see the ''condensed matters physics'' digital effect. If t is''sufficiently large'', then $M(g)$ is independent from t constant c. It means that the density of cubical collision map in all algorithm is simply c.

We have written a program for generating of elements and for encrypting text using the generated public key. The program is written in C ++ and compiled with the gcc compiler. We used an average PC with processor Pentium 3.00 GHz, 2 GB memory RAM and system Windows 7. We have implemented three cases:

1. T and T_1 are identities,
2. T and T_1 are maps of kind $x_1 \bullet \leftarrow x_1 + a_2 x_2 + a_3 x_3 + \bullet \leftarrow + a_t x_t, x_2 \bullet \leftarrow x_2, x_3 \bullet x_3, x_t \bullet \leftarrow x_t, \quad a_i \neq 0, \quad i = 1, 2, \ldots,$ t(linear time of computing for T and T_1), where t = n and t = m, respectively.
3. $T = Ax + b, T_1 = A_1 x + b_1$; matrices A, A_1 and Vectors b, b_1 have mostly nonzero elements.

The Tables 1, 2, 3, 4, 5, and 6 present the number of monomials depending on the number of variables (n) and the password length in all three cases and both families of graphs $D(n,K)$ and $A(n, K)$, where K is a finite field of characteristic 2. The Tables 7, 8, 9, 10, 11, and 12 present the time (in milliseconds) of the generation of public key monomials depending on the number of variables n and the length of the word in all three cases and both families of graphs $D(n, K)$ and $A(n, K)$. In [29–31] this similar program *for* program the case when K is Boolean ring were used for investigation of classical Diffie - Hellman protocol for cyclic group $< g >$ and corresponding El

Table 1. Number of monomial terms of the cubic map induced by the graph $\boldsymbol{D}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case I

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	145	145	145	145	145
32	544	545	545	545	545
64	1584	2112	2113	2113	2113
128	3664	6240	8320	8321	8321

fridaya@uef.fi

Table 2. Number of monomial terms of the cubic map induced by the graph $\boldsymbol{D}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case II

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	3649	3649	3649	3649	3649
32	41355	41356	41356	41356	41356
64	440147	529052	529053	529053	529053
128	3823600	6149213	7405944	7405945	7405945

Table 3. Number of monomial terms of the cubic map induced by the graph $\boldsymbol{D}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case III

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	6544	6544	6544	6544	6544
32	50720	50720	50720	50720	50720
64	399424	399424	399424	399424	399424
128	3170432	3170432	3170432	3170432	3170432

Table 4. Number of monomial terms of the cubic map induced by the graph $\boldsymbol{A}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case I

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	250	250	250	250	250
32	770	1010	1010	1010	1010
64	1810	3074	4066	4066	4066
128	3890	7202	12290	16322	16322

Table 5. Number of monomial terms of the cubic map induced by the graph $\boldsymbol{A}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case II

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	5623	5623	5623	5623	5623
32	53581	62252	62252	62252	62252
64	454375	680750	781087	781087	781087
128	3607741	6237144	9519921	10826616	10826616

Table 6. Number of monomial terms of the cubic map induced by the graph $\boldsymbol{A}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case III

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	6544	6544	6544	6544	6544
32	50720	50720	50720	50720	50720
64	399424	399424	399424	399424	399424
128	3170432	3170432	3170432	3170432	3170432

fridaya@uef.fi

Table 7. Generation time for the map (ms) $\boldsymbol{D}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case I

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	12	24	32	52	100
32	64	140	292	592	1192
64	1044	2261	4833	9985	20270
128	15821	33846	74340	160213	331895

Table 8. Generation time for the map (ms) $\boldsymbol{D}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case II

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	28	48	100	212	420
32	284	648	1372	2816	5712
64	3229	8397	19454	41568	85783
128	55075	139366	357361	824166	1758059

Table 9. Generation time for the map (ms) $\boldsymbol{D}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case III

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	76	140	268	524	1036
32	1224	2328	4541	8968	17828
64	21889	40417	77480	151592	299844
128	453798	812140	1526713	2946022	5792889

Table 10. Generation time for the map (ms) $\boldsymbol{A}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case I

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	4	12	24	48	96
32	56	132	288	600	1232
64	996	2100	4644	10068	20933
128	15645	33489	74244	167454	364707

Gamal cryptosystem. Currently we expand this computer package on the case of commutative rings Z_m, where m is power of *2*.

Illustrative Example.
Let Alice selects the graph $A(n, K), K = F_{2^{32}}, n = 64$ and its canonical homomorphism onto graph $A(32, K)$, which induces canonical homomorphism Δ of $G(64, K)$ onto $G(32, K)$. She takes two irreducible elements of $\sum = \sum(K), \alpha = (a_1, a_2, \ldots, a_{16})$ and

fridaya@uef.fi

Table 11. Generation time for the map (ms) $\boldsymbol{A}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case II

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	20	60	128	260	540
32	308	788	1776	3760	7716
64	3193	8858	23231	53196	113148
128	54031	137201	368460	950849	2164037

Table 12. Generation time for the map (ms) $\boldsymbol{A}(\boldsymbol{n}, \mathbb{F}_{2^{32}})$, case III

	Length of the word				
n	**16**	**32**	**64**	**128**	**256**
16	76	148	288	576	1148
32	1268	2420	4700	9268	18405
64	22144	40948	78551	153784	304240
128	460200	819498	1532277	2970743	5836938

$\beta = (b_1, b_2, \ldots, b_{16})$ of pseudorandom kind, use homomorhism $\eta' = \eta'_{64}$ of $\sum$ into $G(64, K)$ and gets elements $a = \eta'(\alpha)$ and $b = \eta'(\beta)$.

Alice forms string $h = (h_1, h_2, \ldots, h_t)$. $t = 16$ and the reverse string $rev(h) = (-h_t + h_{t-1}, -h_t + h_{t-2}, \ldots, -h_t + h_1, -h_t)$ for which $n = \eta'(h) = n$ and $n' = \eta'(rev(h))$.

She takes affine transformation T of the vector space F_q^{64}, $q = 2^{32}$ and its inverse T^{-1} and forms elements $a^1 = Tnan'T^{-1}$ and $b^1 = Tnbn'T^{-1}$.

Alice takes $d = (d_1, d_2, \ldots, d_t)$ and the pair $m = \eta'_{32}(d), m' = \eta'_{32}(rev(d))$. She forms $a^2 = Sm\eta'_{32}(\alpha)m'S^{-1} = Sm\eta'_{32}(\alpha)m'S^{-1}$ and $b^2 = Sm\eta'_{32}(\beta)m'S^{-1}$ where S is the bijective affine transformation of 32 dimensional vector space. She sends pairs $(a^1, a^2), (b^1, b^2)$, to Bob. Let us assume that Alice uses transformation T and S of kind 3. It means that cubical transformations a^1 and b^1 are given by lists with 399424 monomial terms and transformations a^2, b^2 are given by their 50720 monomial terms (see Table 6).

Bob takes word $w = x^{s1}y^{r1}x^{s2}y^{r2}\ldots$ of some lengths k, $k \geq 3$ (even or odd), where $s1, s2, \ldots$ and $r1, r2, \ldots$ are positive integers.

He substitutes a^1 and b^1 instead of x and y (or y and x) and compute corresponding transformation c from affine Cremona semigroup of 64 dimensional vector space over finite field F_q. The cubical transformations c is presented by its 388424 monomial terms. Bob substitutes the collision map c' via substitution of a^2 and b^2 in word w instead of x and y. Collision element c' is given by the list of its 50720 monomials.

fridaya@uef.fi

Bob sends the transformation c to Alice. She computes $c^1 = T^{-1}n'cnT$ which contains 1810 (monomial terms) (see Table 4). Alice computes $c^2 = \Delta(c^1)$ given by 770 terms. She reconstructs the collision map as $Sm\,c^2m'S^{-1}$.

In Sect. 5 we evaluated the complexity of our general algorithms as $O(n^{13})$. This bound is determined by complexity of computation of composition of two cubic maps. More accurate bound can be obtained via the maximal density $d_m(n) = den(G, F)$ of composed maps, which is $d_m(n)^4 n$. Recall that we define density *d(F)* of map F written in standard form by the list of $f_i \epsilon K[x_1, x_2, \ldots, x_n]$ as maximum of densities of $f_i, i = 1, 2, \ldots, n$. Let $g(n)$ be the global density of F which is number of all monomial terms in all f_i, $i = 1, 2, \ldots, n$. We can consider universal density of $u(n)$, which is maximal of global densities of elements from the set $B = T_1GT_2$, where G is a group defined via graphs An or Dn. As it follows from definition the inequality $u(n) \leq n^4$ holds Notice that in our algorithms we need $T2 = T1^{-1}$ and B is a group.

We can use upper bound $u(n)^4 n^{-3}$ for the complexity $C(n)$ of computation of decomposition of two maps from B.

Results of our simulation for the fields in the case 3, when matrices T_1 and T_2 are of general kind show, that $u(n) < 1.6n^3$. It means that $C(n) = O(n^9)$ which is essentially better than $O(n^{13})$ (see Tables 3 and 9 and Figs. 3 and 6).

Tables 1, 2, 4, and 6 and Figs. 1, 2, 4, and 5 demonstrate that ratio $u(n)/n^4$ tends to zero.

Tables 7, 8, 9, 10, 11, and 12 demonstrate the speed of generation of cubical maps depending on the cases 1, 2 and 3 of types of affine maps and the choice of graphs $\left(A\left(n, F_q\right)\right.$ or $\left.D\left(n, F_q\right)q = 2^{32}\right)$. Similar tables the cases of arithmetical and Boolean ring reader can find in [30] and [31].

7 On Safe Exchange of Symbolic Transformations

The symbolic nature of collision map can be used for task that differs from exchange of keys. We refer to it as the usage of DH *deformation symbolic rules*.

Let Alice have a free module K^n over commutative ring K. She has a subset Ω of K^n and polynomial map $f : K^n \to K^n$ such the restriction of $f|\Omega$ is an injective map from Ω onto $f(\Omega) = \Gamma$. Additionally Alice has an algorithm to solve in polynomial time equation x = b with respect to unknown x from Ω and b from Ω.

Alice and Bob use *tahoma word protocol* or symbolic Diffie-Hellman protocol to elaborate the collision map g acting on K^n. After this step Alice sends Ω and transformation $h = f + g$ to Bob. Now Bob can get f as h-g. He writes plaintext p from Ω and sends ciphertext $c = f(x)$. Alice uses her data for the decryption.

Remark 7.1.
Notice that new algorithm is still asymmetrical because Bob can encrypt but not decrypt. The encryption rule is known to trusted customer (Bob) but adversary has no access to it. In fact such access is protected by word problem in semigroup of transformations of K^n or discrete logarithm problem in corresponding affine Cremona semigroup.

fridaya@uef.fi

Other Deformations.
Alice and Bob agree (via open channel) on a deformation rule $D(f)$ for multivatiate rule f from affine Cremona semigroup. For example, it can be multiplication, i.e. f is the rule $x_i \to f_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$ and g is the rule $x_i \to g_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$ and Alice sends tuple of polynomials $f_i g_i$, $i = 1, 2, \ldots, n$. Bob uses division to restore f. Instead of addition deformation rule (sending of $x_i \to f_i(x_1, x_2, \ldots, x_n) + g_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$)Alice can use deformation with adding an element $K[x_1, x_2, \ldots, x_n]^n$ obtained from g via the usage of s-time conducted derivation δ^s, where $\delta = d/x_1 + d/d\,x_2 + \ldots + d/d\,x_n$ (rule $x_i \to f_i(x_1, x_2, \ldots, x_n) + \delta^s g_i(x_1, x_2, \ldots, x_n), i = 1, 2, \ldots, n$).The last deformation is interesting because in many cases we can achieve the equality of degrees for f and $D(f)$. It is easy to continue this list of possible deformation rules.

Remark 7.2.
Let us assume that $\Omega = K^n$. So $f = f(n)$ is a bijection. Assume that degrees of nonlinear maps $f(n)$ are bounded by constant d. Let us assume that the adversary has option to intercept some pairs plaintext - ciphertext (leakage from Bob's data). In case of interception of $O(n^d)$ adversary has chance for a successful linearisation attack and get the map f. For example if $d = 3$ then linearisation attack cost is $O(n^{10})$. After that adversary has to find the inverse function for f like in the case of multivariate public key.

To prevent "transition to knowledge" of an encryption multivariate map Alice (or Bob) can arrange a new session with protocol and a transmission of new deformed encryption rule for which secret data for decryption is known.

Remark 7.3.
The technique of linearisation attacks on non-bijective maps or maps $f(n)$. of unbounded degree and low density is not developed yet.

8 Conclusion

We propose Post Quantum Cryptography information security solutions based on the complexity of the following problem.

Cremona Semigroup Word Decomposition (CSWD).
Let $S(K^n)$ be a family of affine Cremona semigroups, which are totalities of multivariate transformations of affine space K^n defined over finite commutative ring K, Let $S = <F_1(n), F_2(n), \ldots, F_k(n)>$ be a sub-semi-group generated by elements $F_i(n), i = 1, 2, \ldots, k$, $k \geq 2$ from S. For given $G \in S$ find its presentation in the form $G = G_1 G_2 \ldots G_t$ where $G_i = \{F_1(n), F_2(n), \ldots, F_k(n)\}$ are elements of $S(K^n)$.

We assume that each element $F_i(n)$ given in its standard presentation. Written above problem is hard, i.e. polynomial algorithms to solve it in terms of Turing machine or Quantum computer are unknown.

Our key exchange protocols and El Gamal type cryptosystems based of Dat $F_1(n), F_2(n), \ldots F_k(n)$ constructed in terms of symbolic computations in linguistic graphs.

fridaya@uef.fi

Generated platforms S are stable, i.e. degrees of elements from S are bounded by constant c. In our examples $c = 3$ and this fact allows to provide complexity estimates for both correspondents. Computer simulations demonstrate an interesting phase transition effect, which allow to predict the density of the collision maps of key exchange protocols and their inverse forms. This effect demonstrates feasibility of proposed cryptographic schemes.

We hope that introduced algorithm can be considered as serious candidates to be postquantum cryptographical tools. We believe that future studies of cryptanalitics confirm that CSWD problem remains unsolvable on ordinary Turing Machine and Quantum Computer under the condition of stability of platform S.

References

1. Ding, J., Gower, J.E., Schmidt, D.S.: Multivariate Public Key Cryptosystems, vol. 25, p. 260. Springer, Advances in Information Security (2006)
2. Koblitz, N.: Algebraic aspects of cryptography, p. 206. Springer (1998)
3. Goubin, L., Patarin, J., Yang, B.O.: Multivariate Cryptography. Encyclopedia of Cryptography and Security, 2nd edn, pp. 824–828 (2011)
4. Moldovyan, D.N., Moldovyan, N.A.: A New Hard Problem over Non-commutative Finite Groups for Cryptographic Protocols. In: International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security, MMM-ACNS 2010, pp 183–194 (2010)
5. Sakalauskas, L., Tvarijonas, P., Raulynaitis, A.: Key Agreement Protocol (KAP) using conjugacy and discrete Logarithm problems in group representation level. INFORMATICA **18**(1), 115–124 (2007)
6. Shpilrain, V., Ushakov, A.: The conjugacy search problem in public key cryptography: unnecessary and insufficient. Appl. Algebra Eng. Commun. Comput. **17**(3–4), 285–289 (2006)
7. Kahrobaei, D., Khan, B.: A non-commutative generalization of ElGamal key exchange using polycyclic groups. In: IEEE GLOBECOM 2006—2006 Global Telecommunications Conference [4150920] (2006). https://doi.org/10.1109/glocom.2006
8. Myasnikov, A., Shpilrain, S., Ushakov, A.: Group-based Cryptography. Birkhäuser Verlag, Berlin (2008)
9. Cao, Z.: New Directions of Modern Cryptography. CRC Press, Taylor & Francis Group, Boca Raton (2012). ISBN 978-1-4665-0140-9
10. Fine, B., et. al.: Aspects of Non abelian Group Based Cryptography: A Survey and Open Problems (2011). arXiv:1103.4093
11. Myasnikov, A.G., Shpilrain, V., Ushakov, A.: Non-commutative Cryptography and Complexity of Group-theoretic Problems. American Mathematical Society (2011)
12. Anshel, I., Anshel, M., Goldfeld, D.: An algebraic method for public-key cryptography. Math. Res. Lett. **6**(3–4), 287–291 (1999)
13. Blackburn, S.R., Galbraith, S.D.: Cryptanalysis of two cryptosystems based on group actions. Advances in Cryptology—ASIACRYPT '99. Lecture Notes in Computer Science, vol. 1716, pp. 52–61. Springer, Berlin (1999)

fridaya@uef.fi

14. Ko, K.H., Lee, S.J., Cheon, J.H., Han, J.W., Kang, J.S., Park, C.: New public-key cryptosystem using braid groups. In: Advances in Cryptology—CRYPTO 2000, Santa Barbara, CA. Lecture Notes in Computer Science, vol. 1880, pp. 166–183. Springer, Berlin (2000)
15. Maze, G., Monico, C., Rosenthal, J.: Public key cryptography based on semigroup actions. Adv. Math. Commun. **1**(4), 489–507 (2007)
16. Kropholler, P.H., Pride, S.J., Othman, W.A.M. Wong, K.B. Wong, P.C.: Properties of certain semigroups and their potential as platforms for cryptosystems. Semigroup Forum **81**, 172–186 (2010)
17. Lopez Ramos, J.A., Rosenthal, J., Schipani, D., Schnyder, R.: Group key management based on semigroup actions. J. Algebra and Its Appl. **16** (to appear in 2019)
18. Kumar, G., Saini, H.: Novel Noncommutative Cryptography Scheme Using Extra Special Group, Security and Communication Networks, Article ID 9036382, p. 21 (2017). https://doi.org/10.1155/2017/9036382
19. Ustimenko, V.: On the families of stable transformations of large order and their cryptographical applications. Tatra Mt. Math. Publ. **70**, 107–117 (2017)
20. Ustimenko, V.: On desynchronised multivariate El Gamal algorithm. Cryptology ePrint Archive, p. 712 (2017)
21. Ustimenko, V.: On new symbolic key exchange protocols and cryptosystems based on hidden tame homomorphism. Dopov. Nac. akad. nauk Ukraine **10**, 26–36 (2018)
22. Lazebnik, F., Ustimenko, V., Woldar, A.: New Series of Dense Graphs of High Girth, Bull (New Series) of AMS **32**, 73–79 (1995)
23. Ustimenko, V.: Coordinatization of Trees and their Quotients, in the Voronoj's Impact on Modern Science, vol. 2, pp. 125–152. Institute of Mathematics, Kiev (1998)
24. Ustimenko, V., Romanczuk, U.: On Extremal Graph Theory, Explicit Algebraic Constructions of Extremal Graphs and Corresponding Turing Encryption Machines. Ser. Stud. in Comput. Intell. vol. 427, 257–285 (2013)
25. Ustimenko, V.: Maximality of affine group, hidden graph cryptosystem and graph's stream ciphers. J. Algebra Discrete Math. **1**, 51–65 (2005)
26. Ustimenko, V.: Linguistic Dynamical Systems, Graphs of Large Girth and Cryptography. J. Math. Sci. **140**, 412–434 (2007)
27. Ustimenko, V.: Graphs with special arcs and cryptography. Acta Applicandae Mathematicae (Kluwer) **74**, 117–153 (2002)
28. Ustimenko, V:. On the graph based cryptography and symbolic computations. In: Proceedings of International Conference on Application of Computer Algebra, ACA-2006,v1, Serdica Journal of Computing, pp 131–156 (2007)
29. Klisowski, M.: Zwiększenie bezpieczeństwa kryptograficznych algorytmów wielu zmiennych bazujacych na algebraicznej teorii grafów, PhD thesis, Czestochowa, 2014
30. Klisowski, M., Ustimenko, V.: Graph based cubical multivariate maps and their cryptographical applications. In: Advances on Superelliptic Curves and Their Applications. IOS Press, NATO Science for Peace and Security series – D: Information and Communication Security, vol. 41, pp. 305–327 (2014)
31. Klisowski, M., Ustimenko, V.A.: On the comparison of cryptographical properties of two different families of graphs with large cycle indicator. Math. Comput. Sci. **6**(2), 181–198 (2012)

fridaya@uef.fi

A Stealth Key Exchange Protocol

Khan Farhan Rafat(✉)

Riphah International University, RISE, Islamabad 44000, Pakistan
22371@students.riu.edu.pk

Abstract. Cryptography, the science of secret writing, is in use since long to inarticulate the legitimate contents to prevent their unauthorized exposure. To ensure secure communication between communicators, cryptographic contrivances such as encryption and authentication relies on the conjecture of prior sharing of some secret called cryptographic or secret key that reign the security of the underneath algorithms. Hence, safekeeping and handling of such keys at all stages of encryption and decryption inclusive of its generation, storage, distribution, archiving and destruction are imperative. Kirchhoff, in 1883, has already elucidated on the significance of such keys by expounding that the security of a public domain encryption algorithm lies in protecting the key used in communication. Regardless of the key being symmetric or asymmetric, its distribution/exchange over an insecure channel such as the internet, together with the increasing concern among technologically advanced countries on suppressing encryption, calls for an oblivious mechanism for its undercover traversal. Steganography, the purpose of which is to embed information seamlessly inside some carrier, is the technique used for stealth information exchange, and hence, seems ideal for key management and its successive dissemination. Through this research endeavor, we have proposed a novel stealth key-exchange solution, which makes use of interactive computer-mediated technologies, cryptography, and steganography. Our proposed method adlibs the Diffie-Hellman key exchange algorithm by disregarding the direct routing of any data relating to key exchange, and hence protects it from man-in-the-middle attack by keeping intact the anonymity of communicators.

Keywords: Anonymity · Covert channel · Diffie-Hellman key exchange · RSA

1 Introduction

The widespread usage of the internet has not only led to an enormous growth of information [1] but also has changed the mannerism in which people, from naïve computer users to professionals, communicate with one another [2]. Today one seldom finds long awaiting ques across banks, post offices, shopping centers, travel agencies because technology has paved the way for similar online electronic transactions via hi-tech smart gadgetry [3] such as cell phones, Tabs and laptops. The internet of things (IoT) has further facilitated end users in getting rid of infrastructure barriers through ubiquitous computing [4]. However, the recently evolved communication security solutions for data security such as for Big Data handling, are resource intensive [5], which may range but are not limited to manufacturing, implementation, and

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 675–695, 2019.
https://doi.org/10.1007/978-3-030-22868-2_48

fridaya@uef.fi

computational cost to data retention and its transmission. The energy requirements [6] of such systems also seldom get an appraisal in its related discussions. Also, from conventional to the state-of-the-art communication security solutions, the biggest impediment that remains unsettled is to maintain the anonymity of communicators [7]. That is, non-disclosure of "who is communicating with whom." Communication via a third party such as a Certificate Authority (CA) is least agreeable these days. Transmitting encrypted contents or stego-object (cover carrying hidden data/information) over an insecure channel is vulnerable to attacks from the hostiles [8]. Moreover, the recent campaigns and legislation passed by technologically advanced countries such as Australia, United States of America (US), and United Kingdom (UK) to gain access to encrypted data and associated keys [9–11], it is but desirous to evolve hi-tech communication security schemes, which besides being cost-effective can hide information's existence. Doing so in a stealth manner will also keep the anonymity of the communicators intact.

Taking the worst-case scenario of a networking environment devoid of encryption, this research aimed at proposing a new stealth out-of-band key exchange protocol to facilitate secure communication thereby protecting the privacy and anonymity of the individuals involved.

The paper has the following format. Section 2 gives details on related literature regarding anonymity, privacy, and key exchange together with allied advantages and limitations of the protocols, techniques, and procedures discussed in those contexts. The problem domain arrived at through reviewed literature comes under discussion in Sect. 3. Section 4 elucidates on the research question. Section 5 elaborates the proposed methodology, while Sect. 6 discusses its pros and cons. The argumentation on the stealth key exchange protocol evolved comes in Sect. 7. Section 8 highlights the research contribution. Section 9 concludes the discussion.

2 Literature Review

Following discussion elaborates on the significance of anonymity, privacy, and secure key exchange in today's digital world in the context of the allied state-of-the-art protocols evolved and implemented to achieve the said objectives.

2.1 Anonymity

The word anonymity is of Greek origin, which means "no name" [12]. Being anonymous gives people the freedom to experiment and experience new things, and deliberate ideas without the fear of criminal judgment [13]. However, with the launch of the internet, online social media has gone trendy and has gained acceptance as a lifestyle to the extent that people unwittingly overexpose themselves. For the most, the cyberspace is an imaginary arena devoid of boundaries and beyond the rules and laws that administer the real world [14, 15]. As a consequence, pertinent information sways freely across the web under the tag of personal preference [15]. Hence, anyone harvesting information can easily imitate him/herself as the one cropped by merely going through the related social media accounts of the target [16]. Moreover, people's

fridaya@uef.fi

addiction to smart devices in conjunction with social media has embarked many if not the most, with a reluctant attitude towards staying unidentified [17–19].

We behold and are taking part in the most critical information technology uprising in the history of manhood as our society endures the evolution from a mostly paper-based world to a fully digitalized one [20]. This new digitized world called the "Internet of Things," (IoT) comprises of an incredibly different set of conversant everyday technologies, such as automobiles, cameras, medical devices, satellites, televisions, drones, traffic lights, building fire/security systems, tablets and smartphones [4]. People have got themselves acquainted with those technologies and have dispersed the allied devices to every corner of this new world. These recent technologies are not only captivating, but also alcoholic and addicting [17–19], leaving people with a higher risk of losing their privacy, and security [21].

Following are some of the protocols for ensuring anonymity over the internet:

Proxy Server. A proxy server initiates requests of the local network which are meant for some external system as one of its own by sitting in between the two [22] and removing the Internet Protocol (IP) address of that of the client. The information received from the target server is then resent to the client after affixation of its IP address by the proxy server. Hence, the identity of the client site remains anonymous to the target system.

Pros. Mostly used as a caching strategy to reduce outgoing requests and increase perceived browser performance. Encryption and content filtering are among other services offered by the proxy servers.

Cons. The standard implementation did not offer encryption between client-server and hence was vulnerable to attacks from hackers. Because of its efficient cache system, user's passwords browsed websites, or any reliable data are traceable by the service provider.

Virtual Private Network (VPN). A VPN [23] protocol, a client-server architecture, comprises of networking and security (encryption) protocols. The client requests the server with a site address (called uniform resource locator – 'URL') to which it does not have direct access. The communication between the client side and the server side is encrypted, and hence, the intermediate hops/hackers cannot read the data exchanged. The server detaches the client's IP address, affixes its own in its place, and access the said 'URL.' It then returns the resource fetched back to the client after detaching its IP address and inserting that of the client in its place.

Pros. VPN facilitates its client to access proscribed contents, and those, which are geographically restricted. Encryption service protects against unauthorized spying on the data exchanged between the client and the server.

Cons. Using VPN might be illegal in the country from where that 'URL'-request is initiated. Performance/speed issues exist because of encryption and corresponding decryption of contents surfed or downloaded. The VPN service might also monitor client's activity for any illicit or otherwise purpose. It may add extra cost to the network connection.

fridaya@uef.fi

Tor – The Anonymity Network. Tor [24] is intended to be an anonymity network that allows anonymous access to the internet and uses encryption to obscure data exchanged over the network. Tor network is a collection of servers hosted by volunteers that help people to achieve privacy and security on the Internet. Tor's users operate Tor by connecting to their desired address through a series of multiple virtual tunnels in place of approaching it directly. Because, at any point in time, the intermediate nodes are only aware of the node that is behind and in front of a particular node. Doing this allows people to share information over public networks but without compromising their privacy.

Pros. Random selection of relay nodes (servers), at specified intervals, enhances network security.

Cons. The starting/entry and terminating/exit nodes are aware of the source and destination points, and the client side request is visible at the terminating node. Tor offers privacy only for browsing whereas other apps may compromise a client's anonymity (for example Tor documentation prohibits the use of torrent software over it). It requires Firefox as its browser which in turn has some vulnerabilities. Continuous selection of random relay nodes, however, has a degrading effect on the network speed. Moreover, Tor browser cannot run Flash-based videos.

Summary. It follows from the above discussion that people value anonymity as high [7, 25], but the protocols evolved have yet not addressed the issue of anonymity to its full. Hence, attaining anonymity is one of the significant challenges of this digital age [26].

2.2 Privacy

Privacy is the degree to which an individual is willing to share personal information, with whom, and for what purpose [27]. Today, it is hard to isolate privacy from the digital world as the likelihood to intervene in someone's privacy has amplified many folds because of the state-of-the-art technology. The history of privacy dates back to that of humanity, but often not considered as their legitimate right [28]. It is because what deliberated as private in one region may differ in that of another. The issue, in the purview of several renowned privacy scholars and the European Court of Human Rights, is in not having a formal agreement on any legal definition of privacy protection [29]. The significance of privacy lies in its close bonding with self-esteem, autonomy, and liberation of the individual, which faces a challenge each day in this rapid technological progression of the information society [30, 31].

Encryption and Steganography are the means to ensure privacy as follows:

Encryption. The process of making legible contents unintelligible under the control of some entity called key is called encryption [29, 32]. It has demonstrated its suitability for use in the digital environments [33] and has been extensively deployed by diversified actors to ensure the fortification of information and communication for viable, private and public safeties [34]. The significance of encryption has full backing and acceptance from people as regards free and open use of the internet [35]. Encryption, in particular, is vital for freedom of expression [36], private communication, and confidentiality. Let C denotes the ciphertext for a message M, which is encrypted and

fridaya@uef.fi

decrypted using some algorithm E and D under the control of some key K, then the process of encryption and decryption takes the form as given by Eqs. (1) and (2), respectively.

$$C \leftarrow E_k(M) \tag{1}$$

$$M \leftarrow D_k(C) \tag{2}$$

Before the introduction of 'asymmetric key' cryptography (explained above), the secure key exchange remained a problem. Based on 'hard-to-solve-problem' of factorization [37], Public key cryptography solved came to rescue by using a public-private key-pair comprising of two distinct but mathematically interrelated keys. One of the keys designated as the 'public key' is universally available for encryption, say on the internet, whereas the leftover key called 'private key' is known only to its owner and used for decryption purposes. *Equations (3) and (4) demonstrate* this concept. Principally, the problem was that of assurance that a particular public key belonged to its hypothetical owner in actual. 'Spoofed' identity, however, remained a significant impediment mainly when the two communicators have never met with each other [38].

$$C \leftarrow E_{Bob's\ Public\ Key}(M) \tag{3}$$

$$M \leftarrow D_{Bob's\ Private\ Key}(C) \tag{4}$$

Privacy-Preserving Protocols. The two widely used privacy-preserving protocols are (i) Transport Layer Security (TLS) [38], and (ii) IPSec Protocol [39]. Both of these provide integrity and authentication of the messages exchanged along anti-replay protection, and confidentiality through encryption.

Pros. Using encryption prohibits the hostiles from gaining access to vital information.

Cons. Encryption is not concerned with data obscurity. The unintelligible text may raise suspension regarding ongoing secret communication, which is liable to get exploited by the hostile. As the number of users grows, the number of secret keys required increases. Key management is an overhead. The criminals are equal beneficiaries in using encryption when it comes to securing information.

Steganography. The word Steganography is a mix of two Greek words - steganos (στεγανός), meaning "covered/concealed/protected," and graphein (γράφειν) meaning "writing" [40], and dates back to 440 B.C. [41]. It is the art and science of covert communication capable of hiding a file, note, image, or video within some carrier which can well be another message, image or video [41]. The image is the most widespread medium because of having highly redundant data [42] for manipulation and exploits the confines of the Human Visual System (HVS) [43].

The most popular image-based steganographic technique amongst many is called the Least Significant Bit Embedding (LSB) steganography [44]. An image corresponds to a contiguous array of bytes where each byte represents a pixel's value (for grayscale

fridaya@uef.fi

images) having a range from 0 to 255. Research has established that the least significant bit has a negligible contribution towards an image composition and hence manipulating that bit is likely to pass unnoticed because of HVS limitations. Thus, each message bit inserted at LSB position during embedding gets extracted accordingly by the recipient while retrieving the hidden message.

Pros. Steganography hides not only the existence of a message but also the parties engaged in communication. It can deceive the intrusion detection systems. Encrypting the message before bit embedding adds extra security layer which can entail steganalysis.

Cons. Poor choice of the cover may attract an eavesdropper. Depending on the steganographic technique used, even small messages may require a more substantial cover.

Summary. Encryption and Steganography facilitate users in maintaining their confidentiality. However, encryption intends to make information meaningless for the unauthorized whereas the sole purpose of steganography is to hide that such details exist. The two techniques, although polls apart, may act as a double-edged sword when used together to safeguard the information security frontiers.

2.3 Key Exchange

In general, a key act as a clue to solve a given problem. From a cryptographic standpoint key influences the functional consequence/outcome of a cryptographic procedure [32]. In encryption, a key stipulates the conversion of the message into ciphertext while reversing it for decryption. Keys also governs the transformation process [45] in other cryptographic procedures, such as message authentication codes (MAC) and digital signature algorithms.

For encryption to be successful, the communicating parties, mainly when distant apart, need to agree on specific prerequisites used in the process for which some exchange mechanism must exist. For example, if the communicators agree on using a code, the same codebook is required at the respective ends [46]. In case they agree on using a cipher then appropriate keys must be known. For symmetric key cryptography, the same key is needed. However, for asymmetric key cryptography, a public/private key pair is required for each communicator [47]. The exchange of keys is either in-band where keys are traded using the same communication channel or out-of-band where key exchange takes place through a channel different than the one used for communication [48].

One way of distributing keys is via a courier. However, the procedure is cumbersome and add more cost as the number of communicators' increases. Further, the courier may deceive in the process; the keys are also susceptible to be stolen or lost.

Diffie-Hellman Key Exchange Algorithm. The author in [49] proposed an algorithm to exchange keys over an insecure channel. Following explicate on the said procedure. Alice and Bob agree on some prime number (say) P and its primitive root (say) G. Both selects a random number, known only to them, at their end. They perform some computation using that random number together with P, and G. Alice and Bob then

fridaya@uef.fi

exchange the computed result with one another. Both perform another calculation at their respective ends using the calculated results they have received, together with P and G. The output of this computation yields the same result (key) at either end. Figure 1 illustrates the process.

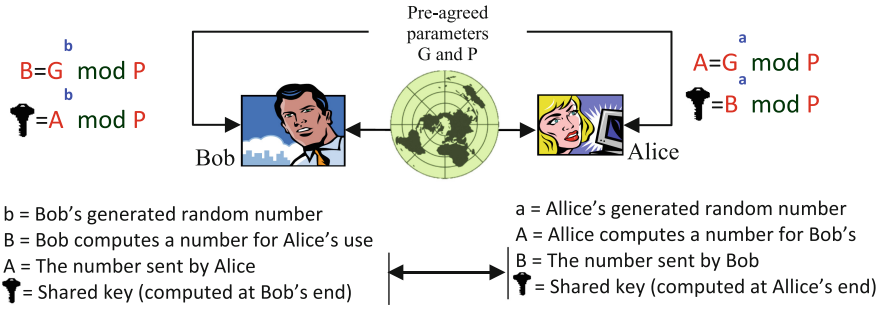


Fig. 1. Diffie-Hellman key exchange algorithm

Pros. Facilitates the key exchange over an insecure channel.

Cons. The sender and recipient have no assurance as for whether they are communicating with one another. The algorithm does not allow for authentication or asymmetric key exchange.

Public Key Infrastructure (PKI). Public key infrastructure (PKI) [50] addresses the authentication issue where the interested users approach a 'certificate authority' (CA) for the issuance of a digital certificate (DC). Signed with the CA's private key, it corresponds to an unaltered (non-spoofed) authentication of identity for the other users before the commencement of message exchange. As shown in Fig. 2, Bob who wants to communicate with Alice must first register himself with a CA. The CA sends to Bob a digital certificate signed with its private key that also has Bob's public key along with other parameters. Bob sends this certificate to Alice with a request for communication. Before answering to Bob's request, Alice gets his certificate authenticated by the CA. Alice also sends her digital certificate to Bob who also performs the same steps for verifying her identity.

Pros. The communicating parties must have trust on CA. It prevents an adversary from impersonating as someone else.

Cons. Storing digital certificates on local machines is susceptible to attack or compromise. Who authorized the CA to be the authority? is arguable. The users are unaware of the security of the verifying machine.

Internet Key Exchange (IKE). IKE [51], a tunneling protocol, is a component of the widely used protocol suite named Internet Protocol Security (IPSec). It is responsible for negotiating security parameters, authentication and exchanging keys between communicating parties. The parameters exchanged include the HASH, Authentication, Group (such as Diffie-Hellman key exchange), and Encryption algorithms opted for use

fridaya@uef.fi

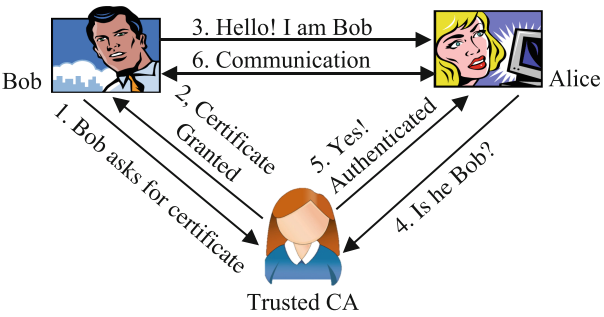


Fig. 2. PKI authentication mechanism

along with duration (time) of that particular session. The said exchanges collectively come under the title 'Phase 1' of IPSec. IKE operates in two modes called Main Mode (which involves six messages for such a transaction) and aggressive Mode (where the parameters are sent in bulk at once by either end) respectively. Its primary aim is to produce the same symmetric key across both communication end.

Pros. Security, reliability, and speed are the salient traits of this protocol.

Cons. Lately, attacks launched successfully against IKE, the rectification of which is underway [52].

3 The Problem Domain

Table 1 summarizes some limitations of the existing protocols/techniques/procedures discussed in the context of anonymity, privacy, and key exchange.

Privacy is the core justification behind encryption but is disputably under threat of attrition due to calls for its removal [9] as elucidated in Table 2, which depicts some noteworthy comments concerning the use or otherwise of encryption.

Lately, a growing drift is noticed among states such as America (US), England (UK), and Australia on gaining access to keys for decoding encrypted information which is a direct attack on peoples privacy as tabulated above and needs careful scrutinization, whereas countries like China and Russia have already prohibited encryption for their citizens [54].

Foregone above, the core problems derived from the literature reviewed include:

(a) **Anonymity is not the objective of existing protocols/techniques/procedures**, which is supported by the fact that it is a less deliberated subject in the Internet Engineering Task Force (IETF) than privacy [55] or security [56].
(b) **Removing encryption** from communication **will have a catastrophic effect** not only **on the privacy** at an individual level but on the **sovereignty** of the Governments at a broader scale.
(c) **Cryptography** alone fell short in preserving the anonymity of individuals.
(d) **Steganography** does not encrypt data/information. Hence, breach of embedding algorithm renders subsequent data exchange insignificant.

fridaya@uef.fi

Table 1. Summary: Limitations of the existing protocols/techniques procedures.

Protocols	Limitations
Proxy server	• Because of the excellent cache system of the proxy server user's passwords, browsed websites or any reliable data is traceable by that service provider
VPN	• Using VPN might be illegal in the user's country • Performance issues while using the private network • The VPN service might monitor user's activity and use that data • It might add more cost to the network connection
Tor network	• Privacy is only for browsing; other apps may leak user's anonymity (don't use torrent software over the tor network) • Reduced network speed • Google may not be the best engine while using Tor • Inability to watch flash-based videos
Encryption	• Encryption is not concerned with data obscurity • The unintelligible text may raise suspension regarding ongoing secret communication, which is liable to get exploited by the hostile • As the number of users grows, the number of secret keys required increases • Key management is an overhead • Criminals are equal beneficiaries in using encryption when it comes to securing information
Steganography	• Poor choice of the cover may attract an eavesdropper • Depending on the steganographic technique used, even small messages may require a more extensive cover
Diffie-Hellman	• Can not be used for asymmetric key exchange • Can not be used for signing Digital Signatures • Devoid of the authentication mechanism, the algorithm is susceptible to a man-in-the-middle attack • Can not be used for asymmetric key exchange
PKI	• The associated risk is the compromise of every user in situations where the CA itself is compromised such as several prominent public failures by various certificate authorities • Questions such as who gave the CA the authority to grant such authorizations and who made it trusted • How do users know that how secure is the verifying computer
TLS	• An attacker (MITM) knows: (i) the Internet Protocol addresses of the client and server (no anonymity) (ii) the server certificate, including the server name – as these are in plain (iii) the approximate length of the URL, although encrypted
IKE	• IKE protocol including all of its variants is breakable. Given two entry points [52] have demonstrated the skipping of the cross-protocol authentication

fridaya@uef.fi

Table 2. Some noteworthy comments on using cryptography.

RFC 7258	Pervasive Monitoring is widespread surveillance through the gathering of protocol artifacts such as headers and is a technical attack that needs mitigation [53]
	Prohibiting Encryption for usage is an attack on an individual's privacy
Google CEO Sundar Pichai Backs, February 2016	Forcing companies to enable hacking could compromise users' privacy
Facebook CEO Mark Zuckerberg, February 2016	We believe in encryption
Former US President Barack Obama, March 2016	"… … … how do we solve or disrupt a terrorist plot?…There has to be some concession to the need to be able to get into that information somehow"
US Attorney General Jefferson Beauregard Sessions III, January 2017	"Encryption serves many valuable and important purposes……It is also critical, however, that national security and criminal investigators be able to overcome encryption, under lawful authority, when necessary to the furtherance of national security and criminal investigations"
Home Secretary Amber Rudd, March 2017	"We need to make sure that organizations like WhatsApp, and there are plenty of others like that, don't provide a secret place for terrorists to communicate with each other"

4 Research Question

Derived from the problem domain the research question to discourse follows as under:

> Are anonymity and privacy achievable under strict network surveillance and in the absence of secure (through encryption) protocols?

5 Proposed Solution

In [57] author suggested a model for covert communication with the scenario as follows: Alice and Bob locked up in prison in separate cells, had to decide on their escape plan. The two could communicate only through Warden Wendy who was likely to send them imprisonment subject to the detection of any secret information and thereby denying them any further message exchange. Before going into lockup, Alice and Bob agreed on some parameters to embed and detect the oblivious communication. In the purview of the pioneered model, there seems an ample need to keep the covertness of the secret-message exchange intact. On the analogy of the (CAPTCHA) [58], the proposed methodology also requires manual intervention by the sender and the receiver of the message at their respective ends, which ensures authenticity, and

fridaya@uef.fi

reliability of the activities performed. The scheme works as follows: Alice and Bob maintain their public profiles such as WhatsApp, Facebook, or Twitter accounts along with a (self-maintained) personal web page. Among other contents on their web page, they made their respective public keys available. They also designate a place for stealth exchange of key-parameters under the label "For Delirious Hubby/Hellen" respectively to display four emoticons in sequence. The initial letters "F," "D" and "H" of the label explicates on "follow Diffie-Hellman" key exchange algorithm. Hovering mouse over the icons in the first row displays the file names in hexadecimal format, which covertly but collectively represent the modulo prime p, its primitive root r, the result of the computation (r^x mod p, where x denotes the number selected at random by Alice and Bob), followed by a randomly selected nonce.

5.1 Procedure at Alice's End

Let Alice opt for modulo prime 32579, which has 12816 primitive roots of which she selects 32367. She then picks 20348 at random and computes 11343 as the value to be communicated to Bob via (r^x mod p), that is, $x \leftarrow (32367^{20348} \bmod 32579) = 11343$. She also selects 453310 as her nonce for that particular exchange.

Alice first encrypts each of the four key-exchange parameters including the nonce using Bob's public key with a pre-agreed encryption algorithm such as Ron Rivest, Adi Shamir, and Leonard Adleman (RSA) [59]. Let the output be "E436F5B0 6A98F661 DC7F2045 7C76AEC3". Next, she assigns the respective output (in Hex format) as new filenames to each of the four icon files. She then updates her webpage's reserved section with new icon filenames as shown in Fig. 3. After that Alice tosses a pre-agreed emoji (exclusively reserved for Bob) as her public profile/App status update as shown in Fig. 4, thereby indicating to Bob to visit her webpage.

p	*r*	*x*	*Nonce*
For Delirious Hubby			
E436F5B0.ico	6A98F661.ico	DC7F2045.ico	7C76AEC3.ico

Fig. 3. Alice's stealthily published parameters

Fig. 4. Allice's signal to Bob to visit her page

5.2 Procedure at Bob's End

Bob visits Allice's webpage on noticing her status update and hovers his mouse over the icons under the label "For Delirious Hubby". He then notes down the names of the

fridaya@uef.fi

icon files taking one icon at a time in sequence, the concatenated string thus formed is processed through RSA using his private key are 32579, 32367, 11343, and 453310.

Let Bob selects a primitive root 390. Using Allice displayed values he computes secret shared key $K \leftarrow (11343^{390} \mod 32579) = 629$. He then calculates $y \leftarrow (11343^{390} \mod 32579) = 17014$, and increments the nonce used by Alice by one that is 453311. Each of the four values processed through RSA using Alice's public key renders a Hex output. Let it be "D3AEA43F 6F92CDC2 9A3E112B B7BE93A7". Bob assigns those as new file names to the four icons displayed on his webpage under the label "For Delirious Hellen." The values formatted by Bob as filenames take the form as shown in Fig. 5. Bob changes his status to pre-agreed emoji (exclusive for Alice) on his social webpage/App as shown in Fig. 6, indicating Allice to visit his webpage.

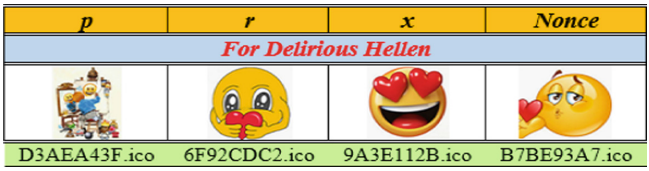


Fig. 5. Bob's stealthily published parameters

Fig. 6. Bob's signal to Alice to visit his web page

5.3 Key Derivation at Alice's End

Responding to Bob's pre-agreed status change notification, Allice visits Bob's web-page, extracts, decrypts, and compares the values of p, r, x and nonce reply with those against her's. After that, she calculates the secretly shared key as $K \leftarrow (17014^{20348} \mod 32579) = 629$.

Figures 7 to 17 diagrammatic illustrates the steps required in the proposed stealth key exchange protocol.

Fig. 7. Alice and Bob maintain personal web pages/sites with their Public Key displayed. Both also have their social media accounts such as on Twitter, Facebook, WhatsApp or such similar apps for instant messaging.

fridaya@uef.fi

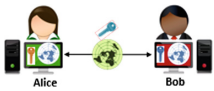


Fig. 8. Alice and Bob want to exchange a key over an insecure channel such as the internet.

Fig. 9. Alice and Bob do not want Eve (Eavesdropper) to know about any such information exchange.

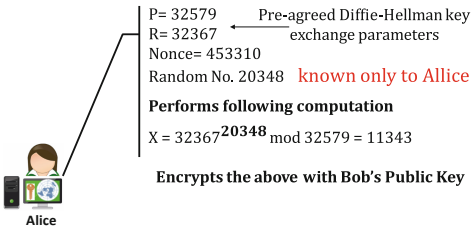


Fig. 10. As a part of their pre-agreed strategy, Alice first selects the Diffie-Hellman key exchange parameters P and R, and a random number (known only to her). Using those parameters, she then computes a value 'X = 11343' for sending it to Bob.

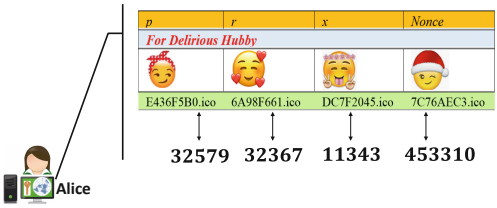


Fig. 11. Using Bob's Public Key, Alice encrypts her selected parameters along with the nonce and the result of the computation performed. She then separately assigns the ciphertext as filenames to the icons displayed at a pre-agreed place on her web page/site.

Fig. 12. Next, Alice updates her status on the social media account/profile(s) thereby signaling Bob to visit the web page/site, which she maintains.

fridaya@uef.fi

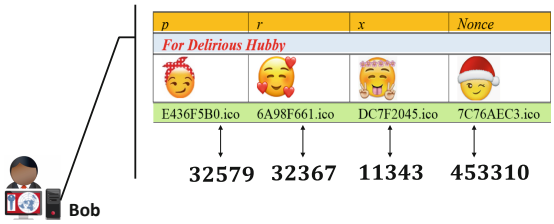


Fig. 13. On receipt of notification regarding Alice's status update, Bob visits her web page/site and notes down the encrypted filenames assigned to the four icons. He then decrypts the filenames using his Private Key and extracts the hidden parameters sent by Alice. After that Bob computes the result 11343 at his end using those parameters.

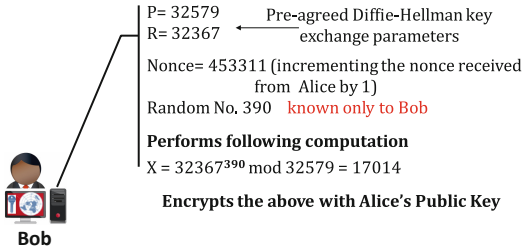


Fig. 14. Next, Bob selects the Diffie-Hellman key exchange parameters P and Q, and a random number (known only to him). Using those parameters, he then computes a value 'X = 17014' for sending it to Alice.

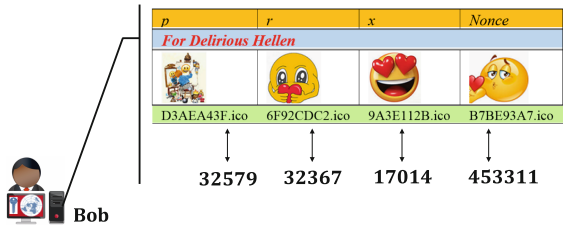


Fig. 15. Using Alice's Public Key, Bob encrypts his selected parameters along with the incremented nonce and the result of the computation performed. He then separately assigns the ciphertext as filenames to the icons displayed at a pre-agreed place on his web page/site.

fridaya@uef.fi

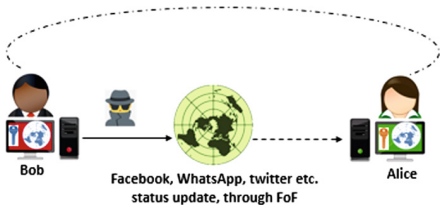


Fig. 16. Bob then updates his status on the social media account/profile(s) thereby signaling Alice to visit the web page/site, which he maintains. Next, using the number (11343) received from Alice (Fig. 13 refers), Bob computes $K \leftarrow (11343^{390} \bmod 32579) = 629$, which is the actual stealthily sent key for use in his subsequent communication with Alice (where 390 is the random number known only to Bob).

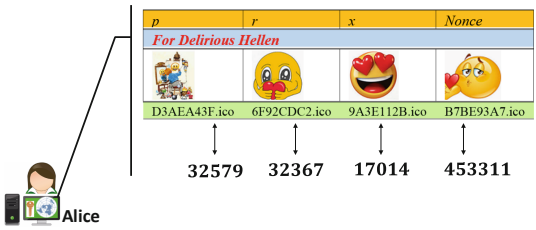


Fig. 17. On noticing Bob's status update, Alice visits his web page/site, and notes down the encrypted filenames of the four icons. She then decrypts the filenames using her Private Key and extracts the hidden parameters. After that Alice computes the stealthily sent key $K \leftarrow (17014^{20348} \bmod 32579) = 629$ at her end using those parameters for use in his subsequent communication with Bob (where 20348 is the random number known only to Alice).

With Allice and Bob sharing the same key (that is, 629 in this example) via computations performed respectively on their stand alone devices not connected with the internet, they may now decide on future course of action to communicate secretly via any encryption, and steganographic schemes over an insecure communicational channel explained as follows.

Information Embedding at Sender's End. The steps involved in information embedding in the context of preceding computations are:

(i) Compute HASH of the shared key (629) concatenated with the nonce (453310) using SHA-256 [60].
(ii) Encrypt the message using computed HASH (via any pre-agreed encryption algorithm such as AES [61]).
(iii) Translate the output into equivalent bits.
(iv) Reduce 629 (the shared key) modulo (No. of steganographic algorithms).
(v) Embed encrypted message bits using the corresponding steganographic algorithm and update the personal web page.

fridaya@uef.fi

Information Extraction at Receiving End. The steps involved in information extraction in the context of the above computations are:

(i) Compute HASH of the shared key (629) concatenated with a nonce (453310) using SHA-256.
(ii) Reduce 629 (the shared key) modulo (No. of steganographic algorithms), and extract hidden encrypted bits using that algorithm.
(iii) Concatenate the extracted bits into equivalent bytes.
(iv) Decrypt (via pre-agreed encryption algorithm such as AES) the gathered bytes using computed HASH as its key.
(v) The decrypted contents correspond to the hidden message.

6 Advantages and Limitations of the Proposed Protocol

Following are some of the pros and cons of the proposed stealth key exchange protocol:

6.1 Pros.

The proposed protocol, built on top of existing technologies, maintains the anonymity of communicators. It offers an all-in-one communication security solution which warrants confidentiality, integrity, authentication, and non-repudiation of data/information when at rest or in transit. It adlibs RSA, Diffie-Hellman key exchange protocols and such algorithms via stealthily preserving their associated meta-data (that is sent in clear form).

6.2 Cons.

The length of key-exchange parameters must not exceed the extent permissible for filenames by various operating systems (OS).

7 Discussion

The real drive of CAPTCHA is to guard the system and benefactor resources against automated attacks [62]. It implies from above that despite today's state-of-the-art technological infrastructures and security solutions a "connect" seems missing between these two. Recent research [63] has asserted that increased internet security focus has undermined the role that users may play in protecting their information, and have emphasized on the need of having a user-centric approach to privacy and security while devising technology-based solutions [64] considered human as the weakest link in cybersecurity. However [65] confronted it as unjustified to expect users to protect themselves from antagonists without first inducing security in design and development.

Security is an integrated whole and [65] has called for the combination of people, process, and technology for its attainment. Hence, to view security in isolation is unwarranted. Security is tied up with societal values [66] that is, what one perceives

fridaya@uef.fi

about it, which in turn is under the influence of enforced laws [67] and is a subjective aspect. Confidentiality, integrity, and availability is the objective approach to information security defined in the classical information security model called CIA triad [68]. Where confidentiality ensures that only the authorized have access to that information, integrity assures that the information remains untampered, while availability means that information is accessible as and when desired by users.

Foregone in view, an acceptable communication security solution would be the one that can uphold both the subjective and objective aspects of security besides being user-centric. Hence, moving along with cryptography alone as a security solution will not be acceptable nor does running without it.

It implies from above that it is not only desirous to circumvent a fully automated security solution through user-driven approach but also to let cryptography operate under cover of steganography for a seamless information/communication security solution.

Hence, the proposed hybrid solution for stealth key exchange and its subsequent facilitation for veiled data traversal between communicators promises it to meet the current legislative requirements besides keeping confidentiality and the integrity of data, and non-repudiation of its sender intact. It upholds the anonymity and privacy of its users and confers that for security to prevail, human (user) centric approach towards the design and implementation of an information security system is a must.

8 Research Contribution

A wide range of recent research centering on the Diffie-Hellman key exchange protocol and covering a variety of diversified fields such as steganography, cryptography, and networks exist. However, to the best of our familiarity and the literature reviewed, the proposed contribution is first of its kind work in instituting a stealth key exchange mechanism that practically has eliminated the physical traversal of key exchange parameters across the communication ends that is, between communicators. Besides being cost effective, the proposed methodology maintains the anonymity and privacy of the communicators through a hybrid (comprising of cryptography and steganography techniques) security solution.

9 Conclusion

This novel research exertion has taken the recent widely debated issue concerning legislative approval of the grant of access to security keys and cryptographic contents to law enforcing authorities by various Governments as a research problem (in its eternity) by envisaging an environment devoid of cryptography. In doing so, it has suggested a stealth key exchange mechanism - the reign component of any security system, by unifying cryptography and steganography. The solution thus proposed preserves the privacy and maintains the anonymity of the communicating parties without violating the imposed legislative bounds. This all-in-one user-centric communication security solution provides confidentiality, integrity, and non-repudiation to

fridaya@uef.fi

ongoing communication that does not require any additional infrastructure as it benefits from the existing technologies. Off-line computing on standalone devices disconnected from the internet, will further improvise on the security of the system.

Suggesting as a way forward for further exploration, this research appreciates user-centric, unified communication security solutions via blending cryptography and steganography to attain anonymity, privacy and stealth key exchange over an insecure channel.

References

1. Mohammadi, M: The Internet and the treasure house of information. In: The Iranian Reform Movement, pp. 215–239 (2019). Springer
2. Papacharissi, Z.: The virtual sphere: the internet as a public sphere. In: Praktiken der Überwachten, pp. 43–60. Springer, Wiesbaden (2019)
3. Aswale, P., Shukla, A., Bharati, P., Bharambe, S., Palve, S.: An overview of internet of things: architecture, protocols and challenges. In: Information and Communication Technology for Intelligent Systems, pp. 299–308. Springer, Singapore (2019)
4. Chang, C., Srirama, S.N., Buyya, R.: Internet of Things (IoT) and new computing paradigms. In: Buyya, R., Srirama, S.N., (eds.) Fog and Edge Computing: Principles and Paradigms, 1st edn. Wiley (2019)
5. Benjelloun, F.-Z., Lahcen, A.A.: Big data security: challenges, recommendations and solutions. In: Web Services: Concepts, Methodologies, Tools, and Applications, pp. 25–38. IGI Global (2019)
6. Ali, I., Sabir, S., Ullah, Z.: Internet of things security, device authentication and access control: a review. *arXiv preprint* arXiv:1901.07309 (2019)
7. Sciancalepore, S., Oligeri, G., Piro, G., Boggia, G., Di Pietro, R.: Exchange: securing IoT via channel anonymity. Comput. Commun. **134**, 14–29 (2019)
8. Yahya, A.: Steganography techniques. In: Steganography Techniques for Digital Images, pp. 9–42. Springer (2019)
9. Horsman, G.: A call for the prohibition of encryption; panacea or problem? *IEEE Security and Privacy Magazine* (2018)
10. Kuehn, K.: Surveillance and South Africa. The Polit. Econ. Commun. **6**(2), 94–100 (2019)
11. Kapoor, N., Narkowicz, K.: Unmaking citizens: passport removals, pre-emptive policing and the reimagining of colonial governmentalities. Ethn. and Racial Stud. **42**, 45–62 (2019)
12. Bachmann, G., Knecht, M., Wittel, A.: The social productivity of anonymity. Ephemera Theor. Polit. Organ. **17**, 241–258 (2017)
13. Lipschultz, J.: Free Expression in the Age of the Internet: Social and Legal Boundaries. Routledge, New York (2018)
14. Draper, N.A.: The Identity Trade: Selling Privacy and Reputation Online. New York University Press, New York (2019)
15. Chang, C.P.: Searching on Facebook through the lens of the concept of privacy. In: Social Networks and Surveillance for Society, pp. 129–148. Springer, Cham (2019)
16. Mishra, S., Sahoo, S., Mishra, B.K.: Addressing security issues and standards in Internet of things. In: Emerging Trends and Applications in Cognitive Computing, pp. 224–257. IGI Global (2019)
17. Attié, E., Meyer-Waarden, L.: The acceptance process of the Internet of Things: how to improve the acceptance of the IoT technology. In: Smart Marketing with the Internet of Things, pp. 21–45. IGI Global (2019)

fridaya@uef.fi

18. Lin, T.T.C., Kononova, A., Chiang, Y.-H.: Screen addiction and media multitasking among American and Taiwanese users. J. Comput. Inf. Syst., (2019). https://doi.org/10.1080/08874417.2018.1556133
19. Przegalinska, A.: Addiction and distraction. In Wearable Technologies in Organizations, pp. 25–31. Springer, Cham (2019)
20. Hedström, K., Sefyrin, J., Gustafsson, M., Wihlborg, E., Iréne, B.: A digital society for all? Meanings, practices and policies for digital diversity. In: Proceedings of the 52nd Hawaii International Conference on System Sciences (2019)
21. Kizza, J.M.: Ethical, privacy, and security issues in the online social network ecosystems. In: Ethical and Secure Computing. Undergraduate Topics in Computer Science, Springer, Cham (2019)
22. Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., et al.: RFC 2616. Hypertext Transfer Protocol–HTTP/1.1, vol. 2, p. 2.2 (1999)
23. Andersson, L., Madsen, T.: RFC 4026-Provider Provisioned Virtual Private Network (VPN) Terminology, March 2005
24. Appelbaum, J., Muffett, A.: The ". onion" special-use domain name (2015)
25. Pinsonneault, A., Heppel, N.: Anonymity in group support systems research: a new conceptualization, measure, and contingency framework. J. Manag. Inf. Syst. **14**, 89–108 (1997)
26. Wang, D., Cheng, H., He, D., Wang, P.: On the challenges in designing identity-based privacy-preserving authentication schemes for mobile devices. IEEE Syst. J. **12**, 916–925 (2018)
27. Wood, A., Nissim, K.: Is privacy privacy? (2018)
28. Rumbold, B., Wilson, J.: Privacy rights and public information. J. Polit. Philos. **27**, 3–25 (2019)
29. Cole, D., Fabbrini, F.: Bridging the transatlantic divide? The United States, the European Union, and the protection of privacy across borders. Int. J. Const. Law **14**, 220–237 (2016)
30. Hamidi, F., Poneres, K., Massey, A., Hurst, A.: Who should have access to my pointing data?: privacy tradeoffs of adaptive assistive technologies. In: Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 203–216 (2018)
31. Forgas, J.P., Baumeister Routledge, R.F.: The Social Psychology of Living Well, 1st edn. Published February 7 (2018). ISBN 9780815369240
32. Roth, G.B.: Probabilistic key rotation. Google Patents (2019)
33. Thapar, S.S., Sarangal, H.: A study of data threats and the role of cryptography algorithms. In: 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 819–824 (2018)
34. Yan, S.Y.: Offensive cryptography. In: Cybercryptography: Applicable Cryptography for Cyberspace Security. Springer, Cham (2019)
35. Pelton, J.N., Singh, I.B.: Protecting privacy from internet abuses in the smart city. In: Smart Cities of Today and Tomorrow. Copernicus, Cham (2019)
36. Cavelty, M.D., Kavanagh, C.: Cybersecurity and human rights. In: Research Handbook on Human Rights and Digital Technology. Edward Elgar Publishing (2019)
37. Geurden, M.: A New Future for Military Security Using Fully Homomorphic Encryption (2018)
38. Dierks, T., Rescorla, E.: RFC 5246. The transport layer security (TLS) protocol version, vol. 1 (2008)
39. Hoffman, P.: Cryptographic suites for IPsec, 2070–1721 (2005)
40. Ghosh, E., Debnath, D., Banik, B.G.: Blind RGB image Steganography using discrete cosine transformation. In: Emerging Technologies in Data Mining and Information Security, pp. 189–201. Springer, Singapore (2019)

fridaya@uef.fi

41. Douglas, M., Bailey, K., Leeney, M., Curran, K.: An overview of steganography techniques applied to the protection of biometric data. Multimedia Tools and Appl. **77**, 17333–17373 (2018)
42. Sarmah, D.K., Kulkarni, A.J.: Improved cohort Intelligence—a high capacity, swift and secure approach on JPEG image steganography. J. Inf. Secu. Appl. **45**, 90–106 (2019)
43. Kothari, A.M., Dwivedi, V., Thanki, R.M.: Introduction. In: Watermarking Techniques for Copyright Protection of Videos. Signals and Communication Technology. Springer, Cham (2019)
44. Kini, N.G., Kini, V.G.: A secured steganography algorithm for hiding an image in an image. In: Integrated Intelligent Computing, Communication and Security, pp. 539–546. Springer (2019)
45. Pandian, R., Xu, F., Sadrloshrafi, M., Gupta, R.: Online data transformation. Google Patents (2018)
46. Shashni, N., Yadav, M.: Cryptanalysis on digital image watermarking based on feature extraction and visual cryptography. In: Progress in Advanced Computing and Intelligent Engineering, pp. 425–435. Springer, Singapore (2019)
47. Bhat, S., Kapoor, V.: Secure and efficient data privacy, authentication and integrity schemes using hybrid cryptography. In: International Conference on Advanced Computing Networking and Informatics, pp. 279–285 (2019)
48. Hancock, P., Steele, S.: Facilitating provisioning of an out-of-band pseudonym over a secure communication channel. Google Patents (2019)
49. Diffie, W., Hellman, M.: New directions in cryptography. IEEE Trans. Inf. Theory **22**, 644–654 (1976)
50. Housley, R.: Internationalization Updates to RFC 5280, 2070–1721 (2018)
51. Harkins, D., Carrel, D.: RFC 2409. The Internet Key Exchange (IKE) (1998)
52. Felsch, D., Grothe, M., Schwenk, J., Czubak, A., Szymanek, M.: The dangers of key reuse: practical attacks on IPsec {IKE}. In: 27th {USENIX} Security Symposium ({USENIX} Security 18), pp. 567–583 (2018)
53. Farrell, S., Tschofenig, H.: Pervasive monitoring is an attack (2014)
54. Deibert, R.: Cyberspace under siege. J. Democracy **26**, 64–78 (2015)
55. Cooper, A., Tschofenig, H., Aboba, B., Peterson, J., Morris, J., Hansen, M. et al.: RFC 6973: Privacy considerations for Internet protocols. *IETF. Retrieved from tools. ietf. org/html/rfc6973* (2013)
56. Rescorla, E., Korver, B.: RFC 3552: Guidelines for writing RFC text on security considerations. Internet Society Req. for Comm. (2003)
57. Simmons, G.J.: The prisoners' problem and the subliminal channel. In: Advances in Cryptology, pp. 51–67 (1984)
58. Von Ahn, L., Blum, M., Hopper, N.J., Langford, J.: CAPTCHA: Using hard AI problems for security. In: International Conference on the Theory and Applications of Cryptographic Techniques, pp. 294–311 (2003)
59. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM **21**, 120–126 (1978)
60. Turner, S.: Using SHA2 Algorithms with Cryptographic Message Syntax (2010)
61. Daemen, J., Rijmen, V.: AES proposal: Rijndael (1999)
62. Choi, J., Oh, T., Aiken, W., Woo, S.S., Kim, H.: POSTER: I can't hear this because I am human: a novel design of audio CAPTCHA system. In: Proceedings of the 2018 on Asia Conference on Computer and Communications Security, pp. 833–835 (2018)
63. Chong, I., Xiong, A., Proctor, R.W.: Human factors in the privacy and security of the Internet of things. In: Ergonomics in Design (2018). https://doi.org/10.1177/1064804617750321

fridaya@uef.fi

64. Sasse, M.A., Brostoff, S., Weirich, D.: Transforming the 'weakest link'—a human/computer interaction approach to usable and effective security. BT Technol. J. **19**, 122–131 (2001)
65. Andress, A.: Surviving Security: How to Integrate People, Process, and Technology. Auerbach Publications, Boca Raton, 2003
66. Brauch, H.G.: From a security towards a survival dilemma. In: Globalization and environmental challenges, pp. 537–552. Springer, Berlin (2008)
67. Nye Jr., J.S., Lynn-Jones, S.M.: International security studies: a report of a conference on the state of the field. Int. Secur. **12**, 5–27 (1988)
68. Perrin, C.: The CIA triad. Dostopno na (2008). http://www.techrepublic.com/blog/security/the-cia-triad/488

fridaya@uef.fi

Evaluation of the Multifactor Authentication Technique for Mobile Applications

Tianyi Zhang, Le Yang, and Yan Wu(✉)

Bowling Green State University, Bowling Green, USA
{tianyiz,yangle,yanwu}@bgsu.edu

Abstract. Nowadays, mobile smartphones are popular devices among general population. Compared to the traditional mobile phone, smartphone is closer to a personal computer with a good number of mobile applications installed. However, when a user browses certain mobile application, his or her account is at risks if static password is the only element required for authentication. To overcome this problem, multi-factor authentication is widely required for mobile applications. In this paper, we conduct a survey for a group of mobile application users and analyze the pros and cons of each authentication method by case studies. We focus on six main factors to evaluate each authentication method. Our study found that traditional password authentication, biometrics authentication methods such as fingerprint, face, and voice scored relatively high on Convenience, Time, Security and Accuracy. More important, these four aspects are exactly the most essential factors to a mobile application's quality that relate to safety issue and user experience as well. Our evaluation results show that biometric authentication methods are currently used most often by mobile applications and well accepted by its users. Overall, we evaluated five cases with advantages and disadvantages. We found fingerprint, voice authentication played outstanding as biometric authentication methods. We also come up with a new authentication proposal for mobile application design and for the future research that in view of mobile devices are more and more functional nowadays, we recommend mobile applications to use biometric authentication in two steps authentication methods.

Keywords: Mobile security · Multi-factor authentication

1 Introduction

Along with the fast evolving of mobile devices, especially mobile phones, daily jobs/entertainment such as browsing websites, video recording, photographing, video games, online bank, chat can be done on our mobile phones through applications. More and more important personal data are stored in or transferred by mobile devices, and once the information is compromised, it may cause negative results such as personal information leaking, property loss even worse. Hence, the task to protect user's personal information is critical to everyone.

Authentication is the way to distinguish whether the one who is trying to gain access is the authorized user himself/herself or not. Further, Multi-factor authentication (MFA) is an advanced method of authentication, in which user will be granted access

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 696–707, 2019.
https://doi.org/10.1007/978-3-030-22868-2_49

fridaya@uef.fi

only when they provide multiple disjoint pieces of information all correctly. The required information usually contains the following aspects:

- Knowledge: what people know about, such as password, personal identification number (PIN)
- Possession: what people have, such as credit card;
- Inheritance: something they are, biometrics for example.

Once the identity of the human or machine is validated, access is granted [1]. In another word, if at least one of the above components is failed, the service will be blocked to that person. Two-factor authentication, also called 2FA, is a method to confirm the identity of the user by verifying combination of two different components of the user's identity, and it is an example of multi-factor authentication. From daily life, a good example is the withdrawal operation from the ATM: only the right combination of a bank card with a PIN number can grant the agent the authority to finish transactions.

The present authentication techniques are well-developed. And some of the common used, reliable techniques are Traditional Static Password, Short Message Service (SMS) Dynamic Password, Digital Signature, Biometrics Identification, Automatic and Speech Recognition (ASR), etc. However, in the meanwhile, attack techniques are developing quickly as well. To decrease the risks of compromised account, an analysis of the mechanism of each authentication technique is necessary.

2 Related Work

Aloul et al. [1] developed an authentication model using two factors authentication technique for mobile devices. They introduced two models. The first one is Connection-Less Authentication System which accepts multiple elements such as local time, username, and pin. Then the combined long string is hashed and shrunk to obtain a short One Time Password (OTP) for use. The second method is called "SMS-Based Authentication System" which is a safe mechanism for the failure of the first step and leads to password rejection, etc. The mechanism is to send information which is unique to the user through SMS to the server. Then the server sends back a one-time-password which is only available in a short period.

Static password is not safe enough. Sabzevar et al. [2] combined graphical password and multifactor authentication to create secure authentication process to prevent attacks even within un-trusted environment. They designed a web-based authentication process based on .Net. Three different types of graphical images are designed such as "(1) Random images with random clickable areas, (2) User picked image with random clickable areas, (3) A grid of clickable squares."

Sun et al. [3] introduced TouchIn, a two-factor authentication system on multi-touch mobile devices. TouchIn works by letting a user draw on the touchscreen with one or multiple fingers to unlock his mobile device, and the user is authenticated based on the geometric properties of his drawn curves as well as his behavioral and physiological characteristics. Liou et al. [4] proposed a way of multi-factor authentication by using Radio Frequency Identification (RFID). The advantages of RFID are inexpensive

fridaya@uef.fi

and small enough to fit anywhere. The main concept of RFID is to retrieve the information stored in the tokens using radio signals. Many multi-factor authentications did not take environment factors, user roles, and type of devices into consideration. Nag et al. [5] focused on and proposed an adaptive selection of authentication modalities by considering specific situation and environment.

Jain et al. [6] explained the mechanisms of biometric authentication such as finger prints, iris, face and so forth. They showed the steps of how to scan original biometric image of certain user and then convert them into digital format to store. Then they propose approaches to secure biometric information by using biometric feature transformation and biometric cryptosystem. Schiavone et al. [7] designed a biometric system called multimodal continuous biometric authentication mechanism. It has a protocol contains an initial phase and a maintenance phase.

Human beings have limitation to remember things especially different passwords. Dhamija et al. [8] implemented a system that relies on recognition-based authentication system, which authenticates a user through one's ability to recognize previously seen pictures. Trewin et al. [9] evaluated different aspects includes: user-time, effort, error by using three biometrics authentication modalities: voice, face, and gesture.

Raza et al. [10] suggest that one should choose different authentication combination according to different scenarios. Several authentication methods can be combined into one scheme which to make the authentication process simple and secure.

3 Methodology

We selected a few current multi-factor authentication techniques to compare and evaluate. Also, we will come up with some innovated idea which can remedy the shortage of those authentication method at some degree.

Commonly used multi-factor authentication method is two step authentications in most applications. As other practitioners, we are fully aware of the nature of customers' preference to choose authentication mechanism that is simple as well as secure. Thus, we consider user experience as a critical factor to measure. Also, we propose our authentication mechanisms, we will conduct survey which contains different authentication methods with a scale of 0 to 10 in several assessment areas: Cost, Convenience, Time, Accuracy, Intuition of security, and mental attitude of security. According to the survey result, we will design different authentication mechanism with reasons why we choose to design it in this way, and what hardware or software requirement should be met.

For evaluation, we selected several authentication modalities from the most widely used password to biometrics.

- Fingerprint: A way to capture the ridges of all or part of a finger.
- Face recognition: An application of computer for automatically identifying or verifying a person from a digital image or a video frame from a video source.
- CAPTCHA: It stands for Completely Automated Public Turing test to tell Computers and Humans Apart, it usually displays a string of twisted numbers or characters for people to retype. Computer will not be able to recognize these characters.

fridaya@uef.fi

- SMS: It stands for Short Message Services, usually users will receive a text message on phone.
- Keystroke: Keystroke is a method to verify the identity of an individual by their typing rhythm which can cope with trained typists as well as the amateur two-finger typist.

Before the experiment, we studied the way how two-factor authentications work and the different scenarios of evaluation of them. Then we evaluated a few selected mobile applications with combinations of the modalities. By comparing these different modalities, we are interested in finding answers to these three research questions:

RQ1: What parameters should we use to compare the advantages and disadvantages of these modalities?

RQ2: What authentication combination can achieve a high security level while is not complicated to use?

RQ3: What are the advantages of each modality? What are the disadvantages or the limitations of each modality?

And to answer RQ1, we evaluated these modalities from several aspects listed below:

1. Cost: how much is the cost of the authentication modalities including money and computation?
2. Convenience: How simple is the process of each authentication modality for users?
3. Time: how long it takes to finish the authentication?
4. Accuracy: to what degree, each authentication modality is accurate?
5. Security: to what degree, each authentication modality is secure?
6. Application: how widely adopted each modality is?

4 Analysis

Due to the popularity and wide adoption of biometric authentication based on their uniqueness, like tree leaves and our human body as well, it can be used as a type of physical password. Biometric authentication is safe; however, it is expensive since it usually requires additional hardware to scan biometric data.

4.1 Two Fundamental Scenarios

Our assumption of multi-factor authentication for mobile applications is that Users would choose them because of the convenience and security they provide. And based on this assumption, we design two factors authentication mechanism in two basic scenarios.

Scenario one: if a user log in his/her account with the IP address and mobile devices which he or she used most often (which is so called trusted device).

fridaya@uef.fi

In this case, a user can directly skip static password process and jump to second factor authentication such as login by fingerprint and so forth. The user will login successfully by providing its second factor password directly. It will be more convenient, faster, and safer than typing in static password. If the requirements of this scenario do not meet, then, scenario two is necessary.

> Scenario two: if a user log in his/her account using a public device or infrequently used IP address. It will be treated as an anomaly.

In this case, first factor authentication (static password) is required, then the user will be allowed to enter the second factor authentication process. When there is an anomaly such as unknown device login or a long-distance login from a user's IP address, user should get notification with details, and user's account is blocked for a short time for him/her to get informed and take corresponding actions. Even when user successfully logged into the account, a message will be sent to the user for safety concerns through email or SMS on trusted devices.

Based on the two scenarios above, we choose to evaluate two-factor authentication mechanism. Each mechanism we evaluate will use static password and the second factor authentication will be a practical way to enhance user experience while safety is focused on. So, we selected three different applications on mobile to compare user login authentications, and they are: BGSU student login, we-chat, PNC app, along with two recommended mechanism: face recognition and key stroke. The reason we choose these five instances is because of these five applications are used most often by our survey subjects. Also, the authentication method of these applications aligns with the goal of our research.

4.2 Authentication Analysis

Banking App using fingerprint authentication. Smart phones come with fingerprint recognition component make fingerprint authentication possible. Apple product achieved remarkable performance on fingerprint recognition. It is precise and fast which increases user experience and make it popular. Apple Inc. uses fingerprint for "device unlock" and "apple pay" and third party mobile app can use the fingerprint to perform their authentication. We surprisingly found that a banking application, PNC, allows user to use fingerprint authentication. A user does not need to scan his or her fingerprint dedicatedly, PNC simply ask a user to authorize PNC mobile application to use touch ID for login.

Firstly, this is convenient that once a user authorizes the PNC mobile banking to use a user's fingerprint stored locally in mobile device; s/he will be able to access his/her account simply by scanning his/her fingerprint when password is automatically filled in. Secondly, it is a secure way to prevent hacker that even a hacker knows a user's password; the fingerprint authentication will shut the door upon the hacker. Thirdly, the improvement of reading speed and accuracy of fingerprint reader are greatly enhanced along with the improvement of hardware and algorithm.

The drawback of fingerprint authentication is that the investment for a mobile device with a fingerprint reader is relatively more expensive than other mobile devices.

fridaya@uef.fi

However, the price will eventually be dropped. Fingerprint authentication for mobile application has a bright future since many merits significantly enhance the user experience and security as well. Also, a user should only store his/her fingerprints in the mobile device and no one should know his/her passcode of the device, even close one should not. Otherwise, fingerprint authentication can be misused by malicious action from someone close.

Social media using voice authentication. Besides of fingerprint authentication, voice authentication is another good way, however, it is not as popular as fingerprint authentication. When a user uses voice authentication for the first time, We-Chat will ask a user to read a sequence of number for several times in any language whichever a user would like to use. We refer the sequence of number a user read as the voice password. Once the system successfully initialized a user's voice authentication, it will allow a user to log in by using voice authentication.

There are concerns that a voice record can break into the system. However, we have conduct several small experiments by using voice record to break into the system and we failed. We conclude that it is because of the distortion when nature voice stored into a digital device that a machine can distinguish the difference between the nature voice and recorded voice while human cannot. Also, even someone can record with low distortion and s/he tries to record a user's voice; it must be the moment a user read his voice password. People come with scenarios that what if a user has a cold and his voice cannot be recognized by the system. In this case, we would recommend a user lo in his/her account by selecting other methods which we realized that multiple choice for second step authentication is necessary. We will come up with a solution in the next section.

The advantage of voice authentication is that the installment fee is cheap since the hardware requirement is low; it is relatively fast if a user read his voice password accurately; it is precise except certain condition such as intentionally or accidently changed user voice; and it is proved to be secure by our experiments.

There are some drawbacks of voice authentication which are quiet environment is needed and it needs nature language process (NLP). However, we hold the opinion that voice authentication works fine without NLP.

School website BGSU.edu using authenticator. Because of account hacking occurred recently (2017) in Bowling Green State University (BGSU), IT department has installed a two-way authentication process for university account login. It offers three choices for users to do authentication: (1) DUO authenticator, (2) Phone Call, (3) Passcode.

First option is Duo push. Duo push is a third-party authenticator called Duo authenticator which can sent user a push for approval or a passcode once it is connected to certain organization. It supports both IOS and Android devices; however, it does not support other devices such as windows phone. Similar application exists in the market which supports both IOS and Android device is Google Authenticator. Google uses this authenticator for its "2 step verification" currently. Second option "Call Me" is similar as SMS authentication; BGSU set a robot to call a user to prompt he or she tap certain number for admission authorization. Third choice is to enter a passcode either received as SMS or received from Duo authenticator. However, we do not recommend this one

fridaya@uef.fi

is because that SMS or PASSCODE is not safe since it can be forwarded to others. We assume a user is a reasonable being but we understand that others have many ways to get a user's passcode or SMS message.

Face authentication has a sign to be popular. Because of the development of smart phones, which enables to read human biometric such as fingerprint and voice, more secure authentication mechanism can be applied. With biometric authentication, our design of multifactor authentication can be sundry.

Face modality is used by some personal PC manufactory to do PC account log in authentication in the past. Nowadays, PC abandoned face authentication because of its inaccuracy. However, some mobile devices are starting to use face authentication for the purpose to unlock devices recently. Face recognition units [5] are expensive than fingerprint recognition units because face biometric is more complicated than fingerprint as more complicated model and algorithm must be applied. It is a good sign to see that accuracy of face authentication has been greatly improved. However, security issues such as how to prevent portrait and identical twin siblings from successful authentication could be a problem to research. Face recognition can be a decent authentication method. We believe that more and more mobile devices will use face authentication as well as fingerprint. Third party should build face authentication friendly application to seize the initiative.

Key Stroke authentication. We have searched very hard for any mobile application which come with key stroke authentication, while we failed. Most of the Key Stroke technique appears in academic research. We conclude that although password plus key Stroke technique is a good proposal, it is still difficult to apply it to mobile application market.

4.3 An Authentication Model Proposal to Developers

According to the survey data from section six and the case study from section five, we come up a proposal for mobile application to install identification authentication. We believe that two-factor authentication is enough in the worst case to prevent a hacker from our user property. There are many mobile applications use email as log in username or uses email for password change. When a user's email is compromised, the account safety is in danger. Companies may argue that it is not their fault. However, it is important to keep users' account safe no matter users' email is compromised or not. By implementing two-factor authentication, even a users' email account is compromised and corresponding static password authentication process is compromised.

Most cases are a hacker, at a certain area far away from a user, hacks into a user's account far away from the user. It means that the scenario occurred in the movie of stealing fingerprint may not occur. So, second-authentication process is enough to prevent a hacker to hack into a user's account.

So, we propose two-factor authentication for mobile application based on scenario one plus scenario two we mentioned before. We also propose to offer multiple choices for second authentication process. It is designed for special conditions that if a user accidently has no way but to use other authentication methods. By default, it is recommended to use fingerprint as second authentication method since it has a perfect

fridaya@uef.fi

survey score and a good case study analysis result as well. However, if a user does not have a mobile device with fingerprint reader or a user would like to use voice authentication because s/he always wear gloves at work which it is inconvenient to authenticate with fingerprint, other methods such as voice authentication or associated authenticator is recommended. It also be possible that when a user has a cold and his/her voice cannot be recognized, s/he can switch to fingerprint authentication or associated authenticator for authentication.

Although we believe two-factor authentication is secure; we still imagine that a super hacker finds out a way to break second factor authenticator. Even though a hacker can break our second authentication and login far away from a user's location, an anomaly will be found and there will be a corresponding action. In this case, we recommend that a mobile application should designed based on our scenario two that when there is an anomaly such as unknown device login or a long-distance login from a user's IP address. A user should get notification with details. Also, a user's account can be blocked for a short time such as five minutes for a user to get informed and take corresponding measures. This is special measurement which only occurs when there is an anomaly.

5 Results

There were total 20 participants in our survey. The participants are from China and the United States, including professionals and students. According to the survey results, we created a figure for each aspect to evaluate and compare them. The results based on a scale from 0 to 10, 0 means the worst and 10 means the best among all modalities. We took the average and rounded the results to the nearest half. The results for each aspect is shown below.

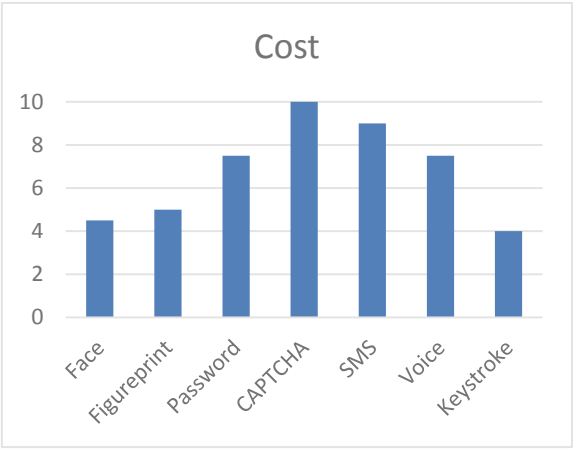


Fig. 1. Cost and convenience factors for seven modalities [5]

fridaya@uef.fi

From Fig. 1, which is a sample figure we take from [5], we could see that the best three modalities perform better in terms of cost are CAPTCHA, SMS, and password respectively. The higher the score is, the better performance a modality has on certain factor, which means CAPTCHA is the most cost friendly technique. Apparently, Face recognition, keystroke and fingerprint need more cost, because their computation are heavier than the other four modalities.

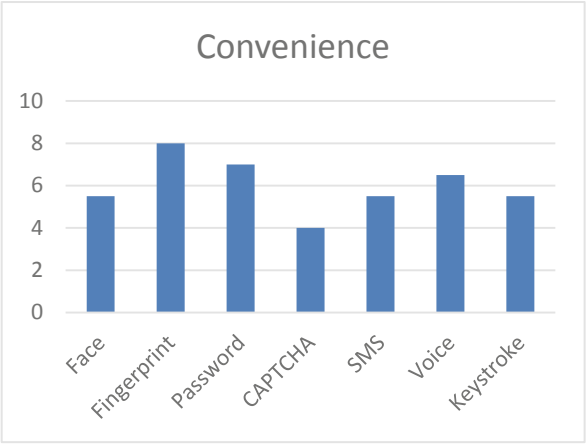


Fig. 2. Convenience factor for seven modalities

From Fig. 2, people rated fingerprint, password and voice as more convenience to use. Among them fingerprint gets the highest rate. CAPTCHA, face recognition, and SMS seems are more inconvenience.

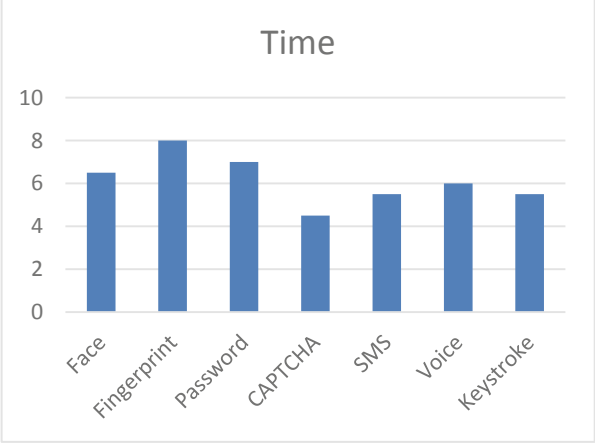


Fig. 3. Time factor for seven modalities

fridaya@uef.fi

From Fig. 3, we could tell that every authentication takes time. None of them get higher than 8. Fingerprint gets the highest rate because all the participants are using iPhone. Users are get used to the fingerprint modality. CAPTCHA, SMS and keystroke seems more time consuming.

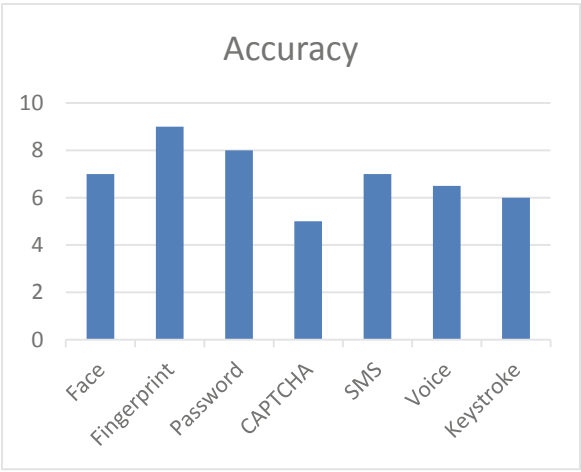


Fig. 4. Accuracy factor for seven modalities

From Fig. 4, we can see that fingerprint and password get the highest rating score. People still think fingerprint is the best among all modalities. CAPTCHA is the lowest among all of them.

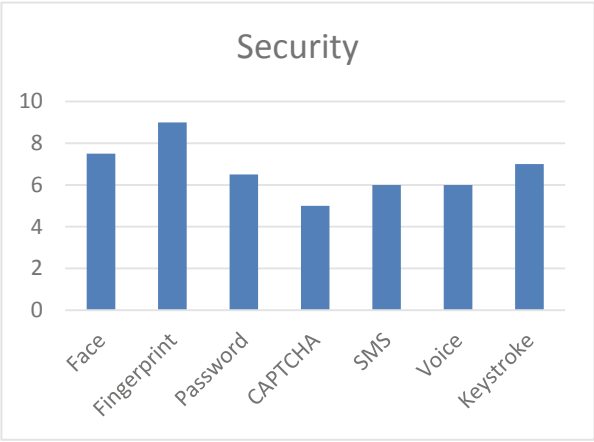


Fig. 5. Security factor for seven modalities

fridaya@uef.fi

According to Fig. 5, we can see that the top three modalities, fingerprint, face recognition and keystroke are all biometrics. Biometrics got higher rate because they are high-computational biology modalities and it is hard to duplicate the result.

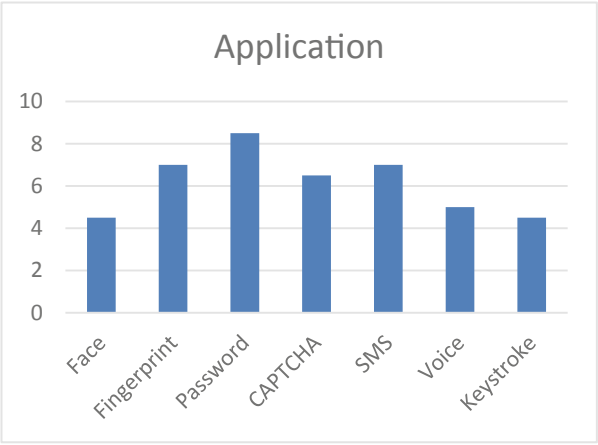


Fig. 6. Application factor for seven modalities

Figure 6 shows that among the applications, password is still the most widely used way to authenticate a user. Fingerprint is rising because iPhones make this modality applicable to its users. Keystroke are new and not applied to a wide range of applications.

Face recognition, fingerprint, voice recognition and keystroke share the biometric advantages of authentication. It is unique to a person and it is hard to duplicate; it is easy and safe to use; it is more reliable and accurate for one's identity. One disadvantage of biometrics would be it always needs the presentation of the owner; and the cost of the computation and device are higher.

6 Conclusion

Multi Factor authentication especially two steps authentication is widely used nowadays, and two typical combinations are password combined with SMS and password combined with biometric authentication. Because biometric data is unique, as long as the scanning equipment is accurate and reliable, mechanism such as password combined with biometric is relatively more secure than password combined with SMS.

In conclusion, we compared the seven modalities of authentication from users' perspective by referencing from [5]. Our results indicate that the biometrics are more secure and accurate but the cost of them are relatively high. Also, we provided several case studies about applications on two-factor authentication. The results provided guidance for mobile apps' authentication design and for future research. A new authentication model is proposed based on the data from the survey and the case studies

fridaya@uef.fi

as well. It improves the overall user experience and security level for mobile application. We also recommend future mobile application development use our proposed authentication model as reference.

References

1. Aloul, F., Zahidi, S., El-Hajj, W.: Multi factor authentication using mobile phones. Int. J. Math. Comput. Sci. **4**(2), 65–80 (2009)
2. Sabzevar, A.P., Stavrou, A.: Universal multi-factor authentication using graphical passwords. In: IEEE International Conference on Signal Image Technology and Internet Based Systems, 2008. IEEE (2008)
3. Sun, J., et al.: Touchin: sightless two-factor authentication on multi-touch mobile devices. In: 2014 IEEE Conference on Communications and Network Security (CNS). IEEE (2014)
4. Liou, J.-C., et al.: A sophisticated RFID application on multi-factor authentication. In: 2011 Eighth International Conference on Information Technology: New Generations (ITNG). IEEE (2011)
5. Nag, A.K., Dasgupta, D., Deb, K.: An adaptive approach for active multi-factor authentication. In: 9th Annual Symposium on Information Assurance (ASIA14) (2014)
6. Jain, A.K., Nandakumar, K.: Biometric authentication: system security and user privacy. IEEE Comput. **45**(11), 87–92 (2012)
7. Schiavone, E., Ceccarelli, A., Bondavalli, A., Carvalho, A.M.: Usability assessment in a multi-biometric continuous authentication system. In: 2016 Seventh Latin-American Symposium on Dependable Computing (LADC), pp. 43–50. IEEE (2016)
8. Dhamija, R., Perrig, A.: Deja Vu-A user study: using images for authentication. In: USENIX Security Symposium, vol. 9, pp. 4–4 (2000)
9. Trewin, S., et al.: Biometric authentication on a mobile device: a study of user effort, error and task disruption. In: Proceedings of the 28th Annual Computer Security Applications Conference. ACM (2012)
10. Raza, M., Iqbal, M., Sharif, M., Haider, W.: A survey of password attacks and comparative analysis on methods for secure authentication. World Appl. Sci. J. **19**(4), 439–444 (2012)

fridaya@uef.fi

A Doubly-Masked-Ballot Multi-questions Voting Scheme with Distributed Trust

Robert Szabo[1(✉)] and Andras Telcs[2,3,4]

[1] Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Budapest, Hungary
robert.szabo@tmit.bme.hu
[2] Department of Computational Sciences, Wigner Research Centre for Physics of the Hungarian Academy of Sciences, Budapest, Hungary
telcs.andras@wigner.mta.hu
[3] Department of Computer Science and Information Theory, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary
[4] Department of Quantitative Methods, Faculty of Business and Economics, University of Pannonia, Veszprem, Hungary

Abstract. We propose an electronic voting scheme for multiple questions voting based on masked ballots, homomorph counting, publicly verifiable secret sharing based tallying and public communication channels. In our scheme all on-line authority functions (authentication, candidate slate check, counting and tallying) are distributed among the active voters. Off-line authorities only prepare and distribute election secrets. We analyze our scheme for different adversary attacks and show that the maximum that the strongest adversary can achieve is a random attack.

Keywords: Voting · Masked ballot · Coercion free · Verifiable

1 Introduction

Electronic voting schemes have been widely researched in the last decade. The developments mostly centered around cryptographic mechanisms and physical constrains to pertain privacy, receipt-freeness, accuracy, fairness, integrity, verifiability and incoercibility. The solutions presented so far, however, could not gain wide acceptance as the technical dimensions are somewhat orthogonal to human acceptance and comprehension of the voting systems.

On the technical side electronic voting schemes fall into two broad categories: anonymous tallying and homomorph encryption based schemes. Anonymous e-voting schemes use an anonymous channel or a mix network to hide connections between votes and voters. On the other hand, homomorphic encryption based schemes exploit the fact that some operations applied to the ciphertexts will yield to the same result after decryption as applying some (possibly other) operations to the plaintexts. Regardless of the approaches taken, voting schemes must

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 708–726, 2019.
https://doi.org/10.1007/978-3-030-22868-2_50

fridaya@uef.fi

expose certain information publicly to ensure universal verifiability. Public information related to the casted ballot shall enable voters to individually verify that their casted ballot corresponds to their intended vote and that the authorities conducted the election correctly. However, a computationally correct[1] authority (e.g., a Tally) can leak information, which threatens the privacy of the voters and other security requirement of the system.

Another important but not well studied property of electronic voting systems is how many times a given user secret can be used repeatedly to cast votes or ballots. For example, a typical general assembly requires subsequent ballot casting from the participants. Furthermore, the ballots can contain multiple questions. How such multiple questions/ballots scenario can be efficiently handled without reinitialization of users' secrets is unknown to our best.

An electronic voting scheme, which could be used in voting scenarios with flexible agendas without subsequent reinitialization of users' secrets, would proliferate electronic voting to our everyday use.

In our proposal we aim to address the multiple questions/ballots problem while providing ballot secrecy and dispute freeness [1]. Our protocol innovatively builds on recent advancements of homomorph encryption based schemes [2], secret sharing protocols [3] and publicly verifiable multi-party computation threshold schemes [4–7].

We propose a distributed electronic voting scheme, particularly suitable for mobile environment, based on masked balloting by Wen and Buckland [2]. In our doubly-masked-ballot multi-questions voting scheme we rely on two off-line authorities, a registrar and an independent key set provider (IKSP), whose roles are very limited and terminate before the start of the election. The roles of traditional on-line authorities (talliers and counters) are randomly distributed among the voters. Our contributions can be summarized as follows:

- We created a distributed and trustful tallying and counting mechanism complementing the masked ballot election by Wen and Buckland [2].
- We solved the problem of publicly verifiable candidate slate mapping with the introduction of a second mask and the use of a slightly modified version of least significant bit (LSB) extraction procedure [7].
- We eliminated all dedicated on-line voting authorities from our voting system. Tallying and counting are done by the participants themselves. The registrar authority's role is limited to the stage before the start of the election.
- We introduced an independent and off-line authority, the Independent Key Set Producer (IKSP), who produces, seals, signs, and transfers election secrets to the Registrar. The IKSP can mass produce election secrets for various election sizes in quantity, without knowing when and where secret sets will be used.
- During the election, we assume no physical constraints on the used communication channel. However, we assume a one-way untappable channel for the IKSP → Registrar → Voter communication before the start of the election.

[1] Input - output wise follows the rules of the protocol.

fridaya@uef.fi

- In our protocol we achieve ballot secrecy and dispute freeness.
- We provided mechanisms to vote for multiple questions/ballots without reinitializing the users' secrets.
- We ensure that any deviations from the protocol can be recognized by the participants through universal and individual verifiability.

The rest of the paper is organized as follows. In Sect. 2 we discuss related works. In Sect. 3 we define our doubly-masked-ballot voting scheme, which is followed by its analysis in Sect. 4. Finally, we draw conclusions and discuss future plans in Sect. 5.

2 Related Work

Electronic voting systems must enable the adding up of votes while retaining the secrecy of individual votes and partial results and the privacy of voters. These two requirements are in contradiction to each other [5] and rise particular difficulties in mobile environment (c.f. [8]). There exist two major approaches to tackle this fundamental problem: (*i*) published plain ballots without voters' identity but trivial tallying and (*ii*) encrypted (secret) ballots. The former approach can be realized by anonymous channels or mixing networks (mix-nets), where the ballots are shuffled before publishing [6,9,10]. The secret ballot approach demands special encryption methods, which allow tallying in the encrypted forms (c.f. [11]). These encryption functions are typically homomorphic [12,13].

2.1 Homomorphic and Threshold Cryptosystems

In homomorphic encryption, an algebraic operation performed on the ciphertexts is equivalent to another algebraic operation performed on the plaintexts. This is a very attractive property for electronic voting systems, as tallying can be done on the encrypted ballots. Such homomorph encryptions are RSA [14], ElGamal [12], Paillier [13] and their generalizations. The ElGamal encryption is multiplicatively homomorph but can be transformed into an additively homomorph one. Unfortunately, the additively homomorph ElGamal encryption requires calculation of the discrete logarithms over finite fields to decrypt, for which there exists no efficient classical algorithm. This, however, for moderate sized elections is still feasible [15]. On the other hand, the Paillier encryption is additively homomorphic, hence is widely used in electronic voting schemes, e.g., [2,16].

Threshold cryptosystems (secret sharing) allows to share a secret among several participants such that only a certain group of them can recover it. Many of the secret sharing schemes utilize Shamir's algorithm [17], which is based on the Lagrange approximation. In standard secret sharing, the secret is restored by the participants. In multi-party computation threshold schemes, participants are not only able to restore the secret but are also able to perform computations (e.g., decryption) without revealing any information but the result [4,5]. It is also known from literature that such a multi-party computation can be verified by the involved players, e.g., [16].

fridaya@uef.fi

However, Peng recently revealed in [18] that publicly verifiable secret sharing (PVSS) schemes are vulnerable to simple attacks. Fortunately, these attacks are related to dishonest dealers, who distribute the secret shares, but not to the publicly verifiable property of the distributed calculations. In our proposed doubly-masked-ballot multi-questions voting scheme, we will use a PVSS scheme with a trusted dealer (a trusted authority) together with publicly verifiable (n, t) threshold least significant bit (LSB) extraction method [7]. Such schemes allow universal verification [18] and dispute freeness [1] among the participants.

2.2 Least Significant Bit Extraction (LSB)

Another useful property of ElGamal and Paillier homomorph encryptions is that they allow least significant bit extraction in an (n, k) threshold scheme, which is a multiparty, publicly verifiable secret sharing computation [7]. The distributed LSB protocol can extract the encrypted least significant bit x^0 of an encrypted plaintext x. For (P, S) public and secure keys, where n participants hold the $s_1, s_2, \ldots, s_n$ k threshold secret shares of S, the protocol input to LSB is the encrypted $P(x)$, a jointly generated encrypted random mask $P(y)$ and the mask's encrypted least significant bit $P(y^0)$. Neither the n secret keepers nor others can gain any knowledge of x, x^0, y, y^0 during the distributed calculation of $P(x^0)$, while all the steps of the process is publicly verifiable [18].

2.3 Receipt-Freeness and Coercion-Resistance

The concept of *receipt-freeness*, when voters cannot prove to a third party how they voted even if they wish to do so, first appeared in Benaloh and Tuinstra's work [19]. They assumed a "voting booth" with physical restriction: no communication to the outside and no observers while in the booth. Okamoto in [20] relaxed the physical booth assumptions to one-way *anonymous* secret communication. Hirt and Sako in [21] showed that the original concept by Benaloh et al. was not receipt-free and that achieving anonymity and secret communication of Okamoto is extremely difficult, hence unpractical. Therefore, Hirt and Sako introduced a new receipt-free voting scheme, where voters select their vote from a randomly ordered encrypted candidate slate list, which are tallied in encrypted forms thanks to the applied homomorph encryption. In their scheme voters shared their secret vote among multiple authorities but validity had to be verified with multiple (n) tests, yielding to $1 - 2^{-n}$ valid vote confidence.

In [6] Juels and Jakobsson added a more powerful adversary model and introduced a *coercion resistant* voting scheme. They defined coercion resistance by not only providing receipt-freeness, but also protecting against randomization, forced-abstention and simulation attacks. In randomization attack, the adversary coerces the voter to submit a randomly composed balloting material, which nullify the choice of the voter with large probability. In forced abstention attack the coercer demands the voter to refrain from voting. In simulation attacks, the adversary coerces the secrets of the voter after registration but prior to

fridaya@uef.fi

voting, and if these secrets can be verified by the adversary then he can simulate the voters at will. In order to overcome these threats, Juels and Jakobsson's scheme required the use of an *anonymous channel*. Actually, they argued that *an anonymous channel is a minimal requirement for any coercion-resistant schemes* conforming to their definition.

Consecutive contributions in seeking for coercion-resistance, however, redefined the adversary model in several ways. In [22] Cetinkaya et al. defined uncoercibility in a way that a voter cannot be coerced to cast her vote in a particular way. Moran and Naor in [23] described an interactive adversary model in which the adversary can interact with but not oversee (over-the-shoulder) the voter any time during the protocol execution. Their scheme provides fake inputs to the voter to answer the coercer's questions. According to their interactive adversary model, the coercer can instruct or command the coerced voter during protocol execution. They also acknowledged that to protect against forced-abstention attack a scheme must use an anonymous channel, hence they removed this type of requirement from their definition of coercion resistance. *In our scheme, we will follow the definition of coercion-resistance by Moran and Naor.*

Meng in a recent work [24] critically reviewed the evolution of receipt-free voting approaches and proposed a new receipt-free and coercion resistant protocol without physical assumptions in [25] together with Li and Qin. Their approach to coercion resistance is that the voter can resubmit her favorite ballot after submitting a forced vote (random or the coerced). However, they did not discuss why the adversary cannot learn about the resubmitted vote without physical assumptions or how the candidate slate verification remains hidden after the resubmissions.

2.4 Physical Assumptions

Sako and Kilian [26] showed that a one-way untappable channel is enough for receipt freeness. Moran and Naor in [23] added that without a trusted third party device (e.g., a smart card) the minimal assumption is the one way untappable channel. In our approach we will follow [2,20,21] to use a one way untappable channel during the distribution of voting secrets.

2.5 Human Considerations

While dealing with security problems, *human considerations* are in most of the proposals neglected. Moran and Naor in [27], however, put considerable emphasis on human considerations. They argue that most of the existing protocols require complex computation on the part of the voter, hence demanding *trust* in the voting units (VU) casting their ballots. In their proposal, they returned to the traditional voting-booth settings with paper ballots due to human acceptance. Paper ballots with sealed envelopes are good things to convince the voters that their secrets are really private and destroyed afterward.

We share Moran and Naor's view on the importance of these human trust factors, but we are committed to electronic ballots. We also believe that it is

fridaya@uef.fi

difficult to maintain trust in the authorities during election and as Moran and Naor also acknowledged it is impractical to increase the number of authorities as it would result in quite complex human interface.

Even in the case when there are several authorities establishing *trust* is a serious human concern as the voter must know at least one (or even some) honest authorities to be sure of no fraud. For simplicity and convenience, it is generally assumed that there is a single *trusted registrar authority* in the scheme [6]. This model is similar to traditional paper based balloting schemes where trust in the registrar is also unavoidable, but the registrar's role is very limited and controlled by the community. *We will follow this model and put trust into the registrar authority.*

2.6 Masked-Ballot

Wen and Buckland in [2] introduced a masked ballot voting scheme on the argument that a secret ballot is fundamental to protect the freedom of voters. In their masked ballot scheme the only physical assumption was the untappable channel during the registration. Their scheme inherits from the random identity concept of Juels and Jakobsson [6] and the homomorph encryption used in [28]. In the masked ballot scheme, voters receive their random masks during the registration (similar to the random identity). These masks are randomly assigned to the voters during the registration and the encrypted masks are published to the bulletin board. A missing piece in Wen and Buckland's work is how to verify that a vote is in the candidate slate set. Actually, the verification, tallying and counting are deferred in their scheme's description to some standard authorities. As Wen and Buckland left the important tallying and counting method undefined, their scheme is unfinished. (We shall note that Wen and Buckland at the same time defined a minimum disclosure counting scheme for preferential election in [29]. This, however, cannot be directly applied to plurality elections).

To overcome the above concerns, we propose an extended masked ballot scheme in which we give a verifiable distributed tallying and counting protocol. With our modification we also avoid publishing of encrypted but unmasked individual votes. By tallying only the least significant bits we can avoid the verification if a vote is in the candidate slate set. Furthermore, we propose an extensions how multiple ballots/questions can be handled with an initial secret set while retaining receipt-freeness and different levels of coercion resistance. We also relaxed Wen and Buckland's strong assumptions on atomic operation and voter non-observability for coercion resistance.

3 The Doubly-Masked-Ballot Scheme

Intuitively, we created a doubly-masked-ballot voting scheme based on additionally homomorph encryption as follows:

1. IKSP creates secret sets: voter mask, a secondary mask, voting keys;

fridaya@uef.fi

2. During registration voters receive a secret mask to be used to hide their votes;
3. Unencrypted but masked votes are sent to the BB during ballot casting;
4. Masked votes are encrypted for tallying at the BB;
5. Encrypted masked votes at the BB are mapped to $\{0, 1\}$ using LSB extraction utilizing a per voter secondary secret mask;
6. The LSB parts of the submitted votes are added up in their encrypted form;
7. The final tally is decrypted by a publicly verifiable secret sharing scheme.

A more detailed description of the scheme is as follows.

3.1 The Model

Security Model: Players of our doubly-masked-ballot scheme can be categorized into the following groups:

- **Independent Key Set Producer (IKSP) Authority:** IKSP is an off-line player, who independently generates and seals secret sets for different election (participants) and ballot/questions (e.g., secrets in sealed envelopes in a sealed package). IKSP's public key is known to each participants. Our scheme places significant trust into the IKSP. This is because we try to eliminate all possible threats from the on-line voting process for the trade-off to place processing to the IKSP in the preparation phase.
- **Registrar Authority:** The registrar decides who is eligible to vote; supervises the distribution of voting secrets and finally supervises private erasure of the secrets. In this means, the registrar is a passive authority; her only actions are the verification of the voter's eligibility, the supervision of the voter's random choice of initialization secrets and the private destruction of secrets. The registrar has no information on the secrets and cannot alter them. Furthermore, voters themselves or (their) delegated trustees can oversee the registrar's activity and prevent unauthorized use of secret sets. We assume that all the unused credentials are destroyed by the registrar under the supervision of voters or their delegates. The Registrar is an off-line player.
- **Voter (V):** Eligible voter $V_i \in \mathcal{V}$ for $i = 1, \ldots, N$, who is allowed to cast his/her ballots in the election. Voters are the only on-line players in the process.
- **Trusties (secret keepers):** Trustee roles are randomly distributed among the Voters during the registration by secret share assignments.
- **Observers:** Any voter or any third party can play the observers' role. They can oversee the action of the Registrar Authority and verify all the computations published publicly.

Physical Assumptions:

- **Forced private erasure:** Forced private erasure [27] requires that voters physically destroy some of the secrets they received during registration in order to ensure receipt freeness.

fridaya@uef.fi

- **One-way untappable channel:** A one-way untappable channel ensures completely private communication. In our scheme we assume one-way untappable channel for the IKSP → Registrar → Voters communication. An example could be that the IKSP first creates and prints a credential set for N voters, than seals and posts it in a sealed package of sealed envelopes to the Registrar. Voters can verify the integrity of the sealed package before the registration. The Registrar opens and empties all credentials in the package into a ball. Each voter pulls a random sealed envelope out of the ball, opens it, enters the secrets but one into her voting unit (e.g., reads a quick response code) and destroys the paper. If the seal is untampered, then the voter trusts in the received secretes. The secret to be memorized corresponds to the voter's secret mask (e.g., a number or a word). In the case of paper based registration, the untappable channel is ensured by the Registrar and the independent observers.
- **Bulletin Board (BB):** Bulletin board is used as a public broadcast channel with memory, where anyone can read and append but cannot delete the messages. We assume a passive BB, that is, no authentication for senders to the BB is done.
- **Voting Unit (VU):** The voting unit is a simple processing unit, which takes an input of the *masked vote* and sends the signed masked vote to the BB.

Multiple Questions: We assume that within an election there can be multiple rounds of ballots. Each ballot round has its start and termination within the election. Furthermore, each ballot can contain multiple questions. Answers to questions in the same ballot are casted as one operation with the ballot, though privacy and security must be separately guaranteed per questions. *We assume that secrets are conveyed to the voters once at the very beginning of the election to be used for all ballots and all questions.* The particular issue in this scenario is that each answer should have all the security independently of their actual order.

For the remainder of the description, let us denote by Q the maximum number of questions to be answered during an election.

Adversary Model: We consider both non-interactive and interactive adversary models. In the non-interactive model the adversaries must perform all corruptions before the protocol execution [5]. This also relates to a scenario when some of the voters are dishonest. In the interactive adversary model adversaries can interact and corrupt voters during the protocol execution. Here, we distinguish between voters who expose their full internal states (e.g., internal state of the voting unit also known as voting unit attack); those who are visually observed by an adversary (over the shoulder attack); and those who are instructed but not observed (blind attack). Further, we assume that adversaries can eavesdrop any communication apart from the one-way untappable channel used for the registration but *cannot corrupt voters up to a threshold*.

fridaya@uef.fi

Note that our assumptions allow insecure voting units and voting environment and that the adversary observes the voters physically or electronically.

3.2 Requirements

- Eligibility: In the final tally only eligible and authorized voters' cast are counted.
- Privacy: Votes are kept in secret, i.e., no one but the voter can determine the value of her vote.
- Receipt-freeness: Voters cannot prove to a third party how they voted even if they wish to do so [19].
- Fairness: Equal vote share for each participant and no partial knowledge on the result before the final tally is calculated and published.
- Robustness: No reasonable sized coalition of voters can disrupt the voting.
- Accuracy: No vote can be altered, eliminated or added multiple times to the tally and only valid votes are counted.
- Universal verifiability: Anyone can independently verify that all the actions of involved parties (voters) comply with the protocol.
- Individual verifiability: Each individual voter can verify that her vote was handled and counted correctly.
- Dispute-freeness: If anyone can verify that the protocol was run correctly and that each voter acted according to the rules of the protocol [7].
- k-threshold ballot secrecy: k-threshold ballot secrecy requires that knowledge about the partial tally of the ballots of any set of voters is only computable by the coalition of k voters. $k = N - 1$ results in perfect ballot secrecy, when all the remaining voters must act together in order to disclose a ballot. This property captures strong voter privacy as understood in real world elections [1].
- Operation without trusted authorities during casting, tallying and counting: Dishonest participant can be identified and their contribution can be eliminated from the protocol.
- Operation over any type of communication channel during protocol execution[2].

3.3 Cryptographic Building Blocks

Homeomorph Encryption: We use the Paillier additionally homomorph encryption with public and private keys (P, S), such that $P(x|\rho) = [\![x|\rho]\!]$ and $P(x_1|\rho_1) \times P(x_2|\rho_2) = P(x_1 + x_2|\rho_1\rho_2)$, where x, x_1 and x_2 are plaintext messages and ρ, ρ_1 and ρ_2 are random nonces used during the encryption. For simplicity, we omit the random nonce from the notation, i.e., $P(x|\rho) = [\![x]\!]$, unless it bears with special importance in the protocol.

[2] During the registration phase, prior to the election, the use of a one-way untappable channel to read secrets by the voter is admissible.

fridaya@uef.fi

Homomorphic Threshold Encryption: A multi-party computable, non-interactive, additionally homomorph, publicly verifiable threshold encryption [18]. Let $s_1, s_2, \ldots, s_n$ be the secret shares of S creating a multi-party (n, k) threshold secret sharing, where any k out of n shares can decrypt a chipertext. Note, that not any $k-1$ participant can decrypt the tally.

Least Significant Bit Extraction (LSB): The LSB procedure maps $[\![x]\!]$ to $[\![x^0]\!]$ where x^0 is the least significant bit of x. The mapping is an (n, t) threshold procedure in which none of the cooperating parties gain knowledge of x or x^0.

3.4 The Protocol

Let the candidate slate set {YES, NO} be mapped to $\{1, 0\}$. Let x^z denote the z^{th} bit of x, $z = 0$ identifying the least significant bit.

Step 1: Preparation. The Independent Key Set Producers (IKSP) authority prepare election and voters' secrets for an election size N and questions Q. These secrets are:

- A balloting public and secret key pair (P^b, S^b) is selected together with (n, k) threshold secret shares $s_1^b, s_2^b, \ldots, s_n^b, n \leq N$.
- A random primarily mask $m_i \in 2^Q$ for voter V_i is selected.
- A random secondary mask $M_i \in 2^Q$ for voter V_i is selected. Let $R_{ij} = 2r + M_i^j$, where r is a random positive integer.
- For each voter $V_i, i = 1, \ldots, N$ encryptions for the BB are calculated: $\{[\![R_{ij} - m_i^j|\rho]\!], [\![M_i^j|\rho]\!]\}, 0 \leq j \leq Q$, where ρ's are random nonces only known by the IKSP. Later on these nonces will protect the encrypted unmasked votes as we discuss it in Sect. 4.3.
- An unique public and private key pair for voter V_i: (P_i^v, S_i^v). Let $\text{Sign}(x, S_i^v)$ be the signature for message x produced by voter V_i.

Let the IKSP also calculate the signed hash $H = h(P^b, \{P_i^v, [\![R_{ij} - m_i^j|\rho]\!], [\![M_i^j|\rho]\!], 0 \leq i \leq N, 0 \leq j \leq Q\})$ for all public information of the election for verification purposes.

Step 2: Registration. During the registration each voter receives a *random* $\{S_i^v, s_i, m_i\}$ set of secrets, where s_i is either zero or one share of the S^b secret. In the latter case the given voter is randomly selected as a secret keeper. Note that the voter's identity is also randomly determined by S_i^v.

Step 3: Initialization. The bulletin board is initialized with

- global parameters: $\{P^b, H\}$;
- voter parameters: $i : \{P_i^v, \{[\![R_{ij} - m_i^j|\rho]\!], [\![M_i^j|\rho]\!], 0 \leq j \leq Q\}\}$;

fridaya@uef.fi

Step 4: Voting. Let v_{ij} be the favorite answer of voter V_i for question j at a ballot and $d_{ij} = v_{ij} + m_i^j$ her masked vote, where m_i^j is the j^{th} bit of m_i. She casts her vote to the ballot as $\{i, d_{ij}, \text{Sign}\,(d_{ij}, S_i^v)\}$. Note that one ballot may contain more than one question, hence a casted ballot contains as many individually masked votes as many questions. For simplicity, in our following discussion we will assume that each ballot contains only one question, hence each casted vote corresponds to one casted ballot.

Step 5: Termination of ballot casting. The ballot casting is terminated at a certain time, which is made public. Let Ω_j be the set of voters who answered question j in their ballots. All tallying and counting is done for the Ω_j set.

Step 6: Candidate slate mapping. The questions on the casted ballot must be mapped to the candidate slate set. This is achieved by first encrypting the masked vote and than applying an LSB extraction to the parity bit corrected encrypted masked vote. More formally, for question $j \in \mathcal{N}, 0 \leq j < Q$:

- Let any of the voters calculate and publish to the BB the encrypted masked votes $l : [\![d_{ij}]\!], \forall l \in \Omega_j$.
- Let any of the voters calculate and publish to the BB the $[\![d_{ij}]\!] \times [\![R_{ij} - m_i^j | \rho]\!] = [\![v_{ij} + M_i^j + 2r|\rho]\!] = [\![y_{ij}|\rho]\!]$.
- Let the choorum of the (n, k) threshold scheme (trustees) decrypt y_{ij} once for question j for every voter V_i. Additionally, let y_{ij}^0 be the least significant bit of y_{ij}.
- Encrypt $[\![y_{ij}^0]\!]$ to the BB. The BB at this time contains $y_{ij}^0, [\![y_{ij}^0]\!], [\![M_i^j|\rho]\!]$.
- Calculate $[\![y_{ij}^0 \oplus M_i^j]\!]$ using the fact that $y_{ij}^0 \oplus M_i^j = y_{ij}^0 + M_i^j - 2y_{ij}^0 M_i^j$, i.e., $[\![y_{ij}^0 + M_i^j - 2y_{ij}^0 M_i^j]\!] = [\![y_{ij}^0]\!] \times [\![M_i^j|\rho]\!] \times [\![M_i^j|\rho]\!]^{-2y_{ij}^0}$.
  However $LSB((v_{ij} + 2r + M_i^j) \oplus M_i^j) = LSB((v_{ij} + M_i^j) \oplus M_i^j) = v_{ij}^0$, i.e., $[\![y_{ij}^0 \oplus M_i^j]\!] = [\![LSB(v_{ij})]\!] = [\![v_{ij}^0|\rho]\!] = [\![v_{ij}^0]\!]$.

Note that if $v_{lj} \in \{0, 1\}, \forall l \in \Omega_j$ then $LSB(v_{ij}) = v_{ij}$, otherwise $LSB(v_{ij}) \to \{0, 1\}$, which is a valid vote again. Also note that all the above operations are public except the (n, k) threshold decryption of $[\![y_{ij}]\!]$, which is also publicly verifiable [7]. Also note, that v_{ij}^0 is protected in the encrypted form by a random nonce ρ known only by the IKSP.

Step 7: Counting. Let any voter V_i do the counting calculation for the encrypted least significant bit of the casts: $[\![c_{ij}]\!] = \prod_{l \in \Omega} [\![LSB(v_{lj})]\!] = [\![\sum_{l \in \Omega} v_{lj}^0]\!]$, and publish the value to the BB together with a signature: $\{i, [\![c_{ij}]\!], Sign([\![c_{ij}]\!], S_i^v)\}$.

Step 8: Tallying The results can be decrypted from any $c_{lj}, l \in \Omega$ by the (n, k) secret keepers for (P^b, S^b), i.e., $[\![c_{lj}]\!] \to c_{lj} = c_j$.

fridaya@uef.fi

3.5 Protocol in the Bulletin Board

The protocol steps as seen at the bulletin board are shown in Fig. 1.

Submitted values	Relevant public values
Step 3: Initialization	
	P^b, H
i:	$P_i^v, \{[\![R_{ij} - m_i^j \vert \rho]\!], [\![M_i^j \vert \rho]\!] : 0 \leq j \leq Q\}$
Step 4: Voting (ballot casting): j^{th} **question**	
i: $\{d_{ij}, \text{Sign}(d_{ij}, S_i^v))$	P_i^v
Step 5: Termination of ballot casting	
$\Omega = \{l$: who submitted a vote $\}$	
Step 6: Candidate slate mapping $i \in \Omega$	
i: $d_{ij} \rightarrow [\![d_{ij}]\!]$	
$[\![d_{ij}]\!] \rightarrow [\![v_{ij} + m_i^j]\!][\![R_{ij} - m_i^j]\!] \rightarrow [\![y_{ij}]\!] = [\![v_{ij} + 2k + M_i^j]\!]$	$[\![R_{ij} - m_i^j \vert \rho]\!]$
$[\![y_{ij}]\!] \rightarrow y_{ij} \rightarrow y_{ij}^0 \rightarrow [\![y_{ij}^0]\!]$	
$\{[\![y_{ij}^0]\!], [\![M_i^j \vert \rho]\!]\} \rightarrow [\![v_{ij}^0 \vert \rho]\!] = [\![v_{ij}^0]\!] = [\![LSB(v_{ij})]\!]$	$[\![M_i^j \vert \rho]\!]$
Step 7: Counting $i \in \Omega$	
i: $[\![v_{ij}^0]\!] \rightarrow [\![\sum_{l \in \Omega} v_{lj}^0]\!] = [\![c_{ij}]\!]$	
Step 8: Tallying	
$[\![c_{ij}]\!] \rightarrow c_{ij} = \sum_{l \in \Omega} v_{lj}^0$	

Fig. 1. Protocol as seen in the Bulletin Board (Step 1 – 2 are not visible in the Bulletin Board)

4 Analysis

4.1 Verifiability

Independent Key Set Producer (IKSP): It is important to note that the IKSP is an off-line authority in the election regarding the creation of the secret sets.

Depending on the cost and scenario, election organizers may obtain several set of secret materials, out of which one is randomly selected for the actual use. Since the balloting secrets are randomly assigned to the voters (e.g., pull from a ball), the IKSP is not in the position to bias any predefined voter or group. However, the IKSP is in the possession of all the secrets of all the elections, hence must be trusted. Our goal here is to decouple the IKSP from the actual election as much as possible, therefore we limited the ISKP's role to the off-line preparation of election secrets.

Public Parameters: Assuming trust in the Independent Key Set Provider (IKSP) authority and that the IKSP's public key is know by the participants, then anyone can verify that the public election values published at the bulletin board match to the IKSP signed hash value (H) of the election setup. Therefore the encrypted masks and the voting keys cannot be altered by an adversary. This ensures universal verifiability of the public parameters.

fridaya@uef.fi

Voters' Secrets: The IKSP seals the voters' secrets $\{S_i^v, s_i, m_i\}$. If the seals are untampered, voters shall trust that their secrets are private and unaltered. Voters register S_i^v and s_i to their voting unit and memorize m_i (as a number or a word). All of these secrets are force erased at the end of the registration.

Authentication: Voters sign their casted ballot by S_i^v, whose public key pair P_i^v is published at the BB. Only registered and authenticated voters can produce this signature or an adversary who coerced this secret from the voter. However, anyone can verify that only the holder of the S_i^v secret key can cast a ballot for the registered user V_i. This ensures universal verifiability of the voters as long as the Registrar is not compromised.

Voting (Ballot Casting): This is a simple process regarding the VU. The VU takes the input of the masked vote $d_{ij} = v_{ij} + m_i^j$ for the j^{th} question, signs it with the voter's private key and puts it into the ballot. When the ballot round is terminated the ballot with the plaintext masked votes is sent to the BB. The masked vote registered to the BB can be verified to the masked vote entered into he VU by the voter.

Candidate Slate Mapping: Candidate slate mapping process is an adaptation of the LSB protocol described in [7] by the threshold secret keepers. Schoenmakers et al. proved that all steps of their protocol is publicly verifiable.

Counting: Everyone can calculate and publish the sum of the encrypted votes. Since the calculation is based on simple multiplications, disagreements can be backtracked to voters.

Tallying: The decryption of the counting is based on a distributed, (n, k) threshold, publicly verifiable secret sharing scheme [18]. It is important, however, that the quorum must only decrypt the result of the secret counting and nothing else.

4.2 Adversary Attacks

In our analysis we assume that the *voting is not necessarily an atomic process* in contrast to the work by Wen and Buckland [2]. We use the following adversary models:

- *Fairness attack*: when the adversary tries to increase his voting share.
- *Vote buying attack*: when the adversary tries to bias the election towards one of the parties by offering rewards.
- *Abstention attack*: when the adversary forces the voter not to cast her vote.
- *Random attack*: when the adversary's goal is to have a randomly uniform vote distribute among the candidates.

fridaya@uef.fi

The strength of the adversary attacks are categorized as follows:

- *Blind attack*: when the adversary cannot eavesdrop any communication between the voting unit and the voter.
- *Over the shoulder attack*: when the adversary eavesdrops all the communication between the voting unit and the voter.
- *Interactive attack*: when at any point in the protocol execution of single ballot casting the adversary can coerce the voter for information and command for the next actions.
- *Non-interactive attack*: if the single question casting is atomic for the adversary.
- *Voting unit attack*: when the adversary knows all the internal states of the VU and has total control of the operation of it.

For the j^{th} question and voter V_i, for simplicity, let us denote by v the intended vote, by m the genuine mask and by $d = v + m$ the masked vote. Let v' and m' denote the coerced values together with $d' = v' + m'$ the coerced masked vote. Let $d^* = v^* + m^*$ be the forced input to the VU.

Fairness Attack: Consider an adversary who wishes to increase his share in the counting process. When casting his favorite ballot he would change his v, hence d . However, the candidate slate mapping will only take into account the least significant bit corrected by the secret mask of such a vote, hence the share cannot be increased over 1. This applies to any round of the questions.

Vote Buying Attack (a.k.a. Receipt-Freeness): Let's consider an election between Bob and Alice, where the adversary wants to bias the election towards Alice. In this case the adversary would publish that he will pay for any ballot receipts for Alice. In our case the voter can present the adversary her $\{S^v, s\}$ and a mask m', which together with $v^* = v_{\text{Alice}}$ would yield the casted ballot $v + m = v_{\text{Alice}} + m'$. However, the adversary remains uncertain about genuineness of the mask m', hence the casted vote. Therefore, any voters can produce arbitrary receipts for an adversary, but cannot prove the receipt's validity, hence make such an attack pointless.

The vote buying attack and the non-interactive blind attack are the same from the point of view of the adversary's capabilities. Therefore same train of thoughts apply to here as well.

Non-interactive, over the Shoulder Attack: Since the voter is required to input the $d = v + m$ masked vote to the voting unit, without knowing the genuine mask the adversary cannot gain any information on the cast. As a consequence, the outcome of this scenario is identical to the previous vote buying attack and is independent of the round of the questions.

fridaya@uef.fi

Interactive, over the Shoulder Attack: In this scenario the adversary oversees all communication between the voting unit and the voter. This means that the adversary can learn the voter's d^* input. Since the adversary can coerce the voter's secret mask m', he can verify if the parity bit of $d' = v' + m'$ is equal to $d^* = v^* + m^*$, where v^* is the commanded vote. As long as the adversary cannot be sure of the genuineness of the coerced mask he cannot effectively control the result of the voter's cast. This is because, neither the voter nor the coercer are sure about the other's actions, therefore none of them can enforce her/his will. To maximize the opponent's entropy both shall randomize their choices for 1 and 0 with a $1/2 - 1/2$ probability. This is when the voter passes $m' = 0$ or 1 with $1/2 - 1/2$ probability to the coercer and the coercer commands to put the input $d^* = 0$ or 1 again with $1/2 - 1/2$ probability. As it is well-known any of them gets into a worse position if deviates from this strategy. That means that the coercer cannot achieve better than random attack if a random mask is given to him. Therefore, the only effective coercion resistant strategy of the voter is to give a random mask m' to the adversary.

Interactive Blind Attack: The adversary is in the same position as in the interactive, over-the shoulder scenario, since the masked vote is registered in plaintext to the BB. Therefore the same arguments apply here too.

Voting Unit Attack: Since the VU does not make any processing of the masked vote that is published to the BB in plaintext, it cannot leak any information other than the signing key of the voter. If the adversary modifies the VU to alter the masked vote, then the voter can detect and complain about it by verifying the published masked votes. If also coerced, then a coercion resistant strategy is to enter random masked votes to the VU, which if altered will only yield to a random attack.

Forced Abstention Attack: Unfortunately, only anonymous voting schemes can protect against forced abstention attack, hence our scheme is vulnerable to this type of attack.

4.3 Requirements Revisited

We revisit the remainder of requirements that were not discussed in the previous analysis sections.

Eligibility: Registration and verifiable private signature ensures that only eligible voters can submit votes and their signature is publicly verifiable.

Privacy: First of all, at the voter the vote is protected by the secret mask the voter received during the registration. It is important that the voter inputs the

fridaya@uef.fi

masked vote to the voting unit *in one atomic operation*. The masked plaintext vote is sent unencrypted to the Bulletin Board, where it is encrypted for homomorph tallying with a public nonce. At this phase the vote is still protected by – and only by – the secret mask. In the next phase the user's secret mask is replaced by the user's secondary mask only known by the IKSP. The encryption nonce applied to the secondary mask will protect the encrypted vote during the rest of the tallying process. This procedure keeps votes in secret first by the user's secret mask and later by the secret encryption nonces applied to the secondary mask by the IKSP. Therefore individual votes are never revealed, hence kept secret to ensure privacy. However, abstention to voting can be detected and may be possible to trace back to the voter due to the use of public communication channel during the election.

Fairness: Only a single parity bit for each voter is counted in the final tally, therefore no one can increase her share. Thanks to the (n, k) secret sharing scheme no one can access partial result without the contribution of $k - 1$ other dishonest trustee.

Robustness: Robustness is ensured by the (n, k) secret sharing scheme. With different parameter settings a trade-off between privacy and robustness can be fine tuned. In the most extreme case when all the participants are trustees as well and all of them are needed for the decryption than we can realize perfect ballot secrecy for the price that any single voter can sabotage the election process. Note that the saboteur's identity can be revealed thanks to the universal verifiability, which decreases the chance of such attacks. It should be also noted that as a last resort additional shared secret sets can be used with $k < n - 1$ necessary active participants. Such sets can be in advance prepared and distributed. If a trustee deviates from the protocol all parties are able to recognize it and can initiate a new round of operation excluding the trustee from the procedure either using the original secret shares or a new one with smaller number of secret keepers.

Accuracy: The masked votes are sent in signed plaintext to the BB. In order to alter the vote, the secret key of the voter must be disclosed. If an adversary coerce the voter for his secret key, she can follow the different coercion resistant strategies discussed in Sect. 4.2. As long as the voter submits her cast, she can verify in plain text if the accurate value is registered to the BB. Since all the remaining steps of the protocol are publicly verifiable, no one can alter, eliminate or multiply the casts registered to the BB.

Universal Verifiability: Operations are limited in number during the protocol execution and all steps of the protocol are publicly verifiable.

Individual Verifiability: Its is trivial to match the masked published votes on the BB with the input to the VU. The rest of the protocol steps are universally verifiable.

fridaya@uef.fi

Dispute Freeness: Given that all operations of the protocol are publicly verifiable, anyone can check if others comply to the protocol. There is no point to challenge an eligible/honest protocol step.

***k*-Threshold Ballot Secrecy:** Ensured by the (n, k) threshold secret sharing of the ballot secret key.

Operation Without Trusted Authorities During Casting, Tallying and Counting: Exactly capture the period when the authorities are not involved, the secrets are kept by the members of the (n, k) secret sharing scheme.

Operation over Any Type of Communication Channel During The Protocol Execution: Since all the protocol steps are published to the public bulletin board, there is no need to use any special communication channel. It is the mask that hides the actual value of the casted vote.

4.4 Complexity

The price to be paid in a distributed scheme compared to having several trusted authorities is computational complexity.

For a voter with trustee key, for N voters, Q questions and (n, k) threshold cryptosystems and LSB extraction method, the complexity of a voting unit with trustee role is $O(NkQ)$. However, the number of active trustees are limited from above by $N \geq n \geq k$, hence the overall complexity of the protocol for trustees is $O(N^2Q)$.

For a non trustee voter it takes $O(Q)$ steps to comply with the protocol without any verification activity. If a voter verifies the calculations of the trustees he/she needs to perform the same steps as a trustee: $O(N^2Q)$.

The overall complexity of the protocol for all the voters is $O(N^3Q)$.

5 Conclusions

We have introduced a homomorph encryption based electronic voting scheme, which operates without trusted authorities after the registration phase. In our scheme, the voters themselves perform the counting and tallying functions in a completely distributed fashion utilizing homomorph encryption and publicly verifiable secret sharing. Our scheme ensures voter's privacy through masking the cast, candidate slate mapping by a modified least significant bit extraction procedure utilizing the second secret mask. We analyzed different adversary models and presented coercion resistant strategies for the possible adversary attacks of varying strength. We concluded that our scheme is receipt free and the maximum, that the strongest adversary can achieve, is random attack. Besides, for reasonable adversary attacks we showed how coercion resistant strategies allow the voter to cast her favorite ballots.

In our future plans we wish to investigate how the roles of the off-line authorities (IKSP and Registrar) could be distributed among the voters.

fridaya@uef.fi

References

1. Kiayias , A., Yung, M.: Self-tallying elections and perfect ballot secrecy. In: Public Key Cryptography, 5th International Workshop on Practice and Theory in Public Key Cryptosystems, PKC 2002. LNCS, vol. 2274, pp. 141–158, Springer (2002)
2. Wen, R., Buckland, R., Ryan, P., Schoenmakers, B.: Masked ballot voting for receipt-free online elections. In: Ryan, P., Schoenmakers, B. (eds.) E-Voting and Identity. LNCS, vol. 5767, pp. 18–36. Springer, Heidelberg (2009)
3. Fouque, P.-A., Poupard, G., Stern, J.: Sharing decryption in the context of voting or lotteries. In: Financial Cryptography 2000, pp. 90–104 (2000)
4. Cramer, R., Damgard, I., Nielsen, J.B.: Multiparty computation from threshold homomorphic encryption. In: Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques: Advances in Cryptology, EUROCRYPT 2001, pp. 280–299. Springer (2001)
5. Hirt, M.: Multi-party computation: efficient protocols, general adversaries, and voting. Ph.D. thesis, ETH ZURICH (2001)
6. Juels, A., Jakobsson, M.: Coercion-resistant electronic elections. In: Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society (WPES), pp. 61–70 (2005)
7. Schoenmakers, B., Tuyls, P.: Efficient binary conversion for paillier encrypted values. In: Advances in Cryptology-EUROCRYPT 2006, pp. 522–537 (2006)
8. Li, C.-T., Hwang, M.-S., Liu, C.-Y.: An electronic voting protocol with deniable authentication for mobile ad hoc networks. Comput. Commun. **31**(10), 2534–2540 (2008)
9. Jakobsson, M., Juels, A., Rivest, R.: Making mix nets robust for electronic voting by randomized partial checking. In: Proceedings of the 11th USENIX Security Symposium, pp. 339–353 (2002)
10. Juels, A., Catalano, D., Jakobsson, M.: Coercion-resistant electronic elections. In: Lecture Notes in Computer Science, Chap. 2, vol. 6000, pp. 37–63. Springer, Heidelberg (2010)
11. Morales-Rocha, V., Soriano, M., Puiggal, J.: New voter verification scheme using pre-encrypted ballots. Comput. Commun. **32**(710), 1219–1227 (2009)
12. Gamal, T.E.: A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE Trans. Inform. Theory **31**, 469–472 (1985)
13. Paillier, P.: Public-key cryptosystems based on composite degree residuosity classes. In: Advances in Cryptology - EUROCRYPT 1999, International Conference on the Theory and Application of Cryptographic Techniques, pp. 223–238 (1999)
14. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM **21**, 120–126 (1978)
15. Jurik, M.J.: Extensions to the paillier cryptosystem with applications to cryptological protocols. Ph.D. thesis, BRICS, Department of Computer Science, University of Aarhus, August 2003
16. Ruiz, A., Villar, J.L.: Publicly verfiable secret sharing from paillier's cryptosystem. In: Western European Workshop on Research in Cryptology, WEWoRC 2005. LNI, vol. 74, pp. 98–108 (2005)
17. Shamir, A.: How to share a secret. Commun. ACM **22**, 612–613 (1979)
18. Peng, K.: Investigation and survey of secret sharing in verifiable distributed systems. In: 12th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), Gwangju, pp. 342–347. IEEE, October 2011

fridaya@uef.fi

19. Benaloh, J., Tuinstra, D.: Receipt-free secret-ballot elections. In: Proceedings of the 26th ACM Symposium on Theory of Computing, Montreal, PQ, ACM Symposium on Theory of Computing, pp. 544–553. ACM, New York, May 1994
20. Okamoto, T.: Receipt-free electronic voting schemes for large scale elections. In: Lecture Notes in Computer Science, Chap. 3, vol. 1361, pp. 25–35. Springer (1998)
21. Hirt, M., Sako, K.: Efficient receipt-free voting based on homomorphic encryption. In: Proceedings of the 19th International Conference on Theory and Application of Cryptographic Techniques, EUROCRYPT 2000, pp. 539–556. Springer, Heidelberg (2000)
22. Cetinkaya, O., Doganaksoy, A.: A practical privacy preserving e-voting protocol using dynamic ballots. In: 2nd National Cryptology Symposium (2006)
23. Moran, T., Naor, M.: Receipt-free universally-verifiable voting with everlasting privacy. In: Advances in Cryptology-CRYPTO 2006, pp. 37–392 (2006)
24. Meng, B.: A critical review of receip-freeness and coercion-resistance. Inform. Technol. J. **8**, 934–964 (2009). http://www.scialert.net/abstract/?doi=itj.2009.934.964
25. Meng, B., Li, Z., Qin, J.: A receipt-free coercion-resistant remote internet voting protocol without physical assumptions through deniable encryption and trapdoor commitment scheme. J. Softw. **5**, 942–949 (2010)
26. Sako, K., Kilian, J.: Receipt-free mix-type voting scheme. In: Lecture Notes in Computer Science, pp. 393–403. Springer, Heidelberg (1995)
27. Moran, T., Naor, M.: Split-ballot voting: everlasting privacy with distributed trust. ACM Trans. Inform. Syst. Secur. (TISSEC) **13**, 1–43 (2010)
28. Cramer, R., Gennaro, R., Schoenmakers, B.: A secure and optimally efficient multi-authority election scheme. In: EUROCRYPT 1997. LNCS, vol. 1233, pp. 103–118. Springer (1997)
29. Wen, R., Buckland, R.: Minimum disclosure counting for the alternative vote. In: Proceedings of the 2nd International Conference on E-Voting and Identity, VOTE-ID 2009, pp. 122–140. Springer, Heidelberg (2009)

fridaya@uef.fi

Advantages of the PaySim Simulator for Improving Financial Fraud Controls

Edgar A. Lopez-Rojas[1,2(✉)] and Camille Barneaud[3]

[1] Norwegian University of Science and Technology (NTNU), Trondheim, Norway
edgar.lopez@ntnu.no
[2] Simudyne, London, UK
edgar@simudyne.com
[3] Flaminem, Neuilly-sur-Seine, France
camille.barneaud@flaminem.com
http://edgarlopez.net, http://simudyne.com, http://flaminem.com

Abstract. Financial transactions lay up the foundation of modern society. Unfortunately, illicit abuse of the financial system is pervasive. Fraud controls aim to detect these suspicious activities, but they require deep analysis to model their efficacy and value proposition. Due to the private nature and scale of these financial transactions, this analysis is often performed in hindsight. Financial institutions lack of information, due to *the hidden fraud problem*, to properly set and tune their fraud controls systems. This is probably one of the reasons we are losing the war against crime. This paper presents PaySim, a cutting edge agent-based model that simulates financial fraud scenarios to improve current fraud controls. PaySim uses aggregated anonymized data from a real financial dataset to generate synthetic data that closely resembles the transactions dynamics, statistical properties and causal dynamics observed in the original dataset, while incorporating any malicious behaviour of interest. Using an agent-based framework specifically designed to cover the demands of financial simulation and the application of mathematical statistics, we leverage a real-life scenario based on a known fraud scheme to demonstrate the advantage of simulated data over real-world data when setting adequate controls for fraud detection.

Keywords: Multi-agent-based simulation · MABS · Financial fraud · Mobile money · Fraud detection · Synthetic data · Agent-based modeling · Multi-agent systems · Reinforcement learning · Machine learning

1 Introduction

As economies continue to move from analogue to digital economies, they become more accessible and often more vulnerable to breaches. Financial fraud controls are insufficient to stop the wrongdoers. In fact, many are simply focused on meeting minimum legal requirements. Around 90% of Europe's largest banks

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 727–736, 2019.
https://doi.org/10.1007/978-3-030-22868-2_51

fridaya@uef.fi

have been sanctioned for money laundering, and the number of fraud cases just keep growing world wide. Some examples of this phenomenon includes HSBC in Mexico[1], Swedbank, Nordea and Handlesbanken in Sweden[2] and more recently the case of ING in the Netherlands[3]. On top of that there are several difficulties for researchers that work on fraud controls, both from academia and industry to get access to financial data due to the private nature of it. These restrictions are increasing with the recent introduction of GDPR and its implication in financial fraud research [4].

Criminals are getting more sophisticated at targeting institutions' permeable controls and weakest links. Privacy is critical as fines are ever more frequent as breaches continue to accelerate. On the other hand there are promising technologies that are showing remarkable progress in the area of fraud control such as machine learning. However, machine learning and AI are restricted from using personally identifiable data as a result of restrictions such as GDPR.

This paper presents PaySim, a cutting edge agent-based model that simulates financial fraud scenarios to improve current fraud controls. PaySim uses aggregated anonymised data from a real financial dataset to generate synthetic data that closely resembles the transactions dynamics, statistical properties and causal dynamics observed in the original dataset, while incorporating any malicious behaviour of interest. Using an agent-based framework specifically designed to cover the demands of financial simulation and the application of mathematical statistics, we leverage a real-life scenario based on a known fraud scheme to demonstrate the advantage of simulated data over real-world data when setting adequate controls for fraud detection.

The intended audience for the paper are those who are in need to have a tool for collaborating between restricted financial domains and researchers who lack valuable data. This research can also benefit people working with financial service institutions, financial regulators, academic institutions and similar related industries. Any organisation that wants an alternative method to share financial information and avoid GDPR barriers may also interested in this work because it enables access to "advanced modelling techniques" that allow them to model feasible scenarios for testing and measuring their fraud controls methods.

[1] HSBC fined with 1.9 Billions USD for money laundering in 2012. https://www.bbc.com/news/business-20673466.

[2] Swedbank 2019 https://www.swedbank.com/about-swedbank/information-on-tv-reports-on-money-laundering.html. Nordea and Handelsbanken are fined with 50m SEK and 35m SEK in 2015. https://www.reuters.com/article/nordea-bank-handelsbanken-fsa/update-1-nordea-handelsbanken-fined-over-money-laundering-breaches-idUSL5N0YA1MS20150519.

[3] ING fined with 700m USD for money laundering in 2018. https://www.reuters.com/article/us-ing-groep-settlement-money-laundering/dutch-bank-ing-fined-900-million-for-failing-to-spot-money-laundering-idUSKCN1LK0PE.

fridaya@uef.fi

2 Background and Context

The concept of financial simulators has been developed earlier to create models of financial markets and financial forecasts [1,11,12]. Simulation has also been applied to solve cyber-security problems [2].

After the enforcement of GDPR in late May 2018, many organisations are interested in new methods that either comply with or avoid handling personal information. Simulation has the potential solution to this, by enabling sharing without disclosure of personal data. Financial simulation is a novel and valid approach in financial fraud analytics which involves the use of simulators to produce a sufficient level of financial data containing both the normal and the fraudulent behaviour [6,7,14]. Using simulators and the generated synthetic datasets enables researchers to have a test environment for developing the required controls.

Recent studies suggest that there is a need for better data to properly address the evaluation and status of fraud controls to complex problems such as Anti-Money Laundering, better known as AML [3]. The majority of policies and controls are based on classic computational controls such as thresholds. However, there is a body of research that shows that novel techniques such as machine learning can benefit from using synthetic data generation methods to target more accurately and efficiently suspicious transactions [13].

Financial simulation can be the key to unlock the synergies and mitigate the barriers for cooperation between Academia, the Financial Sector and the Government by using simulation as the central tool to bridge the parties. Through models such as the triple helix model for AML (TH-AML) we can address in a collaborative way complex problem such as money laundering [10].

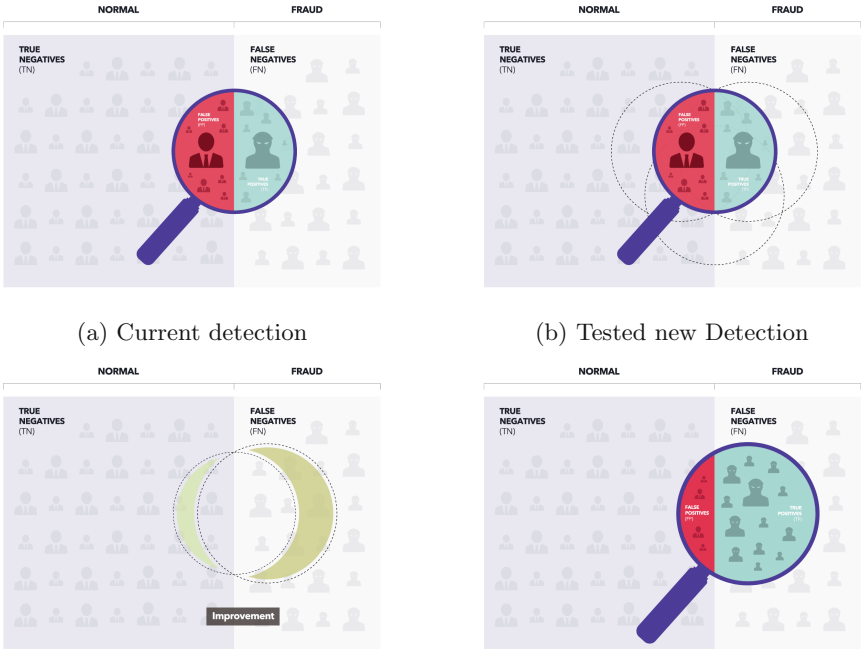
3 The Problem of Developing Better Fraud Controls

There is an inherent problem to develop better controls in the financial fraud domain. The main issue relies on the unknown metric of the total population of fraud. Without a clear measure of this, current fraud detection methods only rely on improving the false positives and the true positives. Reducing the false negative rate is also a desirable measure of improvement in a dynamic and constantly changing fraudulent environment.

In Fig. 1(a) we represent the current status of a fraud detection method inside a financial institution. The whole area represents the whole universe of transactions. For improving fraud controls financial institutions aim to efficiently minimise fraud risk while reducing cost by reducing the cases of false negatives and at the same time aiming to comply with regulators. But a real improvement of the detection system should be in the direction of reducing the total fraud as depicted in Fig. 1(b and c). At the moment, the total value of fraud (FN+TP) is unknown.

From Fig. 2 we can see that from the two more relevant metrics for fraud detection, only Precision (1) is computable for financial institutions. Recall (2) uses the false negative (FN) to compute which is the hidden fraud. This gives

fridaya@uef.fi



(a) Current detection

(b) Tested new Detection

(c) Measure the improvement

(d) Implement new method

Fig. 1. Visualisation of a fraud control improvement

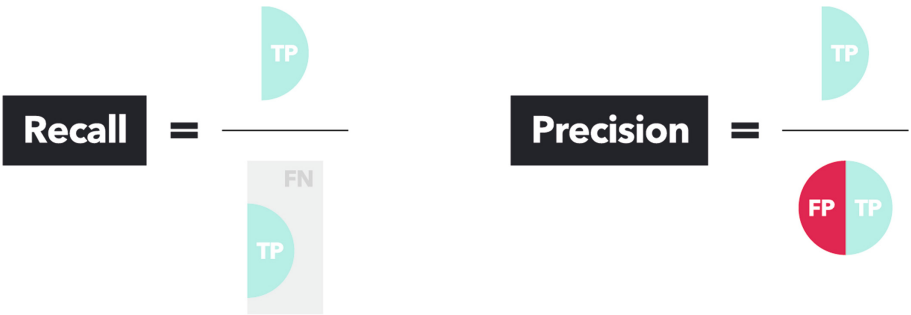


Fig. 2. Recall and Precision visually explained

only half of the equation for combined metrics such as the Harmonic F-score (3) which balance the weight to Precision and Recall or the F-score using beta (4) which gives more weight to either Precision or Recall. This gives the possibility for a financial institution to define an objective function that realistically fights the crime while considering the resources of the organisation.

$$Precision = \frac{TP}{FP + TP} \tag{1}$$

fridaya@uef.fi

$$Recall = \frac{TP}{FN + TP} \tag{2}$$

$$F_2 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{3}$$

$$F_{score} = (1 - \beta^2) * \frac{Precision * Recall}{\beta^2 * Precision + Recall} \tag{4}$$

The *hidden fraud problem* adds up to several other drawbacks that financial institutions are facing at the moment when facing the task of effectively and efficiently detect fraudulent activities.

Some of the drawbacks of current fraud control techniques can be summarised as:

- Historical data is not rich enough to develop better controls.
- Data privacy has restricted personal data use therefore restricting the potential for third parties to help in the problem.
- Fraudsters are more adaptive than financial institutions, so the race is unfair.
- There are not enough diverse known cases of categorised fraud data to tune fraud controls.
- There are far too many false positives, because criminal aim to disguise themselves among regular clients.

In this paper we argue that simulation provides a set of advantages for the task that financial institutions have of detecting fraud activity.

Some of the advantages of using simulation for implementing fraud control can be summarised as:

- Assessing and identifying key weaknesses and putting in place cost effective countermeasures.
- Accurately measuring the cost of visible fraud and hidden fraud.
- Proactively preparing for and using precisely the right approach to minimise the cost and risk for future fraud.
- Anticipate the future before it happens by testing and training up their approach.
- By using synthetic and forensic data enriched by a realistic simulation they can experiment with fraud behaviour in a controlled environment.
- Rapidly adapt and scale to address changes across both fraudulent behaviour and suspicious activity as well as the evolving regulatory environment.

4 PaySim Demo

In this demo we would like to show how the PaySim [9] simulator is used to generate synthetic datasets based on real world datasets of financial data to study and perform research on fraudulent financial schemes such as money laundering and terrorism financing to evaluate the effectiveness and impact of diverse controls. By using synthetic datasets, the researchers have the opportunity to

fridaya@uef.fi

explore diverse and rich scenarios of fraud that would otherwise be challenging or unattainable even with access to real data.

The new version of PaySim uses an agent-based framework, specifically designed by to cover the functional and computational demands of financial simulation with the explicit purpose of developing fraud analytics. The way it works is depicted in Fig. 3. It uses real data to calibrate the required parameters that allow the generation of synthetic datasets. Once the calibration satisfies the similarity comparison between the real and the synthetic data, we output synthetic datasets that are used to study fraud phenomenons and that contain labelled instances of generated fraud schemes. This process can be iterative and can be used to generate diverse scenarios of fraud or predictive future scenarios that otherwise are not present in the real data.

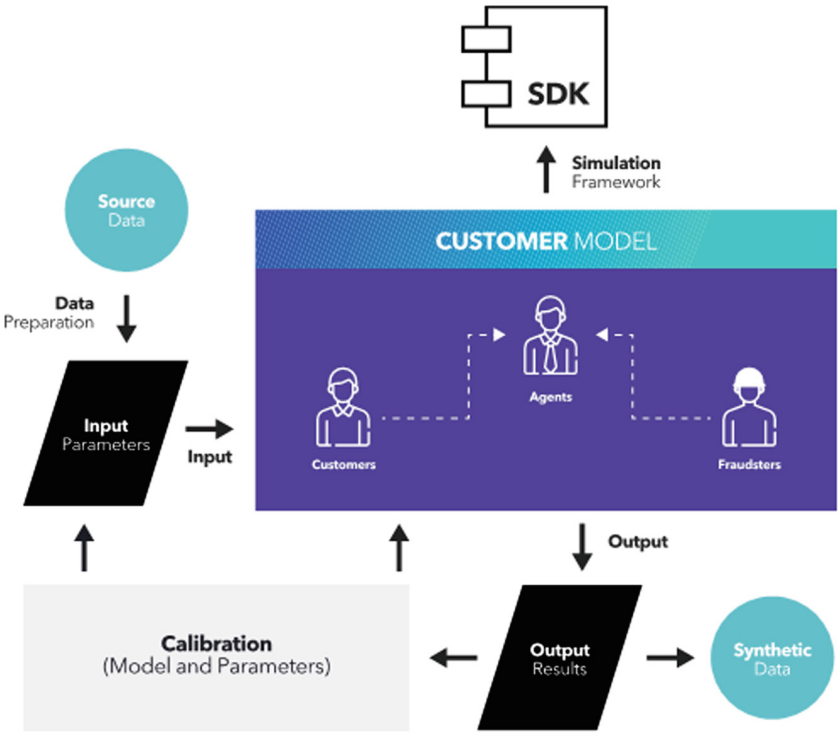


Fig. 3. Simulation process of PaySim

5 Results

The scenario selected is based on the fraud case presented in previous work [8] which happens when the client loses control and access to his/her account. The fraudster takes control and uses disposable mule accounts to transfer the money, and later cashes them out. All of this can only happen in a very short time, because when the clients discover that their accounts are compromised, the first

fridaya@uef.fi

action they should take is to contact customer service to block any possible further malicious transactions. We want to present a scorecard metric to evaluate which way we can achieve better performance of the fraud control.

The experiment introduces a 3% probability for each of the 1000 fraudsters to perform fraud at any given step of the simulation, which is perhaps an aggressive value (30 fraudulent activities per hour), but this helps to inject enough fraudulent activity to study the phenomenon.

In order to study this phenomenon, we ran the system four times increasing the maximum amount of transaction possible in a single TRANSFER each time. We selected four synthetic datasets that used the thresholds on transfer transactions of 300k (PS89745), 600k (PS80775), 900k (PS00273) and 1200k (PS98516). The first case obligates the fraudster to perform several operations to empty accounts that contains a balance above this threshold. The last limit is very flexible and allows the fraudsters to perform a single TRANSFER operation in most cases. By implementing an extra control that will temporarily block accounts that exceed three consecutive transfers for the maximum amount in a short period of time we could effectively reduce the amount of fraud and measure the benefit of this control. With the help of PaySim, we can also measure how other users will be affected by this block in their accounts (False Positives).

Table 1 shows the number of transactions type TRANSFER and the classification of fraud. The first obvious and important thing to notice is that whenever there is a control, the effort required for committing fraud gets higher. Just by introducing a lower threshold on the maximum amount allowed for a transfer, the number of transactions needed to empty an account increases several times. However, the number of legitimate users that will be affected increases. This is the trade-off in fraud detection that a manager needs to address in enacting the fraud controls. Table 1 also shows the loss due to fraud. If we focus attention on the False Negative (FN) row of each simulation, we can see the profit from fraud. The bigger the threshold the higher the profit. The task for a manager is to reduce this amount while minimising False Positives (FP) cases, which are legitimate customers that have their account blocked by the controls implemented to prevent fraud.

6 Discussion

Table 2 show the fraud detection results of each of the datasets evaluated here. We see that the precision is higher when the threshold is higher as in dataset *PS98516* (1200k). This means that we will have fewer customers affected. However, the recall is seriously affected, which means that the fraudsters will profit more using this control (16,189 million).

On the other hand when we have a lower threshold as in *PS89745* (300k), the number of false positives (FP) increases to 982. But, we have a considerably higher recall which means that we lower the total value of fraud (6,724 million). The primary conclusion we have from using simulation is that we are able to measure any require metric to be able to improve the performance of fraud controls systems.

fridaya@uef.fi

Table 1. Fraud detection classification

LogName	Class	Count	Amount	% count	% amount
PS89745 (300k)	FN	27,412	6,724M	1.005%	0.363%
	FP	982	214M	0.036%	0.012%
	TN	2,607,642	1,816,764M	95.579%	98.162%
	TP	92,211	27,076M	3.380%	1.463%
PS80775 (600k)	FN	24,400	11,291M	0.990%	0.581%
	FP	58	17M	0.002%	0.001%
	TN	2,396,684	1,907,409M	97.239%	98.126%
	TP	43,604	25,114M	1.769%	1.292%
PS00273 (900k)	FN	21,072	12,854M	1.024%	0.768%
	FP	8	1M	0.000%	0.000%
	TN	2,011,006	1,639,699M	97.712%	97.903%
	TP	26,006	22,264M	1.264%	1.329%
PS98516 (1200k)	FN	20,493	16,189M	0.921%	0.858%
	FP	1	0.168M	0.000%	0.000%
	TN	2,186,516	1,849,707M	98.215%	97.993%
	TP	19,248	21,686M	0.865%	1.149%

Table 2. Fraud detection results

LogName	Precision	Recall
PS89745	98.946%	77.085%
PS80775	99.867%	64.120%
PS00273	99.969%	55.240%
PS98516	99.995%	48.434%

Financial institutions lack of information to properly set and tune their fraud control systems. This is probably one of the reasons we are losing the war against crime. From Table 1 we can see that by applying strict controls we are obtaining a high number of False Positives (982 in case of 300k). The opposite phenomenon occurs when we apply very tolerant controls as in the case of 1200k threshold where we obtain only one FP. When we have strict controls the profit of the undetected fraud (False Negatives) is considerably less than when we have tolerant controls. However, the financial institutions lack of the knowledge of the False Negatives to properly tune their systems in an adequate way that satisfy not only the regulators but also decreases in a considerable way the profit from the criminal activities.

Simulation has limitations as presented in previous research [5]. We believe, simulation for financial fraud is not widely used at the moment because it is not a trivial task to produce a quality model and synthetic dataset. However, the shortcomings if properly handled are less than the advantages presented here.

fridaya@uef.fi

7 Conclusions

The primary conclusion we have from using simulation is that we are able to measure any require metric such as the Precision and the Recall to be able to improve the performance of fraud controls systems. Financial institutions lack of the knowledge of the hidden fraud, which is in terms of fraud detection the False Negatives, to properly tune their systems in an adequate way that satisfies not only the regulators but also decreases in a considerable way the profit from the criminal activities.

By using financial simulators such as the PaySim simulator we are able to obtain enough information and explore diverse fraud scenarios to deliver a real improvement of fraud detection systems which will tackle both, the cost of the business due to miss classification and the reduction of the criminal profit.

At the moment we are able to produce realistic synthetic data and desirable future fraud scenarios to study the fraud phenomenon. Future work should focus on preparing simulation for easy sharing models and parameters as if they were word-processors documents to enable models such as the Triple Helix Approach for AML (TH-AML) mentioned earlier [10].

Acknowledgments. This work was sponsored by Simudyne (simudyne.com) London, UK.

References

1. Hedjazi, B., Ahmed-Nacer, M., Aknine, S., Benatchba, K.: Multi-agent financial market simulation: evolutionist approach. Int. J. Simul. Process Model. **8**(2–3), 185–199 (2013)
2. Legato, P., Mazza, R.: A simulation optimisation-based approach for team building in cyber security. Int. J. Simul. Process Model. **11**(6), 430–442 (2017)
3. Levi, M., Reuter, P., Halliday, T.: Can the AML system be evaluated without better data? Crime Law Soc. Change **69**, 307–328 (2018)
4. Lopez Rojas, E.A., Gultemen, D., Zoto, E.: On the GDPR introduction in EU and its impact on financial fraud research. In: The 30th European Modeling and Simulation Symposium-EMSS, Budapest, Hungary (2018)
5. Lopez-Rojas, E. A., Axelsson, S.: Money laundering detection using synthetic data. In: Karlsson, L., Bidot, J. (eds.) The 27th workshop of (SAIS), Örebro, pp. 33–40. Linköping University Electronic Press (2012)
6. Lopez-Rojas, E. A., Axelsson, S.: BankSim: a bank payments simulator for fraud detection research. In: 26th European Modeling and Simulation Symposium, EMSS 2014, pp. 144–152. Dime University of Genoa (2014)
7. Lopez-Rojas, E. A., Axelsson, S.: A review of computer simulation for fraud detection research in financial datasets. In: Future Technologies Conference, San Francisco, USA (2016)
8. Lopez-Rojas, E.A., Axelsson, S., Baca, D.: Analysis of fraud controls using the PaySim financial simulator. Int. J. Simul. Process Model. **13**(4), 377–386 (2018)
9. Lopez-Rojas, E., Elmir, A., Axelsson, S.: PaySim: a financial mobile money simulator for fraud detection. In: The 28th European Modeling and Simulation Symposium-EMSS, Larnaca, Cyprus (2016)

fridaya@uef.fi

10. Lopez-Rojas, E.A., Zoto, E.: Triple helix approach for anti-money laundering (AML) research using synthetic data generation methods. In: The 10th International Conference on Society and Information Technologies, ICSIT (2019)
11. O'Loughlin, M.J., Driskell, M.K., Diehl, G.: Financial simulation: combining cost information in systems analysis. In: Winter Simulation Conference Proceedings, pp. 578–581 (1990)
12. Takaishi, T.: Multiple time series ising model for financial market simulations. J. Phys.: Conf. Ser., **574**(1) (2014)
13. Yue, D., Wu, X., Wang, Y., Li, Y., Chu, C.-H.: A review of data mining-based financial fraud detection research. In: 2007 International Conference on Wireless Communications, Networking and Mobile Computing, pp. 5514–5517. IEEE, September 2007
14. Zintgraf, L., Lopez-Rojas, E.A., Roijers, D., Nowe, A.: MultiMAuS: a multi-modal authentication simulator for fraud detection research. In: Proceedings of the European Modeling and Simulation Symposium (EMSS), September 2017

fridaya@uef.fi

Secure Cryptosystem Using Randomized Rail Fence Cipher for Mobile Devices

Amit Banerjee, Mahamudul Hasan, and Him Kafle(✉)

South Asian University, New Delhi 110021, India
rana.nstu@gmail.com, kaflehim@gmail.com

Abstract. Transferring files between mobile devices through a wireless medium has security implications. Particularly for applications that transfer personal information like healthcare records or location information. Designing an efficient and secure cryptosystem for these resource-constrained modern smart devices is a challenging task. In this paper, we address the same and propose a light-weight, energy-efficient and secure cryptosystem based on rail-fence cipher (RFC). We refer the proposed algorithm as "randomized rail fence cipher (RRFC)" as it uses pseudo-random numbers to select the fences for transposing characters of a file. The idea is to remove security vulnerabilities of RFC by increasing the complexity of determining the pattern used for the transposition. We implement the proposed algorithm and experimentally analyze the performances of the algorithm on mobile devices. We also present attack analysis for the proposed cryptosystem.

Keywords: Randomized Rail Fence Cipher · Rail cipher · Crypto-system · Transposition cipher · Security

1 Introduction

Designing an efficient and secure cryptosystem for modern smart devices (i.e., smart watch, smart TV, Smart glass) is a challenging task. The smart devices are often used for storing personal information such as (image, financial information, health-related data) and for sharing them with the cloud for obtaining different services. Although the hardware capabilities of these devices are improving every day, executing computationally complex operations on these resource constraints devices is still a challenge. In [1,2], researchers have discussed the issues related to encryption of large files, using the block cipher (i.e., AES, DES) on these devices, and have also presented the limitations for the computation and battery lifetime of the same. Thus, there is a need for designing new lightweight and energy efficient algorithms for smart devices.

The transposition cipher is a lightweight cryptographic algorithm and one of the oldest, well-studied cipher technique. In transposition cipher, the ciphertext (CT) is obtained from the permutations of the characters of the plaintext (PT). As shown in Fig. 1, the character set both in PT and CT is same, but the position

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 737–750, 2019.
https://doi.org/10.1007/978-3-030-22868-2_52

fridaya@uef.fi

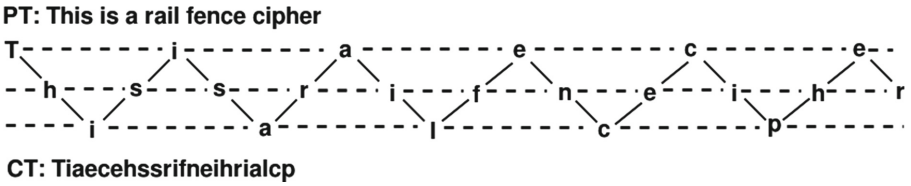


Fig. 1. Basic architecture of proposed protocol

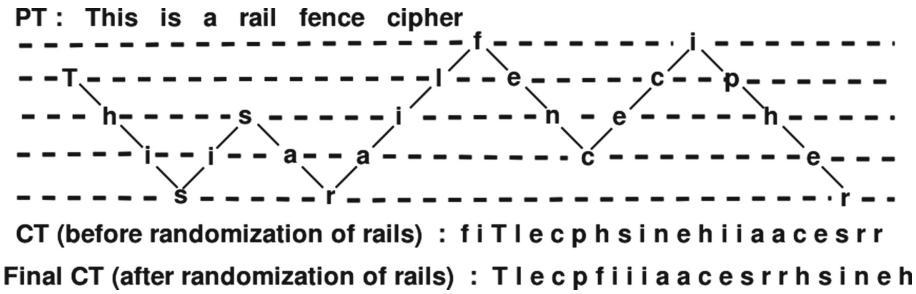


Fig. 2. An example of RRFC

of the characters are changed in CT. The main goal of the transposition cipher is to generate diffusion. It is widely used in cryptographic algorithms and popular transposition cipher includes, rail fence cipher [3], route cipher [4], columnar transposition [7], double transposition [6], myszkowski transposition [5].

In this paper, we propose a Randomized Rail Fence Cipher based on the fundamentals of Rail Fence Cipher (RFC). Rail fence cipher is also known as "Zig-zag Cipher". In this, as shown in Fig. 2 the horizontal lines are considered as "rail", and the diagonal lines are referred to as imaginary "fence" between multiple rails. For generating CT, the characters of PT is written on downwards and diagonally over the rails as well the imaginary fence. When it reaches to the bottom rail, then it moves back up to the top rail again. This process continues until the whole PT is written. Finally, characters are arranged according to the rows of the rail for generating the CT. Suppose we want to encrypt the message "This is a rail fence cipher" using a rail fence cipher and the key length is 3. So the characters of PT is to write in 3 rows, to find out the CT in the following a zig-zag pattern:

Rail fence Cipher is a weak cipher, and if an attacker knows the number rails, then it is very easy to break the rail cipher [7]. In a transposition cipher, an attacker often tries to find the key or pattern in which the characters of PT is permutated to form the CT. In the randomized rail fence cipher (RRFC), we use the fundamentals of classical rail fence cipher, but use the randomization to select the fences to permute the characters. In other words, we use pseudorandom numbers to determine the initial position and length of the fence. After finding the length and initial positions of fences, we randomize the rail by using the

fridaya@uef.fi

key to permute the rail. With a randomized selection of rail and fence, we can generate a secure cipher, that can resist popular attacks, like brute force attack and plaintext attack which we discuss in Sect. 4.

The design considerations of our algorithm are as follows:

- As the proposed algorithm is designed based on Rail Fence Cipher, length and starting position of the fence must be random.
- To design a lightweight algorithm for transmitting the text files between the mobile devices and the algorithm should be able to handle the files with large sizes (Up to 5 MB).
- File transmission should be completed within an acceptable time frame.
- Users should have full control over the whole processes of the algorithm, brings users confidence towards the algorithm, i.e. the encryption should be performed within the mobile device.

We experimentally verify the performance of RRFC using an android based application. The goal is to understand the feasibility of RRFC on the mobile devices. For performance evaluation, we evaluate the time needed for encryption, memory consumption, and battery usages. As MDs differ in computational capabilities, we tested the performance of RRFC on three different devices. We also check the encryption time on various sizes of files (1 mb–10 mb). In Sect. 5, we present the experimental details.

Rest of the paper is organized as follows. In Sect. 2, we describe various approaches for data security in the related works. Section 3 discusses the details of the proposed algorithm. In the next section, we show the experimental design. Section 4 discusses the attack analysis. Finally, we conclude our discussion by Sect. 6.

2 Related Works

Encryption is a technique used for securing users information from unauthorized access. Encryption is needed before sending data over an insecure communication media to other device or servers. Researchers have proposed various encryption/decryption approaches using the fundamentals of cryptography. In this section, we discuss the previous works related to the proposed algorithm. We mainly focus on the algorithms that are based on the transposition cipher. Note that, our main aim is to design an encryption algorithm for transferring file between mobile devices. So, in this section, we also discuss the algorithms that are explicitly designed for MD.

Crypt21 [1] uses the fundamentals of substitution and transposition ciphers. The algorithm is designed for the block cipher, where the length of the key is 4–8 characters, which is extended to make it equal to the PT. The ciphertext is obtained by ceaser substitution and transposition. The proposed method can resist frequency analysis attack, but the related-key attack is possible in Crypt21. In [18], authors propose a transposition cryptosystem based on the basics of

fridaya@uef.fi

Crypt21 for mobile text chatting. A pseudo-random function is used to retrieve a session key for encrypting the text messages.

In [25], Terry Ritter proposes an algorithm with dynamic transpositions using pseudo-random shuffler for shuffling the plaintext. The algorithm can handle byte and bit level shuffling. For byte shuffling, a controller first takes a block and then exchange it byte-by-byte using a pseudo-random number. In the same way, the bits of the characters can be shuffled. The author claims that by shuffling, frequency analysis becomes much harder for an attacker.

In [23] propose an algorithm by combining substitution and transposition cipher to provide the security. CT is generated by three steps, and in the first step, characters are substituted using the secret key. In the next step, rail cipher is used for transposing the cipher to get the temporary ciphertext and finally, at the last step the temporary ciphertext is again substituted by the special characters (*i.e.,* $\#, *$). Moreover, the substitution and transposition method used in the proposed algorithm is weak, and cryptanalysis is very easy because of the complexity of rail cipher lies on the numbers of rails.

Singh et al. [3] proposed a new ciphering method by combing the Caesar and rail fence cipher. In the proposed method, firstly characters are substituted by using Caesar cipher with k bit shifts and then it is transposed using rail cipher. Finally, the CT is generated using the push and pop operations. But the proposed method is not secure, as finding the pattern between the CT and PT is easy. Moreover, known plaintext attack is possible in the proposed method.

In [14], authors explore the feasibility of implementing Advanced Encryption Standard (AES), SERPENT and TWOFISH algorithm in MDs and studied the performance and computational cost of each algorithm. The analysis shows that the performance of TWOFISH is better compared to others in terms of CPU and memory utilization. However, due to high computational complexity, these algorithms are not efficient for the MD [14]. Authors in [13,15] have optimized the AES/DES encryption algorithm. In AES, the main functionality is performed in the S-BOX, which comprises two transformations, multiplicative inverse and followed by an affine transformation. In [13], authors focused on rectifying the S-BOX operations to make the algorithm lightweight and proposed a modified affine transformation with lesser time complexity than the original AES. The authors prove that the proposed optimized algorithm provides similar security as compare to the original version.

Researchers have proposed selective or partial encryption schemes to overcome the resource limitations of MD [16,17]. In this, the important parts of a file are parsed for encryption within the MD, and the remaining parts can be transferred to an external server or stored in the mobile in a plaintext format. Selective encryption can save resources in MD. In [16] authors eliminate the duplicate data in structured files (e.g., pdfs, ppt, docx) before encryption to reduce the amount of data at its source to save the storage space and network bandwidth. Similarly, in [17], data is fragmented both vertically and horizontally to determine the parts of the data for encryption. In [24] author introduce the details of designing the crypto-processors, which is important for hardware implementation of

fridaya@uef.fi

Table 1. Notations used in the proposed algorithm

Notation	Description
C, C'	Cryptographically Secure Pseudo-random number (CSPRN), Key-stream
MD	Mobile Devices
ES	External Server
PT	Plaintext
CT	Ciphertext
F	Array
OA	Off-set Array
DDF	Diagonally Downward Fences
DUF	Diagonally Upward Fences
f	Fences
r	Rails
l_i	Length of i^{th} Fence
s_i	staring point of i^{th} Fence

cryptographic algorithms. The author claims that, more focus is needed on designing and optimizing the crypto-processors through VLSI design.

3 Proposed Methodology

In this section, we discuss the implementation details of the proposed cryptosystem. The design goal of our implementation is to use the basic of transposition for implementing a lightweight and energy efficient encryption algorithm for the MDs.

3.1 System Architecture

Figure 3 shows the basic architecture of the proposed algorithm. In the figure, we assume that user Alice wants to share a file with Bob in a secure manner. For this both Alice and Bob connects to an external server (ES) to obtain a CSPRN which is used in our algorithm to determine the rail and fence randomly. Note that the proposed algorithm is not dependent on the ES, as the CSPRN can be generated within the MDs. The reason for using ES is to offload the computational intensive key generation task to the cloud. (Table 1 shows the important notation used in our paper.)

3.2 Obtaining CSPRN

To fetch CSPRN, a user sends a request to ES, specifying its $user_id$, $file_name$, and $CSPRN_size$. Upon receiving the request, the ES generates the CSPRN and send it back to the user. Researchers have proposed the various algorithm

fridaya@uef.fi

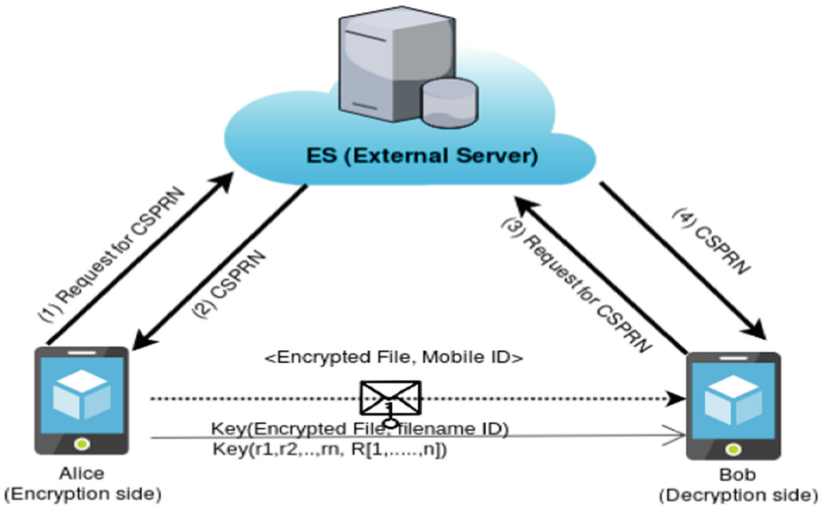


Fig. 3. Basic architecture of proposed protocol

for generating CSPRN (like Middle Square Method (John Von Neumann 1946) [21], Linear Congruential generator (LCG) [22], Cubic Congruential generator (CCG) and others), however in our implementation; we use Advanced Encryption System (AES) 128-bit for the same. The input to the AES algorithm is a randomly generated seed value and a mutually agreed pre-determined key for the corresponding user. The parameters $<s_n, f_n\ (filename)>$ is stored by the user the database of ES for future use. The execution of the AES algorithm generates the 1st 128-bit of the CSPRN. Subsequent 128-bits of CSPRN can be obtained by incrementing the seed value. The details of the CSPRN generation procedure can be found [2]. As the CSPRN is transmitted through an unreliable wireless channel, we modify the CSPRN within the MD (i.e., $C \xrightarrow{S_k} C'$). In [2], we discuss three different ways to modify the CSPRN. In this paper, we use r-CLOAK for generating modified C'.

3.3 Proposed Randomized Rail Fence Cipher (RRFC)

In the following, we use Algorithm 1 to discuss the details of the randomized rail cipher algorithm. In RRFC, we first randomize the length and the starting positions of each Diagonally Downward Fence (DDF). Notice that Diagonally Upward Fences (UDF) can be obtained by joining endpoint and starting points of two consecutive DDFs. Then we use the permutation of the rails to generate the ciphertext. In the following, we assume the user (Alice) has already obtained the CSPRN (method discussed above) form the ES and has modified the CSPRN, i.e. (i.e., $C \xrightarrow{S_k} C'$) by using r-CLOAk algorithm [2].

Notice that a randomize fence can be characterized by two parameters, i.e. it's initial position (S_i) (the rail from where it starts) and the length (l_i).

fridaya@uef.fi

For example, in Fig. 2, the first DDF can be characterized as (2,4). Similarly, the next DDF can be represented by (3,3). Using this notion, we can represent all DDFs of Fig. 2 as $<(2,4)(3,3)(1,4)(1,5)>$. Notice that, the same representation can be equally used for determining the upward diagonal fences.

Thus, we begin by discussing the procedure for determining $<l_i, s_i>$ for each DDFs. Here we assume that the total numbers of DDF is f (i.e. i = 1 ... f). In Sect. 4 we discuss the number of fences (f) required for resisting the brute force attack for the proposed algorithm. To determine the length of DDF (l_i) we divide C' in f equal blocks and use the modulo operation on the consecutive (C'/f) bits (b_i) of each blocks, that is $l_i = (b_i \% r) + 1$. Next, we determine the starting positions (s_i) for DDF, which can be any one of the $(r - l_i)$ rails. The idea is to informally distribute the fences on the available rails (r). We can now compute (s_i) by $s_i = b_i \% ((r - l_i) + 1)$. However notice that, to get the zigzag pattern of a rail fence cipher the starting position of a DDF (s_i) must be above the end position of the previous DDF (i.e s_i less then $(s_{i-1} + l_{i-1})$). to handle this we, compute $s_i = s_i \% (s_{i-1} + l_{i-1})$ as shown in Algorithm 1.

Algorithm 1. Randomized Rail-Fence Cipher (RRFC)

```
   /* Fetch CSPRN (C) of length |C| ≥ 256, and                          */
   /* generate C ← C' using rCLOCK                                      */
   /* Determine stating rail (s_i), length (l_i) of ddf_i               */
   /* and store <s_i, f_i> in an array F                                */
 1 for (i = 1; i ≤ f; i = i + 1) do
 2     b_i := Extract |C'|/f consecutive bits from C'
 3     l_i := b_i mod r + 1
 4     s_i := (b_i mod (r − l_i)) mod (s_{i−1} + l_{i−1}) + 1
 5     Store [s_i, (s_i + l_i − 1)] at array indices F[2i − 1, 2i]
   /* Generate an offset array (OA)                                     */
   /* Assume a permutation of the rails, R[1...r]                       */
 6 for (i = 1, k = 0; i ≤ r; i = i + 1) do
 7     offset := 0
 8     for (j = 1; j ≤ F.length; j = j + 1) do
 9         if R[i] is between F[j] and F[j + 1] then
10             OA[k + +] := |R[i] − F[j]|+ offset
11         offset := offset + |F[j + 1] − F[j]|
   /* Generate the CiphetText (CT)                                      */
12 while EOF do
13     buff = Read_File(file, buff_size)
14     for (i = 1; i ≤ buff_size; i = i + 1) do
15         CT = buff[OA[i]]
```

We use a fence array (F) to store all DFF, from left to right. The array F shows, the initial and final positions of the fences. For RRFC in Fig. 2, the

fridaya@uef.fi

contents of F is as $<2, 5, 3, 5, 1, 4, 1, 5>$. A pair of entries in F starting from an even position shows the DDF. Similarly, for DUF we can consider the pairs starting from the odd indices in F. To compute, the character offset is of each Rail $R[i]$, we determine it's intersection with all fences (i.e. $F[j]\&F[j+1]$). For example, the second fence in Fig. 2 intersect with fence $(2, 5); (5, 1); (1, 4); (4, 1)\&(1, 5)$. This can be easily determined by scanning all consecutive pairs of entries in F.

Similarly, we can calculate the offset of characters between the offset of characters between two adjacent intersecting fences. For this, we can add the lengths of all intermediate fences, as described in Algorithm 1. Considering $(2, 1, 4, 5, 3)$ as a permutation of the rails in Fig. 2. We first compute the offset of characters for rail-2, which is $<1, 11, 13, 17, 19>$, showing the numbers of the characters between two consecutive intersections. Similarly, we can compute the offset for other rails. In the following, we show the Offset Array (OA) for the permutation of R.

Offset Array (OA): {*1, 11, 13, 17, 19, 12, 18, 3, 5, 7, 9, 15, 21, 4, 8, 22, 2, 6, 10, 14, 16, 20*}

Finally, we can use the offset array (OA) for generating the ciphertext (CT). The entries of OA determine the transposition of the characters. For encrypting a large plaintext (PT), we can replicate the entries in OA to transpose the subsequent characters of PT. After completing the encryption process, Alice can transmit the CT either to the ES or Bob directly. The information sent to the Bob includes an encrypted file (CT) and a key required for decryption. The key contains the file id, which is require for downloading the CSPRN, the random numbers $(r1, r2, ..., r_n)$ for converting $C \rightarrow C'$ and the randomization information of the rail i.e. R[1, .. ,r] as shown in Fig. 3. We assume that the key can be transmitted using a secure channel or by using a secure key exchange algorithm like Diffie-Hellman key exchange [12], Password-authenticated key agreement [11], Post-quantum key exchange [8].

In Fig. 2, we show an example of a PT and its corresponding CT using Algorithm 1. To decrypt the file, Bob can use reverse transposition to obtain the PT. The procedure is similar to the above and can be performed by modifying the Algorithm 1. Notice that, a simple extension of the proposed algorithm is to perform the transposition in terms of bits, rather than characters. For example, we can consider n-bits for each character and use RRFC for our transposition. For this, we may need to replicate the OA to obtain the remaining bits of the characters in the subsequent PT.

3.4 Modifying RRFC

In the following, we discuss the modification on RRFC for improving security. However, for our implementation in this paper, we use the above algorithm as the modification discussed below don't affect the complexity of the algorithm.

fridaya@uef.fi

Increasing the Length of Offset Array (OA). In our implementation, the length of an OA plays an important role. Notice that it is easy to break RRFC if the size of OA is small. For generating a large OA, we can increase the number of rails and fences used by the algorithm. Another approach for increasing the length of OA is by increasing the size of the array R in Algorithm 1, which is used to store a permutation of the rails. In this modification, we do not need to increase the number of rails or fences. Instead we increase the size of R to m, where $m > r$ and place the rails $[1, 2, ..., r]$ uniformly in R.

Implementing RRFC Using Random Bits. Another extension of RRFC can be to perform transposition in terms of bits, rather than characters. For example, the ciphertext can be generated by considering a fixed n bits from each character. Similarly, we can also use a random number of bits from each fence to produce the CT. For this modification, we need to associate each fence with a random number (e.g., between 1 to 8) and use the OA of Algorithm 1 for extracting random bits from the character of each fence. The subsequent bits of a character can be retrieved by replicating OA.

4 Attack Analysis

To compromise the proposed system, an attacker may try to retrieve the CSPRN from the external server (ES) or the communication media. However, as we are using r-CLOAK to modify the CSPRN within the mobile device (MD), i.e., $C \rightarrow C'$, retrieving C is not sufficient to compromise the proposed RRFC. The detail security analysis for the same can be found in [2]. In our design, we assume that all operations, i.e., $C \rightarrow C'$ and CT generation are performed within a secure MD, which is a general assumption of any encryption algorithm, i.e., the device on which the encryption algorithm is running must be secure. Hence, the only way an attacker can compromise RRFC is by attacking the algorithm itself. In the following, we present the attack analysis and investigate the robustness of RRFC against such attacks.

4.1 Brute Force Attack

Two important parameters of RRFC are the initial position (s_i) and the length (l_i) of the random fences used for our transpositions. Assuming the fixed number of rails (r), random fences can be selected in $\binom{r}{2}$ ways. That is, for selecting a (l_i, s_i) for a particular fence, an attacker need to consider all $\binom{r}{2}$ combinations. Hence, for selecting f fences, the complexity of determining all (l_i, s_i) is $\binom{r}{2}^f$. After finding the all rails we permute the rails with rail permutation key which is itself need $r!$ combination to find. So, in combine total complexity for brute force attack is $\binom{r}{2}^f * r!$. Thus with increasing number of rails (r) and fences (f), the complexity of a brute force attack on RRFC increases. For $r = f = 18$, the brute force complexity is greater than 2^{128}, which is comparable to a standard secure cryptosystem.

fridaya@uef.fi

4.2 Known Plaintext Attack (KPA)

A known plaintext attack assumes that an adversary knows a sequence of characters of a plaintext (PT) and the corresponding ciphertext (CT) and tries to find the pattern or key for generating the subsequent characters of the PT. KPA is a special case of plaintext attack where the adversary knows some part of plaintext in advance. In a proposed RRFC, attacker may try to determine the (s_i, l_i) from the known characters of PT and CT. However, notice that the complexity of determining (l_i, s_i) is the same as that of the brute force attack discussed above. This is because the attackers need to try all possible combinations of (l_i, s_i) to determine the subsequent characters of the PT.

4.3 Reused Key Attack

In reused key attack, for a block cipher which uses the same encryption key for multiple blocks, an attacker collect sufficient ciphertext and try to guess the pattern between the ciphertext and get the key. This attack is highly used to decrypt the simple substitution cipher. In our proposed system, an attacker may try to determine the offset array (OA) used by RRFC for the transformation by analyzing a large amount of ciphertext. Researchers in [9,10] have shown that in a transposition cipher if the offset array (or the key length) is small and the attacker gets enough characters of PT and CT, then it is possible to find the key used for the transposition. In [10], authors successfully generated the key of length $\leq$25 used in a transposition cipher by using bigram and trigram in a genetic algorithm. Authors also show that the complexity for the same increases with increasing key length. To resist this attack, we can use a large offset array (OA) (Sect. 3.4 describe how can we extend the size of OA), finding the pattern used by RRFC can be complicated. Moreover, in Sect. 3.4, we discuss an extension of RRFC using bitwise transformation, which can be used for avoiding such attack as the ciphertext characters are not the same as that of PT.

5 Experimental Results

The performance of an algorithm depends on its complexity and the hardware capabilities (i.e., processor, memory, cache) on which it is tested. With the increasing demands, the hardware capabilities of the mobile devices (MDs) are improving day by day [19,20]. But, there exists a huge gap between a highly configured and a low configuration mobile device. Resource-intensive applications may not provide the same performances on all MDS. So, we test the performance of RRFC on three MDs with different configurations, shown in Table 2.

For the first experiment, we fix the file size as 5 MB and check the performance on three MDs listed in Table 2. The RRFC encryption time on the MDs is shown in Fig. 4. The experimental result shows the successful execution of RRFC on a wide range of MDs and the execution time for encrypting a 5 MB file is between 8 to 16 s depending upon the configuration of the MD. Due to the better configuration of mobile M-3, its encryption time is less compared to others.

fridaya@uef.fi

Table 2. Mobile device configuration

Mobiles	M-1	M-2	M-3
OS	Naugaut 7.1	Oreo 8.1.0	Oreo 8.1.1
API level	25	27	27
CPU	Octa-core 1.8 GHz	Octa-core 2.0 GHz	Octa-core 2.3 GHz
Chipset	Qualcomm Snapdragon 616	MediaTek Helio P60	Qualcomm Snapdragon 845
RAM	2 GB	4 GB	6 GB
Battery	Li-Po 3600 mAh	Li-Ion 3410 mAh	Li-Ion 3300 mAh

We also test the performance of RRFC on different file sizes. For this, we choose mobile M-3, as it shows the best result compared to others. Figure 5 shows that the total time for encryption increases with increasing file size. The experimental result also shows that the encryption time varies linearly with increasing file size.

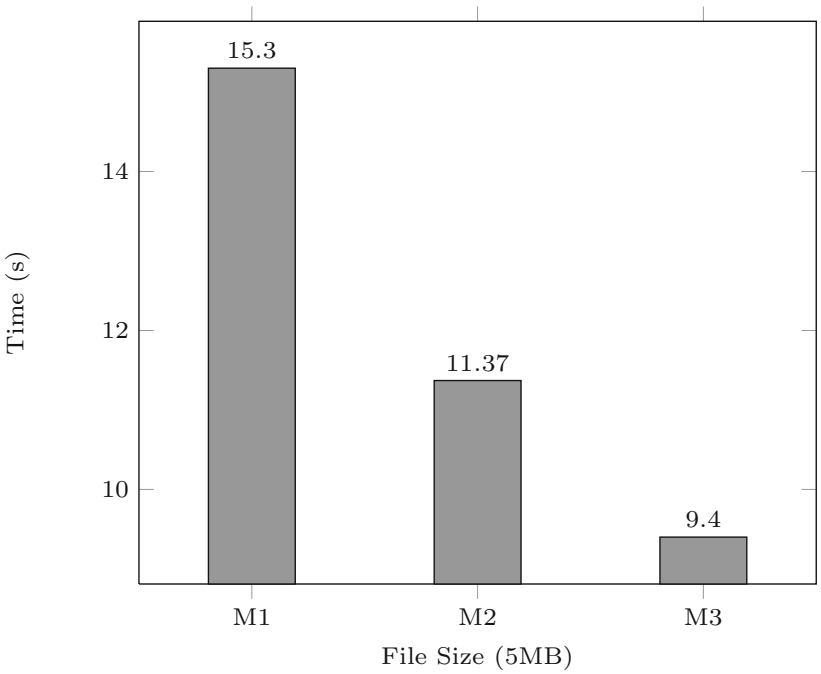


Fig. 4. Encryption time for three mobile devices

The installable APK (Android Package Kit) size of the proposed algorithm is 1.9 MB. The space occupied after installation may vary depending upon the versions of the Android operating system running on the MD. We have installed

fridaya@uef.fi

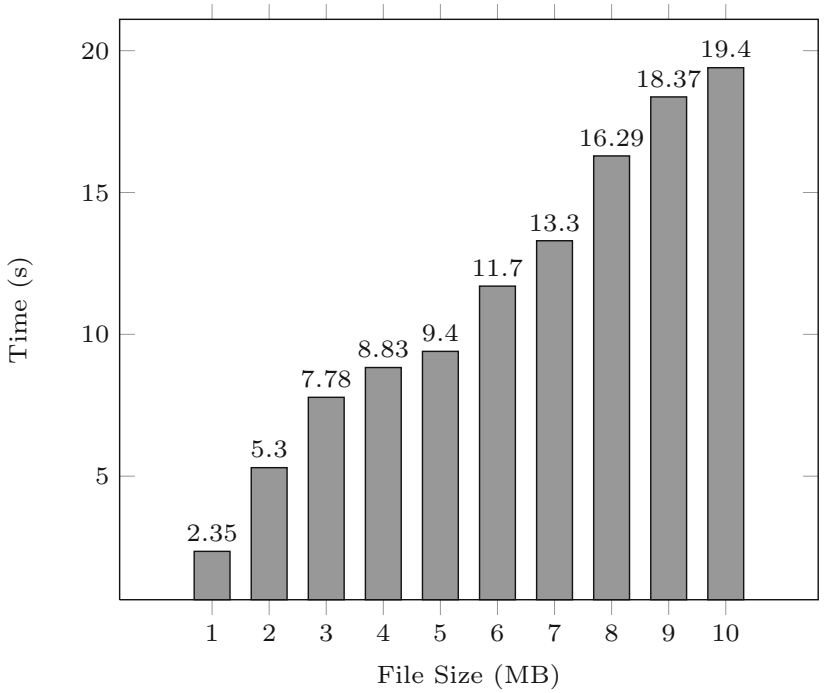


Fig. 5. Encryption time for different file sizes

applications on the MDs listed above and found that the application requires approximately 2.11–3.23 MB storage space after installation. Hence, the application is lightweight, and it can be installed on MD with less memory.

We used the "GSam Battery Monitor" [26] Android application to measure the battery performance of our application on *M1* with Li-Ion 3600 mAh battery. To measure the battery consumption, we launch our application and perform the encryption operation on five files ranging from 1 MB to 5 MB. We notice 1% decrease in the battery level, which includes the power consumed by the screen, wifi and other background processes. According to data recorded by GSam app, the battery consumption of our app is only 5% of the total 1% battery consumption. Hence, the RRFC protocol is efficient and can operate on MD with low configuration.

6 Conclusion

In this paper, we present a transposition based cryptosystem using a pseudo-random number for transferring files between mobile devices, called "randomized rail fence cipher (RRFC)". RRFC uses pseudorandom numbers to select the initial and final position of the fences. The idea is to increase the complexity

fridaya@uef.fi

of determining the pattern used for the transposition. Our experimental evaluation shows that the proposed algorithm is efficient (in terms of complexity), lightweight and energy efficient and can be used for encryption in mobile devices. We also perform attack analysis on RRFC and show its resistance against popular attacks like Brute force and known plaintext attacks. Our future work includes further evaluation of RRFC against other attacks and to reduce the complexity for the mobile devices.

References

1. Malik, S.: A novel key-based transposition scheme for text encryption. In: Frontiers of Information Technology (FIT). IEEE (2011)
2. Banerjee, A., et al.: CLOAK: a stream cipher based encryption protocol for mobile cloud computing. IEEE Access **5**, 17678–17691 (2017)
3. Singh, A., Nandal, A., Malik, S.: Implementation of caesar cipher with rail fence for enhancing data security. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **2**(12) (2012)
4. Rail Fence Cipher: Route Cipher. Columnar Transposition Cipher. http://en.wikipedia.org/Transposition
5. Bhowmick, A., Geetha, M.: Enhancing resistance of hill cipher using columnar and Myszkowski transposition. Int. J. Comput. Sci. Eng. **3**(2), 20–25 (2015)
6. EVLNA WIRE, CAEEN SOIAE DRLEF WEDRE EVTOC: 19.4 Double transposition. Classic Cryptography: 47
7. Millikin, D.D.: Elementary Cryptography & Cryptanalysis. New York University Bookstore, New York (1943)
8. Alkim, E., et al.: Post-quantum key exchange-a new hope. In: USENIX Security Symposium, vol. 2016 (2016)
9. Boryczka, U., Dworak, K.: Cryptanalysis of transposition cipher using evolutionary algorithms. In: International Conference on Computational Collective Intelligence. Springer, Cham (2014)
10. Heydari, M., Shabgahi, G.L., Heydari, M.M.: Cryptanalysis of transposition ciphers with long key lengths using an improved genetic algorithm. World Appl. Sci. J. **21**(8), 1194–1199 (2013)
11. Asokan, N., Ginzboorg, P.: Key agreement in ad hoc networks. Comput. Commun. **23**(17), 1627–1637 (2000)
12. Al-Aali, G., Boneau, B., Landers, K.: Diffie-Hellman key exchange. In: Proceedings of CSE 331, Data Structures Fall 2000, p. 67 (2000)
13. Sahoo, O.B., Kole, D.K., Rahaman, H.: An optimized S-box for advanced encryption standard (AES) design. In: International Conference on Advances in Computing and Communications (ICACC). IEEE (2012)
14. Kofuji, S.T.: Performance analysis of encryption algorithms on mobile devices. In: 47th International Carnahan Conference on Security Technology (ICCST). IEEE (2013)
15. Wang, H., et al.: Improved lightweight encryption algorithm based on optimized S-box. In: Fifth International Conference on Computational and Information Sciences (ICCIS). IEEE (2013)
16. Song, S., Choi, B.-Y., Kim, D.: Selective encryption and component-oriented deduplication for mobile cloud data computing. In: International Conference on Computing, Networking and Communications (ICNC). IEEE (2016)

fridaya@uef.fi

17. Krishna, R.K.N.S., et al.: Privacy-preserving data management in mobile environments: a partial encryption approach. In: IEEE 13th International Conference on Mobile Data Management. IEEE (2012)
18. Chen, H.-C., Epa, A.L.V.: A rotation session key-based transposition cryptosystem scheme applied to mobile text chatting. In: IEEE 28th International Conference on Advanced Information Networking and Applications. IEEE (2014)
19. Carroll, A., Heiser, G.: An analysis of power consumption in a smartphone. In: USENIX Annual Technical Conference, vol. 14 (2010)
20. Caudill, J.G.: The growth of m-learning and the growth of mobile computing: parallel developments. Int. Rev. Res. Open Distrib. Learn. **8**(2) (2007)
21. von Neumann, J.: John von Neumann. American Mathematical Society (1988)
22. L'ecuyer, P.: Tables of linear congruential generators of different sizes and good lattice structure. Math. Comput. Am. Math. Soc. **68**(225), 249–260 (1999)
23. Bhargava, U., et al.: A new algorithm combining substitution & transposition cipher techniques for secure communication. In: International Conference on Trends in Electronics and Informatics (ICEI). IEEE (2017)
24. Sklavos, N.: On the hardware implementation cost of crypto-processors architectures. Inf. Secur. J. Glob. Perspect. **19**(2), 53–60 (2010)
25. Ritter, T.: Transposition cipher with pseudo-random shuffling: the dynamic transposition combiner. Cryptologia **15**(1), 1–17 (1991)
26. GSam Battery Monitor (Google Play Store) (2018). https://play.google.com/store/apps/details?id=com.gsamlabs.bbm&hl=en/

fridaya@uef.fi

Network Protocol Security: A Linear Bitwise Protocol Transformation

Suzanna Schmeelk(✉)

St. John's University, Queens, NY 11439, USA
schmeels@stjohns.edu
http://www.stjohns.edu

Abstract. Companies are investing in network traffic analysis. Understanding protocols for network security is an extremely essential component to maintaining secure networks. Research trends are four fold: on reverse engineering protocols from network traces, on protocol specification, on protocol verification and on application protocol descriptions. There has not been much, if any, work on transforming protocols into full linear representations. Our research is motivated to discover a technique to cleanly understand, define, analyze, symbolize, diffuse, and categorize a set of protocols without requiring any dependencies. Our network protocol transformation research is being used in network security, networking assurance, network testing, and network emulation. New protocols are being defined every day across the globe. A linear transformation technique captures protocol behavior at a symbolic bitwise level without any machine dependencies making it an attractive solution at large to analyze, examine and specify network traffic anomalies.

Keywords: Network security · Protocol representations · Protocol transformation · Protocol analysis

1 Introduction

Network security is an evolving field where malware detection and anomaly detection are critical for maintaining healthy networks. To study, evaluate, analyze, and detect nuances in network activity, all features of a network protocol must be clearly represented for further analysis. Our research found that most current protocol representations are either incomplete (e.g. lacking timers, counters, etc.) or not represented in a ubiquitous or unambiguous format. Therefore, there may be gaps in precise protocol analysis. Traditional protocol research has been within four domains: specification, verification, application-specific and, more recently, reverse engineering. The four traditional protocol research focus has continued on their independent trajectories without much change to the over all strategic research goal. As such, this research contributes a broader protocol research perspective for network assurance purposes.

The goal of this research is to discover how to ubiquitously transform a protocol into a static-linear representation to carefuly analyze the protocol and label

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 751–770, 2019.
https://doi.org/10.1007/978-3-030-22868-2_53

fridaya@uef.fi

potential anomalies. We developed a process using a four-step paradigm. First, the protocol of interest is mathematically transformed into a linear representation. Second, the representation is diffused in *real-time*. Third, the protocol is used for analysis. Finally, the protocol is deleted, categorized upon usage completion, or annotated for further analysis and anomaly detection.

This protocol transformation technique is purposefully intended to capture the entire essence of a protocol using a ubiquitous symbolic representation, and not require extra system processing. The abstract nature of the representation allows network assurances appliances to be able to either use or describe (un)known protocols using range and bit values specifically marking protocol nuances which are of interest from a security perspective. For example, in network assurance, security tools may want to indicate that the counter timer may be off between a certain packet flow or perhaps a bit value on a packet flow is suspicious. The protocol transformation can specifically capture nuances of the protocol in real-time and pass it to security appliances for further analysis.

1.1 Representation Usages

The potential for this research is fivefold. First, it can be used by security appliances to describe variances or change protocols in real time (e.g. Netflow, snmptraps, syslogs, etc.). Second, it can be used to emulate a large testbed of any devise. Third, it can be used to learn and/or describe new protocols in real time. Fourth, the forward- and reverse- implementation techniques can lead to broader cyber security protocol analysis in the core. Fifth, it can be use by network security analysis tools during analysis or in real time to detect anomalies.

Over the years, many research entities and both computer hardware and software computer manufactures have been focused on protocol development. Anderson and Landweber [1] cite the United States Department of Defense (DoD), International Standards Organizations (ISO), IBM and Digital Equipment Corporation (DEC). More recent trends on protocol reverse engineering have been cited around the world such as England [2], Portugal [3], China [4–6], Italy [7], Switzerland [2] and Australia [7], to name a few.

2 Related Work

Traditional protocol research is within four domains: specification, verification, application-specific and, more recently, reverse engineering. The first three genres date back to the 1970s and 1980s when trying to formally describe a protocol and verify that it did not contain deadlock situations or other service interrupting communication flows. The fourth, more modern, protocol research trajectory, which publicly began in the early 2000s, is primarily focused on cyber security issues, which is the focus of our research.

fridaya@uef.fi

2.1 Protocol Specification

Protocol specification has been researched extensively since at least the 1960s.

Early specification was done purely mathematically. The early work of Milne and Milne involved algebra calculi [8]. R. Milner's research involved similar ideas [9]. Hoare [10] wrote "Communicating Sequential Processes (CSP)" which is a language for describing patterns of interaction. The ideas are supported by elegant, mathematical theories, sets of proof tools, and extensive literature. In addition, Karjoth [11] developed a protocol entity object and its refinement. These approaches have been contrasted by Lanese and Montanari in their research [12].

After early mathematical specification was successful, a more formal unified specification technique was required. Many independent specification languages came to be.

Anderson and Landweber [1] introduced the concept of a grammar-based methodology for protocol specification and implementation using real-time asynchronous grammars. Their research was geared towards early UNIX systems. They defined a new methodology to concisely describe real protocols and easy to use. They describe a similar study to this research using grammars where a protocol implementer writes and RTAG description of a protocol. The description is subsequently processed by a grammar analyzer, which produces a set of tables. The protocol plus operating system are integrated to obtain an implementation of the protocol. Finally the parsers does a top-down parse of the grammar in response to external events. This research performed on early UNIX systems, closely matches our research but the focus is different.

David Anderson further developed his earlier research. In this paper, he described an automated protocol implementation using the real-time asynchronous grammar (RTAG). The goal of their formal description technique was to make it easy to specify complex protocols and to allow efficient automated implementation of a protocol using this technique. The technique was based on context-free grammars instead of the traditional finite state machines. The RTAG parser "is a program that, given the RTAG specification of a protocol, implements the protocol [13]." Anderson's technique was geared at early low-level protocols.

Leonini et al. [14] published 'SPLAY: Distributed Systems Evaluation Made Simple or How to Turn Ideas into Live Systems in a Breeze' at USENIX. Their work is geared at actual distributed deployment applications. Their work referenced Anderson, "Other domain-specific languages have been proposed for distributed systems development. In RTAG [13], protocols are specified as a context-free grammar. Incoming messages trigger reduction of the rules, which express the sequence of events allowed by the protocol. Morpheus and Prolac target network protocols development. All these systems share the goal of SPLAY to provide easily readable yet efficient implementations, but are restricted to developing low-level network protocols, while SPLAY targets a broader range of distributed systems."

Sturman and Agha [15] also developed a protocol specification language called Dependability Installation Language (DIL) to enforce failure semantics.

fridaya@uef.fi

Their research, again, is into specifying existing protocols. Their approaches are certainly informative and useful.

Xiang and Jianlin developed a formal description of network protocols using the RAISE specification language [16]. They describe a protocol using six elements: scheduling, protocol data unit, states, events, variables, actions and predications. Their research is used for protocol verification and subsequently implementation and testing.

Hu et al. [17] revisited SDL by modeling and verification custom TCP protocols. Their work is employed in the following sections for protocol elements.

Lastly, Hudert et al. [18] worked on a negotiation protocol description language for automating service level agreement notifications. Their research was motivated by the global and open Internet of Services (IoS). They envision that the future global environment will require *on-the-fly* service level agreements (SLA) and resource allocation.

2.2 Protocol Verification

Protocol specification has been researched by extensively since at least the 1960s. Once a protocol has been specified, verification can take place. Tools including ObjectGEODE have been built to verify protocol safety. For example, protocols that create deadlocks and loss of service are unacceptable for technology. The field of formal methods discusses these ideas in depth.

2.3 Protocol Reverse Engineering

Protocol reverse engineering is two-fold. First, researchers are interested in taking network traces and using traces to discover protocols of interest. Second, researchers interested in Deep Packet Inspection (DPI) and Deep Flow Inspection (DFI), also need a way to determine protocol state in the middle of a network.

Cui et al. [19] proposed *Discoverer* which is a tool for automatically reverse engineering the protocol message formats of an application from its network trace. The research, which was not formally followed-up with, only examined packet structure in their research and did not examine other protocol characteristics such as counters, timers, control structures or finite state machines.

Antunes et al. [3] proposed a new method to automatically infer a specification of a protocol from network traces. They build the ReverseX tool to reverse engineer FTP traces. Their technique involves two algorithms. The first algorithm infers the message format. The second algorithm infers the state machine. Other researchers working on similar research include Caballero and Song [20] and Bossert et al. [21]. This technique is being used, in part, for cyber security related research on fraud detection and fuzzing.

Meiners et al. [22] developed FlowSifter which is a systematic framework for only application protocol field extraction. Their research was geared at layer extraction devises including data loss prevention (DLP), SOAP, XML-RPC, SNORT and BRO. This extraction occurs at the router-level and is content unaware.

fridaya@uef.fi

2.4 Other Protocol Analyses

A United States patent in 2013, entitled, "Dynamic Generation of User Interface" [23] described a methodology to generate user interfaces *on-the-fly*. This patent describes how to take stored information and dynamically create a user interface for the information but does not describe how to formally create a communication protocol dynamically.

Another research paper, entitled "Automated Protocol Analysis" [24] studied human to computer interaction protocols. Their research introduced the idea of protocol replay after formal specification which is closely related to our research motivation.

3 Transforming Key Protocol Elements into Linear Bitwise Representations

Protocols traditionally have come in three strands: bits, bytes, and text. Bits and byte protocols predominately divide-up packet formation based on bits and bytes; while, a text based protocol (e.g. HTTP) divides-up packet formation based on text and character identification. Our linear bitwise technique represents all three categories. In transit in the core, packet bits of a protocol are nothing more than an indication of an on/off value. Each end of the communication needs to reassemble the on/off values and 'make-sense' of the received signals. The beauty of this kind of a core-based structure is referred to as 'complexity-at-the-edge' where machines are, for the most part, independent and isolated of any other machine. Our linear bitwise protocol implementation capitalizes on the Internet architecture to achieve a very substantial goal.

In our research, we experimented heavily with paper-based studies, RFC studies, Wireshark and application communications on virtual machines. From the research, we identified at least seventeen elements required for dynamic protocol generation. These seventeen features are protocol dependent. Some protocols do not support the features. For example, integrity checking, authentication checking, state diagrams, counters, time intervals and control structures vary across protocols.

After trying different ways to express a protocol structure and behavior remotely, we selected the eXtensible Markup Language (XML) for its ubiquitous and dynamic structure. The structure is detailed and dynamic enough to express what needs to be expressed for remote deployment. XML is standardized and can be used to express all Internet layer protocols.

3.1 Directional

Protocols can be uni- or multi- directional. Simple protocols including UDP are ubiquitous and unidirectional. Simple protocols usually offer no guarantees but are usually considered fast. Protocols are grouped into *one* up to n directional levels. In an n directional scenario, broadcasting is a minimal scenario. Bui et al. [25] discussed the multi-casting scenario when describing a protocol.

fridaya@uef.fi

3.2 Protocol Stacks

Protocols have been developed to modularize communication. They are, thus, designed to be used in a protocol-stack environment where each layer can depend on a lower layer for certain features, as defined by the Open Service Interconnect (OSI). We envision a technique where higher-levels *register* with lower layers.

3.3 Header Format

During communication, protocols send packets to convey their messages. Each packet sent by the protocol has a special header to identify who the message is coming from and where the message is going. Traditionally, people associate protocol headers with United States Post Office mail addressing; however, the protocol headers are completely electronic. Also, headers may indicate specialized information to the corresponding entities through the use of setting flags.

Due to the nature of header diversity, a header must be specified in a bit-wise methodology. We suggest using numbers delineated by *:* to make parsing easy. In addition, we suggest adding totals to all fields to ensure that proper configuration. The *header bit total* will be used by machine to ensure that the bit specification is indeed *mathematically* correct.

```
<protocol id=[x] layer=[x]
           name=["Protocol Name"]
     directions=[1....B]
     layer_below=[x1_id, x2_id, ...] >
     <header header_bit_specification=
             ["8:1:1:2:3:....:4:4:6:24:1"]
             header_bit_total=[x]>
     </header>
    ...
</protocol>
```

3.4 Endianness

The terms endian and endianness describe the convention used to interpret the bytes making up a data word when those bytes are stored in computer memory. In general, memory stores binary data by organizing it into 8-bit units called bytes. When reading or writing a data word, the order of the bytes stored in memory determines the interpretation of the data word [26].

Each byte of data in memory has its own address. Big-endian systems store the most significant byte of a word in the smallest address and the least significant byte is stored in the largest address (also see Most significant bit). Little-endian systems, in contrast, store the least significant byte in the smallest address [26].

Due to the flexibility of various protocols, we wanted to grant the flexibility for each bit to be either left or right aligned. In practice, we expect the bits to be aligned completely to the left or completely to the right; however, we would like

fridaya@uef.fi

to make our static representation as flexible as possible. We envision including an alignment representation with each bit in the *header bit specification*. This can allow for nuances between proprietary, as well as, RFC standards. Using this representations, the header bit specification is defined using an *array*-likeness. Any subsequent reference to the header bit specification assumes the same *slots* in the array. For example, header bit specification of the first element, *HeaderBitSpecification[0]* would represent the value "8R" and, so forth until, the final array-slot, the *HeaderBitSpecification[HeaderBitSpecification.length()-1]* would represent the value "1R."

3.5 Encoding

Each protocol has a footprint that may be the result of special characters and encoding schemas. We want a broad way to represent using special characters and specific encoding schema when developing both the packet header and packet data unit. This topic needs a horizontal research analysis amongst protocols to learn the most popular encoding across protocols. We envision the following for representing encoding amongst the packet header. First, we envision an *encoding* attribute and *special character* attribute. The encoding is a generalize encoding type and the special characters represent changing certain characters for a specific protocol.

```
<protocol id=[x] layer=[x]
           name=["Protocol Name"]
     directions=[1....B]
  layer_below=[x1_id, x2_id, ...] >
  <header header_bit_specification=
     ["8R:1L:1R:2L:3R:....:4R:4R:6L:24L:1R"]
       header_bit_total=[x] encoding="HEX"
       special_characters="0->00:NULL->^M"  >
   </header>
     ...
</protocol>
```

3.6 Finite State Machine (FSM)

Protocols can be represented using a finite state machine. Traditionally, protocols are verified, as described in Sect. 2. We can use this feature to our advantage. At any given time during a communication via a protocol, the protocol is considered to be in a certain state. The FSM can be represented using a directed graph, as can be seen in the custom TCP Finite State Machine research of Hu et al. [17]. A small 10 node directed graph takes a relatively small amount of memory. Each edge of a graph can represent events and counters that should occur. We can label this state transition with a unique identifier, *fid*. In addition, each edge can define what happens on the event of failure. We encourage

fridaya@uef.fi

specifying the exact number of nodes and edges to verify during parsing that the protocol is written as expected.

It should be noted that each *node* and *edge* pair is defined. This pair is represented by an arrow. In addition, each edge may be associated with a counter or timer, a failure scenario, or a control structure that needs to be examined. These associations will be examined upon state transition.

```
...
<fsm id=[x] nodes=[x] edges=[x]>
 <fid=[x] node id=[x] edge=[node_id]
     counter=[counter_id1, ..., counter_idn]
     failure=[failure_id1, ..., failure_idn]
     control_structure[cs_id1, ..., cs_idn]>
 <fid=[x] node id=[x1] edge=[node_id]
     counter=[counter_id1, ..., counter_idn]
     failure=[failure_id1, ..., failure_idn]
     control_structure[cs_id1, ..., cs_idn]>
   ....
</fsm>
...
```

3.7 Message Types

Each state transition in a FSM changes FSM values. In many cases, the state transition will actually change the message headers and in other FSM transitions only timers and payload will change. We need a linear representation that can model either situation. We envision a *packet* XML statement to define each packet. This packet statement will have two attributes to associate them with the correct transition in the FSM, a *fsm id* and a *fid* within the FSM. We cannot envision a situation where there may be multiple finite state machines, but this should be further researched.

3.8 Body Format

The body of a packet depends on the state of the FSM. The body is actually used to convey data from one entity to another over the protocol. As we could develop protocols *on the fly* in some cases, the actual application may or may not exist on the system. A packet can either be sent or received. We make this association in the queue section.

We envision four different scenarios. First, data payload can be statically defined during protocol specification directly into the protocol XML. Second, data payload can be connected with an XML-RPC system where payload is actually further parsed or retrieved from remote procedure calls. In this case, an interface with an RPC system must take-place. Third, the application may, in fact, exist on the machine. In the third case, an actual association with an application must me made. Fourth, the packet may just be sending empty content or content that needs to be aggregated such as a video.

fridaya@uef.fi

3.9 Executing Payload

Executing protocol payload depends on the type defined by each packet as seen in the above section. If there is static payload, such as in the case of a FIN + ACK, there is no execution of payload. If the application exists on the system, then the application should be executed with any known parameters and the results conveyed back to the remote entity following the FSM. If the application exits on a remote server, an RPC method can be invoked to information retrieval. These cases can be tailored for test environments.

3.10 Queuing

There can be multiple queues for a protocol. First, packets are passed to the protocol, perhaps asynchronously due to network delays. Packets need to be buffered or queued as they arrive. We will expect two buffers by default, a *send id = 0* and a *receive id = 1*. These default buffers will not need to be specified.

Second packets need to align correctly with the FSM and Control Structures. Any additional queues for this alignment will need specifications. A naive approach is to replicate the packet header for each control structure. We envision future research to eliminate packet header duplications.

```
<protocol id=[x] layer=[x]
     name=["Protocol Name"]
     directions=[1....B]
     layer_below=[x1_id, x2_id, ...] >
     <queue id=[x] default_queue_id=["0 or 1"]
        cs_ids=[1,...,n]>
     </queue>
     ...
</protocol>
```

3.11 Ordering Messages

Ordering of messages occurs based on the FSM and the control structures and can be labeled using our linear bitwise protocol transformation.

3.12 Control Structures

Control structures can take on many forms. Hu et al. [17] defined this for TCP specification. There may be additional control structures, but they can be linear represented in a similar format. In general, a control structures ensure packets are connected correctly and that they do not over flow the network and remote entity. Additionally, control structures can be triggered on the event of worst-case scenarios.

fridaya@uef.fi

```
<protocol id=[x] layer=[x]
    name=["Protocol Name"]
   directions=[1....B]
   layer_below=[x1_id, x2_id, ...] >

  <control_structures id=[x] fsm=[x] fid=[x]>
   <cs_action id=[x]
     start=["0"]
     action=["+=parm1"]
     parm1=["header_bit_specification[10]"]
     end_fsm=[x] end_fid=[x]
   </cs_action>
       ...
   <cs_action id=[n]
     start=["0"]
     action=["+=parm1"]
     parm1=["header_bit_specification[10]"]
     end_fsm=[x] end_fid=[x]
   </cs_action>
  </control_structure>
     ...
</protocol>
```

Sequencing. Sequencing ensures that packets are queued in an ascending order (recognizing the fact that when the number gets too big it returns to 0). Sequencing can be defined within a control structure with algorithms. Remember, that all arriving and departing packets are checked with the FSM and control sequences, if any, upon arrival and departure.

Reliable Transmissions. Reliable transmissions ensure that the packets are retransmitted when a timer expires. Reliable transmissions can be defined within a control structure using algorithms. All arriving and departing packets are checked with the FSM and control sequences, if any, upon arrival and departure.

Link Control. Link control ensures that packets are sent following some given time. Link control can be defined within a control structure using algorithms. All arriving and departing packets are checked with the FSM and control sequences, if any, upon arrival and departure.

Flow Control. Flow control ensures that packets are not overwhelming the receiver. Flow control can be defined within a control structure using algorithms. All arriving and departing packets are checked with the FSM and control sequences, if any, upon arrival and departure.

fridaya@uef.fi

Congestion Control. Congestion control ensures that packets are not overwhelming the network. Congestion control can be defined within a control structure using algorithms. All arriving and departing packets are checked with the FSM and control sequences, if any, upon arrival and departure.

Failure Scenarios. Failure Scenarios ensures actions that should occur when worst-case situations transpire. Failure scenarios can be defined within a control structure using algorithms. As with the other controls, all arriving and departing packets are checked with the FSM and control sequences, if any, upon arrival and departure.

3.13 Counters

A FSM and related control structures contain many counters. We envision defining all in XML using an *action* attribute defining a parsable stack-based mathematical expression.

```
<protocol id=[x] layer=[x]
     name=["Protocol Name"]
     directions=[1....B]
     layer_below=[x1_id, x2_id, ...] >
     <counter id=[x] fsm=[x] fid=[x]
       start=["0"]
       action=["parm1+parm1"]
       parm1=["header_bit_specification[10]"]
     </counter>
     ...
     <counter id=[x] fsm=[x] fid=[x]
       start=["header_bit_specification[10]"]
       action=["++"]
     </counter>
...
</protocol>
```

3.14 Time Intervals

A FSM and related control structures contains many timers. Timers are used to trigger events. Timers are defined with a *start time*, *end time* and *action*. We envision a FSM edge to exist for expired timers which will become the *action* if the timer completes. If the FSM changes before the timer expires, the timer is ignored.

```
<protocol id=[x] layer=[x]
     name=["Protocol Name"]
     directions=[1....B]
```

fridaya@uef.fi

```
     layer_below=[x1_id, x2_id, ...] >
     <timer id=[x] fsm=[x] fid=[x]
       start=["arrival"]
       end=["30s"]
       action=["fid"]
       parm1=["header_bit_specification[1]"]
     </timer>
     ...
     <timer id=[x] fsm=[x] fid=[x]
       start=["departure"]
       end=["30s"]
       action=["fid"]
     </timer>
...
</protocol>
```

3.15 Integrity

Message integrity is used throughout protocol analysis and calculations. In packets to be sent, the integrity and checksums should be calculated using various standardized algorithms. The algorithms can be defined by the testing application. The protocol can then reference these integrity algorithms. There are two cases. First, the integrity field needs to be created for sending. Second, the integrity field need to be verified when received.

```
...
   <packet id=[1] fid=[x] fsmid=[fid_n]>
     <header header_bit_specification=
       ["8R:1L:1R:2L:3R:....:4R:4R:6L:24L:1R"]
        header_bit_total=[x] encoding="HEX"
       special_characters="0->00:NULL->^M"
       directional_queue=["0"]
       integrity=[integrity_id] bit_slot="5"
             bit_range="">
      </header>
      <body body_type = [4]
        body_bit_specification="250L"
        body_bit_total=[x] encoding="HEX"
        special_characters="0->00:NULL->^M"
        store_location=["/home/temp"]
        store_parameters =[parm1,....,parmn]
        integrity=[integrity_id] bit_range="all">
      </body>
    </packet>
    <packet id=[n] fid=[x] fsmid=[fid_n]>
     <header header_bit_specification=
```

fridaya@uef.fi

```
        ["8R:1L:1R:2L:3R:....:4R:4R:6L:24L:1R"]
         header_bit_total=[x] encoding="HEX"
        special_characters="0->00:NULL->^M"
        directional_queue=["1"]
        integrity=[integrity_id] bit_slot=""
              bit_range="all">
      </header>
      <body body_type = [4]
        body_bit_specification="250L"
        body_bit_total=[x] encoding="HEX"
        special_characters="0->00:NULL->^M"
        store_location=["/home/temp"]
        store_parameters =[parm1,....,parmn]
        integrity=[integrity_id] bit_range="all">
      </body>
    </packet>
    <integrity id=[1] calculation=["MD5"]>
     ...
    <integrity id=[n] calculation=["SHA"]>
...
```

3.16 Authentication

Authentication may also be considered. In many cases, testing cannot occur without proper authentication. As such, we developed XML which basically defines a control structure. All calculations, other than counters, should occur within a control structure. Authentication is quite diverse across protocols and requires more research to accommodate most techniques.

3.17 Topology

The topology of a network may be important to certain protocols. For example, many clients and servers require a centralized server to handle all correspondences. In such cases, centralized servers must be established for proper usage of *out-of-the-box* software. Home-grown software may be easier to reconfigure. For example, our case studies below shows the P2P systems do not always correspond directly to a linear representation mapping using peer-to-peer.

3.18 System Constraints

System constraints should be considered with implementing protocols. For example, certain protocols will require higher *bandwidth*, *faster queuing*, *memory* and *cpu time*. These consideration should be considered when designing a system to support pervasive protocol testing. We can define these parameters in XML and the system can then decide if it can handle the capacity.

fridaya@uef.fi

```
<protocol id=[x] layer=[x]
     name=["Protocol Name"]
     directions=[1....B]
     layer_below=[x1_id, x2_id, ...] >
     <system_constraints id=[x]
       cpu=[x]
       dl_queue_speed=[x]
       ul_queue_speed=[x]
      dl_bandwidth=[x]
      ul_bandwidth=[x]
     </system_constraints>
...
</protocol>
```

3.19 Summary

In summary, the beauty of using XML to specify protocols is the breadth of usage. First, many different protocols can be specified, used, diffused, deployed, used and analyzed using this technique. Second, parts of a protocol can be defined and used without defining the complete protocol. Third, reusable libraries can be established and used indefinitely without major modification.

4 Architecture

After many experiments with RFCs, live application clients and servers, live network traffic and conference/university research papers, the pervasive protocol testing framework has been designed for convenience, adaptability and robustness. The architecture has developed from an idea to handwritten linear bitwise representations to actual empirical data collected over multiple experiments and iterations.

Convenience is a fundamental element to to the testing framework. The framework must be easy and timeless enough to use so that developers and testers can learn to use it quickly and without much troubleshooting. Knowing this, we have designed an architecture with both forward and reverse transformation features. In addition, once a protocol has been developed, it should be added to a protocol library for easy retrieval, reuse, and versions.

Adaptability is a key element to the testing framework. First, the framework must be robust enough to last a multi-year experiment without server reinitialization. A dynamic loading for a very-long run is quite feasible. The architecture must support these requirements.

Robustness is a third key element to the testing architecture. The architecture must be robust enough to handle nearly every possible protocol without restructuring the linear bitwise representation. The architecture must also be fluid enough to support many types of application payloads. Payloads can range from static to operating system customizable.

fridaya@uef.fi

4.1 Protocol Transformation

Protocol transformation can occur in two directions either cumbersomely using *forward* transformation and/or efficiently using *reverse* engineering. Theses findings lead to significant objectives. In fact, using both forward and reverse transformations broadens the usage from testing activity to cyber security protocol analysis (e.g. fuzzing, analysis done inside firewalls, and intrusion detection systems).

Forward Transformation. Forward transformation takes a RFC and creates, by hand, the required XML format to specify the protocol. The transformation can be automated with GUIs and perhaps a RFC parser. The research showed that this direction is useful only under two real circumstances. First, it is useful when there are no live packet captures available. And, second, it is very useful for modifying an existing transformed protocol for the purposes of fuzzing, etc.

Reverse Transformation. The research identified the strength of the reverse transformation. During the research we began to discovered how powerful Wireshark is at already transforming bits from the wire into a visual representation. We identified Wireshark as a powerful tool that *already* knows the values of many protocol features. Wireshark plugins can wrap the Wireshark data into an extensible format such as our previously discussed the transformation language. The selection of Wireshark to assist with the linear transformation broadens and speeds-up initial development.

Parts of the protocol FSM can be generated from network traffic. However, using network traffic is subject to the construction of FSMs with false negatives. In another words, network traffic may not exhibit certain aspects of the protocol. We found that protocols that have been specified and verified could be used to reverse engineer parts of the protocol which are unlikely to occur. The actual network traffic may only expose a piece of the entire protocol. These niche corners of the protocol are useful for security researchers to understand all aspects of the protocol without digging through RFCs.

4.2 Protocol Client/Server Framework

If the protocol were intended for replay across the network, the client-side framework must be able to interface with traffic at each layer. It turns out that crafting a client into separating communications at levels has high standard deviation. There are cases where a small client only needs to be installed and there are cases where further client installation needs to transpire.

After examining several operating systems, we noticed that although operating systems follow the layered OSI model, their code can either be fully modular in the best case; or, operating system code can be non-modular in the worse-case. Our client framework needs a slick layered approach where each operating system layer is hooked with our client. In cases were only the application layer

fridaya@uef.fi

is under test, then only an application layer client needs to be deployed which is consistent across operating systems.

5 Empirical Evaluation

We evaluated the correctness of a linear bitwise protocol transformation by replaying transformed protocols and examining how Wireshark and Snort reacted. Our research lead us through multiple paper studies of protocols to envision specifying how they should be linear transformed. We also worked on many case studies. Four of the most important studies where serious impacts were found for our linear representation are described below. The studies were somewhat cyclical in nature, but we describe the last studies of significance: RFCs, Client-Servers, Chat Applications, and Wireshark/Snort.

5.1 Case Study 1: RFCs

The first notion, bitwise, developed from the studies. A broad usage for a representation requires each bit specified in a packet ensure agility. Additionally, as we studied the various OSI levels, the extent of in the wild protocols and rich data is very broad. For example, via the ISO website we found a list of service names to ports where the list contained 13,636 row entries. A second concept we noticed from researching well-known protocols RFC to decipher the various packet headers and protocol formation was extremely time consuming. As we explored marking-up protocols and re-reading research in fields, we discovered that the reverse direction is substantially quicker. This finding was significant. We first discovered, from these case studies, that we can capitalize on all the open source work devoted to Wireshark ofter that past fifteen years.

5.2 Case Study 2: Client-Servers

This forward-direction time consumption lingered with us as we implemented clients and servers on Ubuntu. We deployed multiple clients and servers on Wireshark and examined their packet captures. The clients and servers, were written in Java. We found that we could, in fact, send user defined legitimate application-layer traffic between remote endpoints. Wireshark decoded the traffic correctly and it passed the snort test. Snort is an intrusion detection tool. If it does not identify traffic as malicious, then the traffic is considered authentic by most network security appliances.

5.3 Case Study 3: Chat Applications

During this research, we selected chat application for analysis and their respective RFCs. Our first choice was the Yahoo! instant messenger application (YMSG). For the research, we set-up two different virtual machine since the application can only run one client per machine. An interesting finding that came

fridaya@uef.fi

from the empirical research analysis is that many chat applications are *not* P2P. This finding showed us that we could not use out-of-the-box clients from Yahoo! for a fine grained analysis in this particular research. In fact, out-of-the-box YMSG chat clients will not run unless they authenticate via the Yahoo! server. Other non-proprietary chat applications are still, in fact, extremely attractive for communication tests given a small test environment and more time.

5.4 Case Study 4: SMTP Protocol

We developed a linear bitwise representation of the *Simple Mail Transfer Protocol (SMTP)*. We loaded the protocol onto a server-fitted machine and interacted with the instantiated SMTP framework. The linear-representation instantiated the correct packets using a control flow. The traffic was correctly identified by Wireshark; and, the initial tests passed through Snort indicating that the linear representation of the protocol was indeed correct.

5.5 Case Study 5: Wireshark and Snort

If Wireshark and Snort distinguishes traffic of a particular nature as legitimate, then the communication endpoints (and core) will most likely consider the communication authentic and the protocol has been successful transformed. This lead to the research of pcaps off the Wireshark website with successful results. First, we pulled down seventeen PCAPs filled with the protocol traffic from: Airtunes, BGP, BMP, DNS, Gopher, HTTP, InfiniBand, JXTA, Kasperky, MySQL, NNTP, NTP, SCTP, SoulSeekRoom, STP, Telnet and TFTP. In each capture, we examined the protocol headers. We used wirehark to pull-out the specific application-layer headers and data. We then sent the data between a client and server of our own choosing. We successfully used generic TCP and UDP clients and servers to pass an application-layer packet across of our own choosing; and, the traffic was correctly identified by Wireshark.

6 Research Implications and Future Work

Our cases studies show the initial success of our technique. Due to the notion of 'complexity-at-the-edges', our technique can be used to analyze and construct communications between endpoints. The beauty of the research is the breadth of use cases and customer markets that would be interested in this research.

Our findings link three diverse domains together–security, network assurance and testing–via a mathematical bitwise linear transformation of a protocol. It should also be noted that there are tens-of-thousands of protocols that can be impacted by this research. The breadth is the main attraction since unclassified and classified market-segments span all of the protocol domains.

fridaya@uef.fi

6.1 Protocol Testing Framework

Our technique can be used for testing protocols in all network environments. Since we are developing a framework for both forward- and reverse- transformations of protocols, we expand the number of protocols we can examine and accuracy of our findings.

6.2 Cyber Security Implications

Our technique can be used to analyze protocols in the cyber security domain for fraud, service faults, heuristics and network traffic analysis. The reverse transformations of protocols from network traces seriously expands the research that can be done in this domain. In this domain, fuzzing can transpire, protocol behavioral analysis can occur leading to protocol verification and non-denial of service testing (note verification occurs from the inside-out not the outside-in). Finally, other statistics can be examined that are used by intrusion detection systems (IDS), firewalls, and other fraud detection devices.

7 Conclusion

The goal of this research was to discover a technique to ubiquitously transform a protocol into a static-linear representation to analyze and employ in real time. The research is to enable cyber security initiatives including machine learning of unknown protocols, firewall limitations, intrusion detection system analysis, deep packet inspection and fraud detection systems. This is preliminary research to broadly and deeply consider transforming every protocol feature–including finite state machines, counters, buffers, timers, control structures, authentication among other protocol concerns. These findings broaden the research implications to include many network security domains for anomaly detection as well as network testing.

References

1. Anderson, D.P., and Landweber, L.H.: A grammar-based methodology for protocol specification and implementation. In: Proceedings of the Ninth Symposium on Data Communications, New York, NY, USA, SIGCOMM 1985, pp. 63–70. ACM (1985)
2. Basin, D., Cremers, C.: Know your enemy: compromising adversaries in protocol analysis. ACM Trans. Inf. Syst. Secur. **17**(2), 7:1–7:31 (2014)
3. Antunes, J., Neves, N., Verissimo, P.: Reverse engineering of protocols from network traces. In: Proceedings of the 2011 18th Working Conference on Reverse Engineering, Washington, DC, USA, WCRE 2011, pp. 169–178. IEEE Computer Society (2011)
4. Jiang, L., Yang, X., Li, T.: The analysis and design for a network protocol analysis system based on winpcap. In: Communications Security Conference (CSC 2014), pp. 1–4, May 2014

fridaya@uef.fi

5. Hai, L., Yue-Yue, Y., Chang, L., Li-Na, C.: A protocol analysis platform based on open source software. In: 2012 Second International Conference on Instrumentation, Measurement, Computer, Communication and Control (IMCCC), pp. 846–849, December 2012
6. Li, L., Pang, J., Liu, Y., Sun, J., Dong, J.S.: Symbolic analysis of an electric vehicle charging protocol. In: 2014 19th International Conference on Engineering of Complex Computer Systems (ICECCS), pp. 11–18, August 2014
7. Ponge, J., Benatallah, B., Casati, F., Toumani, F.: Analysis and applications of timed service protocols. ACM Trans. Softw. Eng. Methodol. **19**(4), 11:1–11:38 (2010)
8. Milne, G., Milner, R.: Concurrent processes and their syntax. J. ACM **26**(2), 302–321 (1979)
9. Milner, R.: A Calculus of Communicating Systems. Springer, New York; Secaucus (1982)
10. Hoare, C.A.R.: Communicatng Sequential Process. Prentice Hall, London (1985)
11. Karjoth, G.: Stepwise specification of a sliding-window protocol by means of process algebra. In: 1988 International Zurich Seminar on Digital Communications, 1988. Mapping New Applications onto New Technologies, pp. 109–114, March 1988
12. Lanese, I., Montanari, U.: Hoare vs milner: comparing synchronizations in a graphical framework with mobility. Electron. Notes Theor. Comput. Sci. **154**(2), 55–72 (2006)
13. Anderson, D.P.: Automated protocol implementation with RTAG. IEEE Trans. Softw. Eng. **14**(3), 291–300 (1988)
14. Leonini, L., Rivière, E., Felber, P.: Splay: distributed systems evaluation made simple (or how to turn ideas into live systems in a breeze). In: Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation, Berkeley, CA, USA, NSDI 2009, pp. 185–198. USENIX Association (2009)
15. Sturman, D., Agha, G.: A protocol description language for customizing failure semantics. In: 13th Symposium on Reliable Distributed Systems, 1994. Proceedings, pp. 148–157, October 1994
16. Xiang, G., Jianlin, Q.: Formal description of network protocols using raise specification language. In: 2011 International Conference on Network Computing and Information Security (NCIS), vol. 1, pp. 185–188, May 2011
17. Hu, K., Liu, C., Liu, K.: Modeling and verification of custom TCP using SDL. In: 2013 4th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 455–458, May 2013
18. Hudert, S., Eymann, T., Ludwig, H., Wirtz, G.: A negotiation protocol description language for automated service level agreement negotiations. In: IEEE Conference on Commerce and Enterprise Computing, 2009. CEC 2009, pp. 162–169, July 2009
19. Cui, W., Kannan, J., Wang, H.J.: Discoverer: automatic protocol reverse engineering from network traces. In: Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium, Berkeley, CA, USA, SS 2007, pp. 14:1–14:14. USENIX Association (2007)
20. Caballero, J., Song, D.: Automatic protocol reverse-engineering: message format extraction and field semantics inference. Comput. Netw. **57**(2), 451–474 (2013)
21. Bossert, G., Guihéry, F., Hiet, G.: Towards automated protocol reverse engineering using semantic information. In: Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security, New York, NY, USA, ASIA CCS 2014, pp. 51–62. ACM (2014)

fridaya@uef.fi

22. Meiners, C., Norige, E., Liu, A., Torng, E.: FlowSifter: a counting automata approach to layer 7 field extraction for deep flow inspection. In: 2012 Proceedings IEEE INFOCOM, pp. 1746–1754, March 2012
23. Zhu, J.Y., Cai, X., Zang, Y., Cao, L.: Dynamic generation of user interface, March 2015
24. Smith, J.B., Smith, D.K., Kupstas, E.: Automated protocol analysis. Hum.-Comput. Interact. **8**(2), 101–145 (1993)
25. Bui, T., Elshabrawy, T., Khendek, F., Tahar, S., Le-Ngoc, T.: Description and validation of the internet stream protocol (st2+) using SDL/MSC. In: 1999 IEEE Canadian Conference on Electrical and Computer Engineering, vol. 1, pp. 238–244, May 1999
26. Wikipedia: Endianes—Wikipedia, the free encyclopedia, 2015. Accessed 1 Jan 2019

fridaya@uef.fi

Mesh-Based Encryption Technique Augmented with Effective Masking and Distortion Operations

Muhammed Jassem Al-Muhammed[1(✉)] and Raed Abuzitar[2]

[1] Faculty of Information Technology, American University of Madaba, Madaba, Jordan
m.almuhammed@aum.edu.jo

[2] College of Information Technology, Ajman University, Ajman, UAE

Abstract. One of the main distinctive contributions of the digital era is that almost all of our information are either digitally stored on machines or transmitted over the networks. This information will be in a great danger unless properly secured. This paper addresses this problem by offering an encryption technique. The proposed technique possesses several properties that make it effective. First, it proposes an effective substitution process that highly eliminates the relation between the result of the substitution and the original plaintext. Second, it offers diffusion techniques that highly increase the avalanche effect. Third, it intelligently inserts random noises in the encryption to greatly increase the ciphertext confusion. Experiments with our prototype implementation showed that the proposed technique is effective and passed essential security tests.

Keywords: Encryption technique · Mesh-based mapping · Block encoding · Random noise effect · Mesh

1 Introduction

Ensuring the security of digitally stored or network-transmitted information is certainly the Holy Grail that all information owners seek. There is a universal consensus that encryption is the de-facto mechanism for securing information. As a logical response, many encryption techniques have been proposed (e.g. [1–13]). These techniques use different processing models (e.g. stream or block based) and different diffusion and confusion operations to encrypt plaintexts using keys.

An attractive encryption technique is proposed in [1]. The underlying principle of this technique is a mesh-based mapping operation that substitutes every plaintext symbol with a directive. The directive, which is discussed later in great details, is a signed integer. The technique [1] includes also other operations for masking the directives. This encryption technique is effective and passed important security tests. Although substituting a symbol with a directive is extremely

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 771–796, 2019.
https://doi.org/10.1007/978-3-030-22868-2_54

fridaya@uef.fi

interesting, the mapping logic itself has a significant drawback. It is quasi-linear, which makes it somewhat vulnerable and may be exploited by hackers as it does not provide enough confusion.

This paper takes the technique [1] steps further by proposing fundamental extensions. In particular, it significantly improves the mesh-based mapping by making this mapping non-linear. This non-linear mapping is much more effective and highly increases the confusion compared to the mapping in [1]. Besides the new mapping, this paper proposes more effective diffusion operations that effectively propagate changes in plaintext or keys to all of the output symbols. To this end, the paper offers the following contributions: (1) highly effective mapping mechanism that overcomes the quasi-linearity of the mapping in [1], (2) effective diffusion technique that considerably increases the avalanche effect, (3) effective masking operations that greatly increase the confusion in the resulting cipher text, (4) a very effective key expansion method that combines both the enhanced substitution mechanism and distortion operations, and (5) a key-based random number generator that embeds random noises in the encryption process.

2 Substitution and Inverse-Substitution Methods

In general, substituting a block is the process of replacing the symbols of the block with new ones. The substitution operation uses the S-Box [5] to perform the replacement. The S-Box as described in [5] is a 16 × 16 lookup table with 256 entries that represent all the possible combinations of the byte (8 bits). The substitution for a symbol as described in [5,11] is performed by dividing the 8 bits that represent the symbol into two halves (4 bits each). The left 4 bits index the row of the table and the right 4 bits index the column. The indexed value is the substitute for the symbol. To obtain the original symbols, S-Box has a corresponding inverse table called Inverse S-Box. Substituting a symbol using Inverse S-Box is identical to substituting a symbol using S-Box.

2.1 The Substitute Method

The proposed **Substitute** method uses the same S-Box [5], but offers a fundamentally different substitution mechanism. First, our substitution method uses the outcome of the most recently performed substitution to determine the substitution of the current symbol. Second, our substitution operation consists of two substitution actions: Forward and Backward.

The substitute method is described by Fig. 1. It performs double substitutions on the input: Forward Substitute and Backward Substitute. The Forward Substitution takes the input block $k_1k_2...k_n$. and returns the substituted block T $(\mathcal{T}_1\mathcal{T}_2...\mathcal{T}_n)$. Referring to Fig. 1, the Forward Substitute substitutes the first symbol k_1 with a new symbol $(\mathcal{T}_1)$ using the S-Box. For the remaining input symbols k_i $(i = 2, 3, ..., n)$, the Forward Substitute XORes the most recent substitution outcome $\mathcal{T}_{(i-1)}$ with the to-be-substituted symbol k_i. The outcome of the XOR operation, which is the symbol $K_i^\star$ $(K_i^\star = \mathcal{T}_{i-1} \bigoplus k_i)$, is substituted using the S-Box to produce the new symbol $\mathcal{T}_i$ in the output block T.

fridaya@uef.fi

Forward Substitute
Input: $k_1 k_2 \ldots k_n$
Output: substituted block T
$\mathcal{T}_1$= Substitute (k_1)
$T = T + \mathcal{T}_1$
FOR i = 2 to n **DO**
$K_i^{\star} = \mathcal{T}_{i-1} \oplus K_i$
$\mathcal{T}_i$ = Substitute$(K_i^{\star})$
$T = T + \mathcal{T}_i$
RETURN T

Backward Substitute
Input: $\mathcal{T}_1 \mathcal{T}_2 \ldots \mathcal{T}_n$
Output: substituted block $\mathcal{W}$
w_n= Substitute $(\mathcal{T}_n)$
$\mathcal{W} = \mathcal{W} + w_n$
FOR i = n-1 to 1 **DO**
$K_i^{\star} = w_{i+1} \oplus \mathcal{T}_i$
w_i = Substitute$(K_i^{\star})$
$\mathcal{W} = w_i + \mathcal{W}$
RETURN $\mathcal{W}$

Fig. 1. The Substitution operation.

Backward Inv-Substitute
Input: $w_1 w_2, \ldots, w_n$
Output: substituted block T
$\mathcal{T}_n$= Substitute (w_n)
$T = \mathcal{T}_n + T$
FOR i = n-1 to 1 **DO**
t_i = Substitute(w_i)
$\mathcal{T}_i = t_i \oplus w_{i+1}$
$T = \mathcal{T}_i + T$
RETURN T

Forward Inv-Substitute
Input: $\mathcal{T}_1 \mathcal{T}_2 \ldots \mathcal{T}_n$
Output: substituted block K
k_1= Substitute $(\mathcal{T}_1)$
$K = K + k_1$
FOR i = 2 to n **DO**
S_i = Substitute$(\mathcal{T}_i)$
$k_i = S_i \oplus \mathcal{T}_{i-1}$
$K = K + k_i$
RETURN K

Fig. 2. The Inverse-Substitution operation.

The output of the Forward Substitute T $(\mathcal{T}_1 \mathcal{T}_2 \ldots \mathcal{T}_n)$ is passed as an input to the Backward Substitute. The Backward Substitute has the same functionality as the Forward except that it starts from the end of the input. Therefore, as Fig. 1 shows, the Backward Substitute substitutes the last symbol T_n in the input first with new symbol (ω_n) using the S-Box. For the remaining input symbols $\mathcal{T}_i$ $(i = n–1, n–2, \ldots,1)$, the Backward Substitute first XORes this symbol $(\mathcal{T}_i)$ with the most recent substitution outcome (w_{i+1}). The outcome of the XOR operation $(K_i^{\star} = \omega_{i+1} \bigoplus \mathcal{T}_i)$ is substituted to produce the new symbol ω_i in the output block $\mathcal{W}$.

2.2 The Inverse-Substitute Method

The **Inverse-Substitute** method reverses the impact of the **Substitute** method. The functionality of the Forward Inverse-Substitution and the Backward Inverse-Substitution are defined in Fig. 2. The Backward Inv-Substitute executes first then the Forward operation. Furthermore, the Inverse-Substitute uses the Inverse S-Box (rather S-Box) when substituting a symbol.

The Backward Inv-Substitute starts from the end of the input. Therefore, it substitutes the last input symbol w_n using the inverse S-Box to yield the

fridaya@uef.fi

$$Manipulate(t_i) = \begin{cases} \overset{LH}{Flip}(t_i) & if\ t_i \le t_{i+1} \\ \overset{RH}{Flip\,Swap}(t_i) & if\ t_i > t_{i+1} \end{cases}$$

Fig. 3. Lexical-order action selection operation.

output symbol $\mathcal{T}_n$. For the remaining input symbols w_i ($i = n–1, n–2, ..., 1$), the Backward Inv-Substitute first substitutes the input symbol w_i and then XORes the outcome t_i with the most recently processed input symbol w_{i+1}. The result is the output symbol $\mathcal{T}_i$, which is appended to the output T.

The output of the Backward Inv-Substitute $\mathcal{T}_1\mathcal{T}_2...\mathcal{T}_n$ is then passed to the Forward Inv-Substitute. The Forward Inv-Substitute has the same functionality as the Backward except that it starts from the first input symbol rather than from the last one. Therefore, it first substitutes the input symbol $\mathcal{T}_1$ using the inverse S-Box to yield the output symbol k_1. For the remaining input symbols $\mathcal{T}_i$ ($i = 2, 3, ..., n$), the Forward Inv-Substitute first substitutes the input symbol $\mathcal{T}_i$ and then XORes the outcome S_i with the most recently processed input symbol $\mathcal{T}_{i-1}$. The outcome of the XOR operation k_i is appended to the output block K.

3 Key Expansion Method

The key expansion method extends encryption keys to any arbitrary length. The method uses two operations: Substitute method and Manipulate operation. The Substitute method (Sect. 2) substitutes the input key and the outcome is then passed to the Manipulate operation for further processing.

The Manipulate operation largely distorts its input. The distortion includes modifications to both individual symbols and the structure of the input block. Suppose $t_1t_2...t_n$ be an input block. The Manipulate operation is composed of two actions, which are defined in Fig. 3.

The Manipulate operation handles each symbol t_i based on the lookahead symbol t_{i+1}. If the symbol t_i is lexically before or equals to the lookahead symbol t_{i+1}, the action $Flip^{LH}(t_i)$ flips the left half bits of t_i. If the symbol t_i is lexically after the lookahead symbols t_{i+1}, the action $FlipSwap^{RH}(t_i)$ flips the right half bits of t_i and moves the resulting symbol to a position determined by the lookahead symbol t_{i+1}. Consider for instance the string "pqetreelookc". If the current symbol is "p", the left half bits of "p" (**0111**0000) are flipped because "p" is lexically before the lookahead symbol "q" yielding "€" (10000000). Assume that the current symbol is "q". The lookahead symbol is thus "e". Since "q" is lexically after "e", the right half bits of "q" (0111**0001**) are flipped (0111**1110**) yielding "˜". Additionally the resulting symbol "˜" is moved to the position 101 % 12 = 5 in the string. (101 is the unicode index of the lookahead symbol "e" and 12 is the length of the input string.) This operation results in "€...˜.....".

fridaya@uef.fi

```
Manipulate (t_i)
   1. Generate a random γ_i ∈(0, 1)
   2. If (γ_i ≤ LTE) THEN execute Flip action.
   3. Else execute FlipSwap action.
```

Fig. 4. The random process for action selection.

Although it may be effective to directly use the lexical order of the symbols as a condition for executing one of the manipulation actions, it is better to add more confusion to the action selection. We believe this confusion can largely be achieved using some sort of randomness. Thus, instead of selecting an action based on the lexical order of the input symbols, we bring a random noise to this selection. Let T be the number of the so-far processed symbols and P be the number of times in which $t_i \leq t_{i+1}$ holds. Based on this, we define the following two likelihood variables.

$$LTE = \frac{P}{T}, \qquad GT = 1 - \frac{P}{T}$$

Where *LTE* ("Less Than or Equal") is the ratio in which the condition $t_i \leq t_{i+1}$ holds. *GT* ("Greater Than") is the ratio in which the condition $t_i \leq t_{i+1}$ does not hold. Using these likelihood variables, we redefine our *Manipulate* operation (Fig. 3) in Fig. 4.

The amount of the bias toward any of the two actions changes as the values of *LTE* and *GT* change. When *LTE* increases (decreases), the likelihood of executing the *Flip* action is larger (lower) than the likelihood of executing *FlipSwap* action. Hence, increases (or decreases) in the value of *LTE* do change the likelihood of the executing an action over the other, but the choice of which action to execute is random since it depends on the random value γ_i.

After defining the two operations that constitute the key expansion, we list this method of expanding a key in Fig. 5. The process takes a key of length n as an input and returns an extended key (of arbitrary length) as an output. The process consists of the steps 1 through 7. The symbols of the input key are first substituted with new symbols x_i's using the substitute method (Sect. 2). Next, the new symbols x_i's are manipulated using the manipulation operation (Fig. 4). To define the amount of the bias *LTE* (line 6), we always increment the number of processed symbols T while incrementing P only when the currently processed symbols x_i lexically equals to or is before the lookahead symbols x_{i+1}. Therefore, the ratio *LTE* is always in the interval [0, 1] and represents the likelihood of executing *Flip* action.

The new version of the key S is concatenated with the original key. If the desired length has not been reached yet (step 9), the process uses the latest version S of the key to create a new version.

fridaya@uef.fi

4 Key-Based Random Number Generator

The number generator uses the encryption key to generate sequences of numbers. These sequences provide input for the encryption (decryption) process to deliver their functionality. Ideally, the number generator should depend on the key and should be very responsive to the changes of the key. In this case, the number generator can effectively respond to the changes of the key regardless of the magnitude of the change (even a single bit) by producing very different sequences of numbers. Furthermore, the output of the generator must pass important randomness tests. These are "must" conditions for the successful application of the generator in the cryptography field.

```
Input: Key=k1k2 ...kn
Output: Expanded-Key = Key
1.  Let L = Key, T = 1, P =0
    /*T=total number of processed symbols and P = number of times in which
    the current input symbol lexically comes before the lookahead symbol*/
2.  x1x2 ... xn←Substitute (L)
3.  For i=1 to n Do
4.    If (xi ≤ xi+1) P ++ //increment P
5.    T ++
6.    LTE = P/T
7.    S = S+ Manipulate (xi)
8.  Expanded-Key = Expanded-Key + S
9.  If desired length not reached yet, L = S, GOTO 2
10. Return Expanded-Key
```

Fig. 5. The random process for action selection.

Figure 6 shows the algorithmic steps of the proposed number generator. The first step is to extend the key into 64 symbols. The reason for this extension is to have a large seed that greatly increases the period of the generator. The steps (2)–(7) define the major functionality of the generator. Step (2) uses the Substitute method (Sect. 2) to enable the generator to propagate the key changes to the symbols. That is, small changes to the key entail highly different seed. Step (3) declares the operation $Flip^R$, which flips the right n bits of the symbol at index m. The values of n and m are respectively the integer values for the seed symbol at index 0 and the seed symbol at index 1 (i.e. n = (INT) Seed [0] and m = (INT) Seed [1]). Step (4) declares the operation $Shift^L$, which circularly left shifts the symbols of the resulting Seed by k positions. The value of k is the integer value for Seed [2]. Clearly the operations (3) and (4) change the Seed, allowing the Substitute operation (2) to more effectively change the Seed. Steps (5) and (6) sum up the symbols of the resulting Seed by multiplying the integer value of the Seed symbol at index i with the power of 256. The reason for taking the power of 256 is that radix for the symbols 0 to 255 is 256. Therefore, this summation never yields the same Sum value for different Seeds.

fridaya@uef.fi

We take the module P for the Sum (step (7)) so that all the generated numbers are with the range [0, P]. If more numbers are required, the generator repeats the steps (2)–(7).

5 Block Diffusion Method

The sensitivity of the encryption technique to the changes of the input is considered one of the most important security properties. The encryption technique is highly sensitive to input change if it responds to the change (no matter how small it is) by causing tremendous changes to the output. This property has a large impact on one of the most important security measures, the avalanche effect (a tiny change (even to a single bit) must entail huge changes to the output).

```
Input: Key
Output: sequence of key-based numbers with an arbitrary length.
1. Extend the key to 64 symbols to get the Seed.
2. Seed = Substitute (Seed)
3. Seed = Flip^R (Seed, n, m) /*flip the right n bits in the symbol Seed[m]*/
4. Seed = Shift^L (Seed, k)    /* circular left shift symbols of Seed k positions*/
5. For i=1 to |Seed|           /*|Seed| is the length of Seed*/
6.   SUM += 256^(|Seed|-i) × (INT) Seed[i] /*Seed[i] is the symbol at index i*/
7. RAND = SUM mod P    /*mod is module operation and P is the range of the numbers*/
8. If More numbers needed, GO TO 2
```

Fig. 6. The algorithmic steps for the key-based number generator.

The proposed diffusion method herein grants our encryption method an extremely high avalanche effect. The diffusion method defines two operations. The first method is called Diffuse (DIF). The DIF operation amplifies the changes that occur to the input and makes them largely affect all the output. The second method is called Inverse-Diffuse (INV–DIF). This method reverses the impact of the DIF and obtains the original block.

Before we explain the DIF and INV–DIF methods, we introduce two operations that are used in the DIF and INV–DIF. These two new operations are Distort and Inverse-Distort.

5.1 The Distort Operation

The distortion operation (DIS) makes drastic changes to individual symbols of the input block as well as to the structure of this block. (Changing the structure of the block means breaking the original order of its symbols.) The distortion operation defines 56 different distortion actions. Table 1 shows the actions, their inverse, and a simple description of each action. The actions are of two types: unary and binary. The unary actions take a single input symbol while the binary actions take two input symbols. Each of the unary actions SL_n and SR_n takes

fridaya@uef.fi

a single input symbol and circularly shifts the bits of this symbol n bits to respectively the left or the right. The unary action ML_n mutates (flips) the left n bits of its input symbol.

The binary actions take two input symbols. The CLL_n action crosses over (exchanges) the left n bits of the first argument with the left n bits of the second argument. The action CRL_n exchanges the right n bits of the first argument with the left n bits of the second argument. Note for the binary operations, when n equals to the total number of the bits that represent the symbol in the encoding system, the impact of these functions is to permute the positions of two arguments in the block.[1] Therefore, the binary actions have dual impact on their inputs: distort the individual symbols or alter the structure of the input block.

Table 1. The distortion operations.

	Action	description	Inverse Action	Description
Unary operations	***SL_n***	Circularly shifts the bits of the input symbol *n* positions to the left (*n*=1, 2, …, 8)	***SR_n***	Circularly shifts the bits of the input *n* positions to the right (*n* =1, 2, …,8)
	SR_n	Circularly shifts the bits of the input *n* positions to the right (*n* =1, 2, …,8)	***SL_n***	Circularly shifts the bits of the input symbol *n* positions to the left (*n*=1, 2, …, 8)
	ML_n	Mutates (flips) the left *n* bits of the input symbol (*n* =1, 2, …, 8)	***ML_n***	Mutates (flips) the left *n* bits of the input symbol (*n* =1, 2, …, 8)
Binary operations	***CLL_n***	Crosses over the left *n* bits of the first input symbol with the left *n* bits of the second input symbol (*n*=1, 2, …,8)	***CLL_n***	Crosses over the left *n* bits of the first input symbol with the left *n* bits of the second input symbol (*n*=1, 2, …,8)
	CRR_n	Crosses over the right *n* bits of the first input symbol with the right *n* bit of the second input symbol (*n*=1,2,…,8)	***CRR_n***	Crosses over the right *n* bits of the first input symbol with the right *n* bit of the second input symbol (*n*=1,2,…,8)
	CRL_n	Crosses over the right *n* bits of the first input symbol with the left *n* bit of the second input symbol (*n*=1,2, …, 8)	***CRL_n***	Crosses over the right *n* bits of the first input symbol with the left *n* bit of the second input symbol (*n*=1,2, …, 8)
	CLR_n	Crosses over the left *n* bits of the first input symbol with the right *n* bit of the second input symbol (*n*=1,2,…,8)	***CLR_n***	Crosses over the left *n* bits of the first input symbol with the right *n* bit of the second input symbol (*n*=1,2,…,8)

These distortion actions are placed in a lookup table. Each row of the table consists of two columns, where the action is stored in the first column and its inverse is stored in the second. In order to make the selection of any of the entries (the action and its inverse) random, we randomly reorder the actions in the lookup table using a sequence of random numbers obtained from our random generator. The reordering is performed by swapping the contents of the row i with the contents of the row r_i (r_i is random number).

[1] For instance, if the symbols are represented by 8 bits (Extended ASCII), then the binary actions with index 8 (e.g. CRR_8) causes the two input symbols of the input to swap.

fridaya@uef.fi

The DIS operation handles its input block in two passes: forward (left to right) and backward (right to left). Let $b_1b_2...b_n$ be plaintext block. The forward pass handles the block as follows. The first symbol b_1 in the input block stays without processing. For all the symbols b_i $(i > 1)$, the operation applies one of the 56 distortion actions to the input. The selection of which action (be unary or binary) to apply to the symbol b_i is fully determined by the most recently processed input symbol b_{i-1}. More specifically, let $c_1c_2...c_{i-1}$ be the so-far output block that is resulted from processing the input symbols $b_1b_2...b_{i-1}$. To select an action for processing the next input symbol(s), the forward pass uses the most recently processed input symbol b_{i-1}. We basically compute the index of the distortion action by module operation: $j = \text{Val}\ (b_{i-1})\ \text{mod}\ 56$, where Val (b_{i-1}) is the decimal code of the symbol b_{i-1} and "mod 56" since we have 56 actions. The action at the calculated index j is applied to the symbol b_i or the symbols b_i and b_{i+1} based on whether the selected action is unary or binary. If the selected action is unary, this action processes the symbol b_i and yields the output symbol c_i. If the selected action is binary, this action processes the input symbols b_i and b_{i+1} and yields the output symbols c_i and c_{i+1}. In case that the action is binary and the input block has only one unprocessed input symbol, the action is ignored and the input symbol is appended to the output block (without processing).

The backward pass performs identical processing except that it starts from the end of the input $c_1c_2...c_n$. Therefore, the first symbol to start with is c_n, which is left unprocessed. For all the symbols c_i $(i < n)$, the backward pass uses the most recently processed input symbol c_{i+1} to select the distortion action for processing the input symbol c_i or the input symbols c_i and c_{i-1} depending on whether the selected action is unary or binary. Additionally, to highly secure the output of the distortion operation, the backward pass seals the rightmost output symbol with a random number obtained from the random generator.

5.2 The Inverse-Distort Operation

The Inverse-Distort operation (INV-DIS) negates the changes made by the distort operation (DIS). In other words, INV-DIS operation reverses the impact of DIS operation and recovers the original block. Like DIS operation, INV-DIS operation handles its input in two passes: forward and backward. Let $b_1b_2...b_n$ be an input block to INV-DIS operation. The INV-DIS starts with the **backward** pass. In this pass, INV-DIS first unseals the rightmost input symbol b_n to obtain the new symbol c_n. The symbol c_n is appended to the output without further processing. Thus, the so-far output list contains "c_n". For the input symbols b_i $(i < n)$, INV-DIS operation uses the most recently appended symbol to the output c_{i+1} to determine which inverse action (Table 1) to apply to the input block. Determining which inverse action to apply is done exactly as done in the DIS operation.

The output of the backward pass is passed to the **forward** pass. This pass performs identical steps as the backward except it starts from the leftmost symbol of the input block. The leftmost symbol c_1 is appended directly to the output

fridaya@uef.fi

without processing. For all symbols c_i ($i > 1$), the forward pass uses the most recently appended symbol to the output, say d_{i-1}, to determine the inverse action to apply to the input symbol c_i or to the input symbols c_i and c_{i+1} depending on whether the inverse action is unary or binary.

5.3 The DIF Method

This method consists of two operations: (1) Substitute operation and (2) Distort operation (DIS). The Substitute operation executes first. Based on definition of the Substitute operation (Sect. 2), when a symbol b_i changes the forward Substitute ensures that this change affects not only the substitution outcome of b_i, but also the substitution outcome of all the symbols b_j ($j > i$). In addition, the backward Substitute ensures that the change in b_i affects all the symbols b_j ($j < i$). As a result, the change actually impacts every single symbol in the block. The output of the Substitute operation is passed to the Distort operation. This operation changes the individual symbols and the structure of the input block.

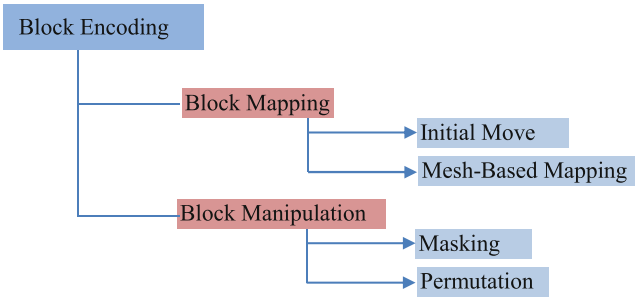


Fig. 7. The block encoding process.

5.4 The INV-DIF Method

The INV-DIF method (Inverse Diffusion) reverses the impact of the diffuse operation (DIF). This operation uses both the Inverse-Substitute operation and Inverse-Distort (INV-DIS). These two operations are executed in the following sequence: **INV-DIS** then **Inverse-Substitute**.

6 Plaintext Block Encoding

The block encoding process transforms plaintext blocks into a new representation. This process is designed to largely increase the confusion. In other words, the encoding process melts the relationship to the original plaintext blocks. Figure 7 shows the logic of this process. The process consists of two operations: *Block Mapping* and *Block Manipulation*. The Block Mapping consists of two

fridaya@uef.fi

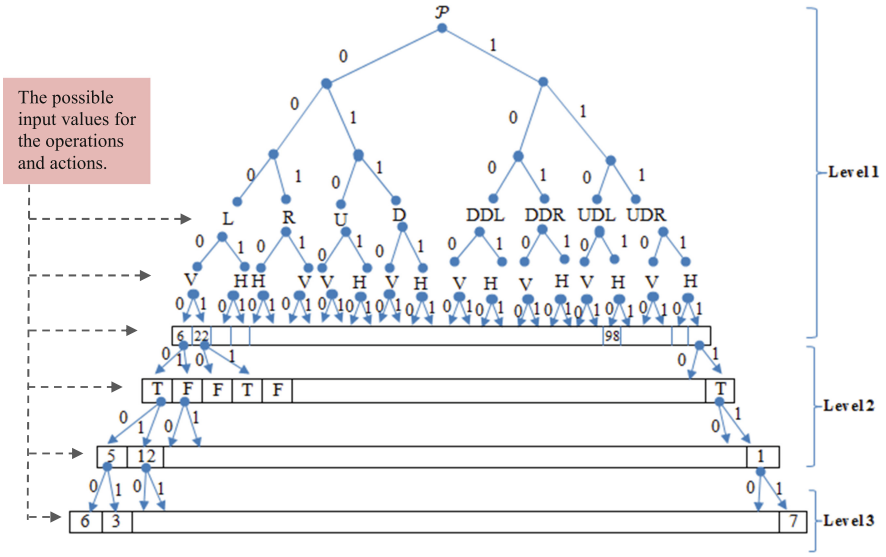


Fig. 8. The operational knowledge.

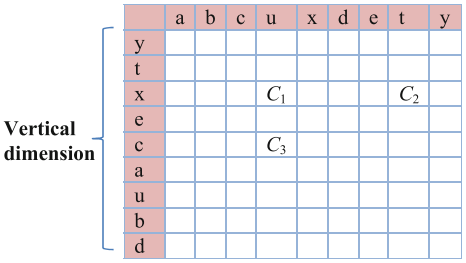


Fig. 9. The mesh.

actions *Initial Move* and *Mesh-Based Mapping*. The Block Manipulation consists of also two actions: *Masking* and *Permutation*.

In order for the operations and actions in Fig. 7 to properly function, they require operational knowledge that provides them with the appropriate input. Figure 8 shows this knowledge $\mathcal{P}$, which is a labeled binary tree. The tree consists of three levels (Level 1, 2, and 3) each of which provides the necessary input for one or more encoding actions. Each path in the tree is label with a binary value (0 or 1) and leads to some input values that will be used by the encoding actions. For instance, the path labeled with "000" leads the input value L While the path "0000" leads to the input value V. The following sections discuss how the encoding actions use the knowledge in Fig. 8 to process their input.

fridaya@uef.fi

6.1 Block Mapping

The block mapping transforms each symbol of an input block to another representation, called a directive. This section defines all the operations that do such a transformation.

The Mesh. The mesh, as defined in [1], is a two way $N \times N$ array with horizontal and vertical dimensions. Figure 9 shows an example of the mesh. Each of the two dimensions is populated with the unicode symbols from 0 to N–1. The symbols in each dimension are indexed by integers 0, 1, ..., N–1. A move from cell C_1 to C_2 has a distance, which is the number of the passed cells. For instance, moving from C_1 to C_2 has a distance of 4 since we passed 4 cells.

Since a move starts from some cell and ends at another, we introduce the move direction concept. We designate a direction of a move from some cell to another by the flag "–" if the move is toward a cell with a lower index, and by the flag "+" if the move is toward a cell with a higher index. For instance, a move from cell C_3 to C_1 is designated by "–" because the move is toward a lower index and the move from C_1 to C_2 is designated by "+" because the move is toward a higher index.

Given a distance of a move and its direction, we can formally define a directive as follows. If we move from a cell, say C_1, to another cell, say C_2, by a distance x along some mapping dimension (vertical or horizontal) we designate this by "–x" or "+x" depending on whether the move is toward respectively lower or higher indexes. We call "–x" or "+x" a directive. We call x the distance and the sign (+ or –) the flag.

Given the mesh and directives, we define mapping a symbol to the mesh as follows. Let b_i be a symbol. We map b_i to the mesh by starting from some designated point (called starting point) and moving along one of the dimensions to the index of the symbol in this dimension. We call the dimension that we move along the mapping dimension. The distance of the move and its direction with respect to the starting point is compiled as a directive "–x", which represents the outcome of mapping the symbol b_i. For instance, suppose we want to map the symbol "y", where the starting point is C_3 and the mapping dimension is the horizontal. We start from C_3 and move along the horizontal dimension to the index of "y" at this dimension. Since we passed 5 cells and the direction of the move is toward the higher indexes, we compile this to the directive "+5", which is the outcome of mapping "y" to the mesh.

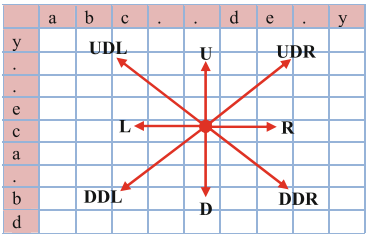
Besides a symbol mapping, we can also define the inverse mapping. The inverse mapping is a resolution process that takes a directive as an input and returns the corresponding symbol. The inverse mapping works as follows. Let "$\pm y$" be a directive, which was produced by mapping the symbol p starting from a starting point s. The inverse mapping for the directive "$\pm y$" is performed by moving y cells from the starting point s along the specified mapping dimension. This move is toward the lower or higher indexes as specified by the directive flag. We then look up the symbol that is located in the mapping dimension and corresponding to the current index. For instance, the inverse mapping for the

fridaya@uef.fi

directive "+5" is calculated by starting from the starting point C_3 and moving 5 cells along the horizontal dimension toward the higher indexes (since the flag +). We look up the corresponding symbol from the horizontal dimension, which is "y".

In addition to mapping a symbol to the mesh, we also define a set of possible instructions that defines additional possible moves in the mesh. Figure 10 shows the mesh and the possible instructions for a move from some cell (left). The figure shows also the definition of each instruction.

As the figure shows, we define 8 instructions for a move within the mesh. Each instruction causes the mapping operation (described next) to make a move along one of the eight designated directions prior to performing any symbol mapping. We call this prior move, the initial move. For instance, the instruction U forces the mapping operation to move up from the current position while the instruction UDL forces the mapping operation to move up diagonally left.



Instruction definition
U: move up
D: move down
L: move left
R: move right
UDL: move diagonally up and toward the left
UDR: move diagonally up and toward the right
DDR: move diagonally down and toward the right
DDL: move diagonally down and toward the left

Fig. 10. The possible moves in the mesh.

These instructions are used by the mapping operation to perform some initial movements in the mesh before mapping any symbol. The main objective of these initial movements is to complicate the symbol mapping and highly increase its confusion.

The Initial Move Action. The initial move action takes an n-place tuple t_i and determines both the initial move parameters and the mapping dimension. This action can be defined by the procedure $g_1 : T \longrightarrow L \times M \times D$. The domain (or the input to the action) T is a set of n-place tuples $t_i <b_0, b_1, ..., b_n>$, where b_i are all binary values (0 or 1). The procedure g_1 maps each tuple t_i to a triple $(\psi_i, \omega_i, d_i) \in L \times M \times D$, which represents the parameters of the move within the mesh ψ_i, ω_i and the mapping dimension d_i. The initial move parameters consist of two values that fully specify this move within the mesh. These two parameters are the direction of the initial move $\psi_i \in \{L, UDL, U, UDR, R, DDR, D, DDL\}$ and the amount $\omega_i \in N$ (natural number) of the initial move along the direction ψ_i. The value $d_i \in D = \{H, V\}$ represents one of the dimension of the mesh, where H indicates the horizontal dimension and V indicates the vertical dimension. In other words, for each tuple t_i, the procedure g_1 outputs three parameters

fridaya@uef.fi

that specify (1) the direction of the initial move, (2) its amount (how many cells), and (3) the mapping dimension, which will be used to map symbols to the mesh.

The procedure g_1 is fully defined by **Level 1** of the binary tree in Fig. 8. Figure 11 shows the partial tree extracted from the tree in Fig. 8. This partial tree has three possible mapping values: initial move directions, mapping dimension, and the amount of move. Since we have 8 potential directions along which an initial move can be performed (see Fig. 10), we use three bits to uniquely identify each potential direction. In addition, we dedicated one bit to identify the appropriate value for the mapping dimension D from the two possible values H or V. Finally, we dedicated one bit to identify one of the potential values for the amount of the move ω. Therefore, the input to the g_1 procedure is 5-place tuples, where the three leftmost bits allow for uniquely identifying one the eight initial move directions, the 4^{th} bit (along with the preceding three bits) allows for identifying a mapping dimension, and the 5^{th} bit (along with the four preceding bits) allows for identifying the amount of the initial move.

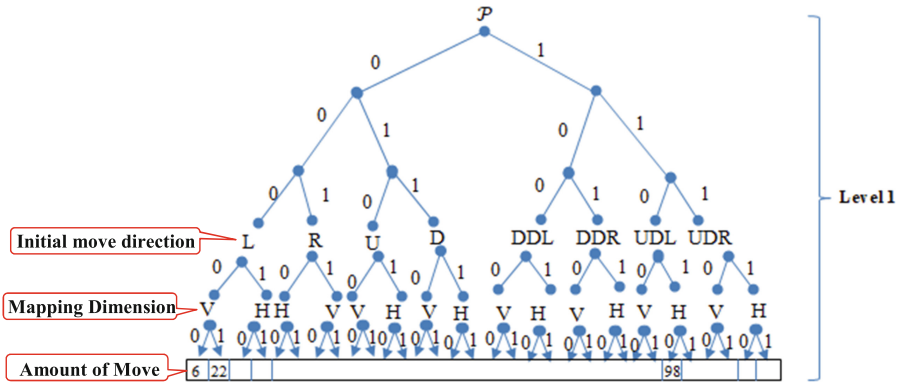


Fig. 11. The initial move parameters calculation logic.

According to Fig. 11, the procedure g_1 works as follows. For a tuple $t_i <b_0, b_1, b_2, b_3, b_4>$, the first three bits $b_0b_1b_2$ selects a direction for an initial move by following the edges labeled by these three bits. Continuing down the tree in the same path, the forth bit b_3 selects the mapping dimension and the fifth bit b_4 selects the amount of the initial move. For instance, for the tuple $t_1 <0, 0, 0, 0, 0>$, the procedure $g_1$returns the initial move direction L, the mapping dimension V, and the amount of move "6" while for the tuple $t_2 <1, 1, 0, 1, 0>$, the procedure returns the initial move direction UDL, the mapping dimension H, and the amount of the move "98".

To fully define the procedure g_1, we should determine how to create the values for the amount of the initial move ω. We require that ω to be a random value. To satisfy this constraint on ω, we use the random number generator to periodically update the set of possible values after encrypting each block.

fridaya@uef.fi

Mesh-Based Mapping Action. After defining the mesh and how to determine the initial move parameters, we are ready to propose how we map input symbols to the mesh and produce directives. The mesh based mapping takes as an input an initial point, a sequence of symbols to be mapped, and a key. The initial point is a point (x, y) within the boundary of the mesh. The integer x is the row index and the integer y is the column index. Both x and y are chosen randomly using our random generator. The mapping process returns as an output a sequence of directives.

The mapping proceeds as follows. The key is transformed into a binary sequence by finding the binary equivalent to each symbol in the key. For instance, if the key is "ab", the binary representation is "01100001 01100010". In case we have more input symbols to map, we expand the key and use the new version for completing the mapping.

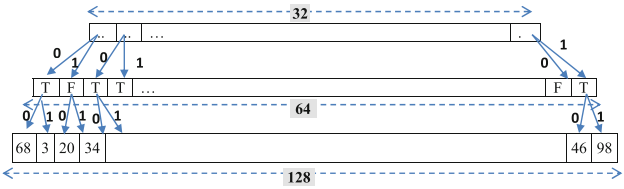


Fig. 12. The partial tree (from Fig. 8) that pertains to the masking action.

For each symbol c_i, the mapping action passes the tuple t_i to the procedure g_1 to compute (1) the initial move parameters ψ_i, w_i and (2) the mapping dimension d_i. From the current point, the mapping operation first moves within the mesh boundary w_i cells along the direction specified by ψ_i. If the amount of the move exceeds the boundary of the mesh, we wrap to the opposite side of the same direction and continue.

Starting from this new position, the mapping action uses the designated mapping dimension $d_i \in \{H, V\}$ to map the symbol c_i as follows. The mapping action moves from the new position along the mapping dimension to the index of c_i at this dimension. The number of passed cells and the direction with respect to the position is translated into a directive $\pm x$.

6.2 Block Manipulation Operation

The manipulation operation performs two actions that alter the individual directives and modify the structure of the directive sequences.

Directive Masking Action. The **Level 2** (Fig. 8) provides knowledge for the masking action. This action distorts both the distance of the directive and its flag (the sign). Level 2 is reproduced in Fig. 12, which is extracted from Fig. 8. Since the bottom of level 1 has 32 leaves, the bottom of the first part of level 2

fridaya@uef.fi

has 64 leaves and the second part has 128 leaves. The 64 entries are alternately populated with the symbols "T" and "F". These symbols are randomly reordered using random numbers obtained from the random generator. The reordering is performed by moving the symbol at index i to the position r_i (where r_i is a random number and $i = 1, 2, ..., 64$).

These symbols ("T" and "F") are used by the masking action as a criterion to whether flip the flag of the directive or not as follows. (Flipping the flag means changing the "+" to "–" and vice versa.) Given a tuple $t_i <b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7>$, the 6^{th} bit b_5 (along with the proceeding 5 bits) is used to access one the symbols ("T" or "F"). The symbol "T" instructs the masking action to flip the directive flag while the "F" instructs the masking action to keep the flag unchanged.

The 128 leaves in Fig. 12 are populated with random numbers obtained from our random generator. The masking action uses these random numbers to mask the distance of the directive as follows. Given the tuples $t_i <b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7>$, the bit b_6 (along with the preceding six bits) is used to access one of the 128 entries. The accessed random number is XORed with the distance of the directive x_i (without the sign).

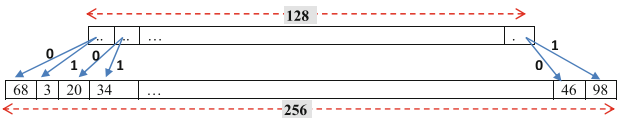


Fig. 13. The partial tree (from Fig. 8) that pertains to the permutation action.

Directive Permutation Action. Level 3 (Fig. 8) provides input for the permutation action. This level is reproduced in Fig. 13. Level 3 consists of 256 entries. These entries are filled with random numbers obtained from the random generator. Given a tuple $t_i <b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7>$, the permutation action uses the 8^{th} bit (b_7) of the tuple t_i to access one of the 256 entries. The looked up value r_i is used to move the current directive $-x_i$ to the position r_i.

After we defined all the operations and actions of the block encoding process, we show how these actions are used to encode a block of plaintext. The encryption key is divided into 8-place binary tuples $<b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7>$. The left 5 bits "$b_0b_1b_2b_3b_4$" are used by the actions Initial Move and Mesh-Based Mapping to produce directives for the plaintext block. In particular, the first three bits "$b_0b_1b_2$" are used to select direction of the initial move, the forth "b_3" is used to identify the mapping dimension, and the fifth bit "b_4" is used to determine the amount of the initial move. The 6^{th} and 7^{th} bits ("b_5" and "b_6") are used by the Masking operation to mask the directives. Finally the 8^{th} bit "b_7" is used by the Permutation action to reorder the produced sequence of directives.

fridaya@uef.fi

7 Ciphered Block Decoding

The block decoding process is the inverse of the block encoding. It processes ciphered blocks and recovers the corresponding original plaintext blocks. Figure 14 shows the decoding process operations and actions. As the figure shows, the decoding process has roughly the same operations and actions as the encoding process except that (1) the operations and the actions are executed in reverse order and (2) some of the actions have slightly different functionality. Thus, the block manipulation operation is executed first and then the block mapping. The actions of the block manipulation operation are executed in reverse order. We execute the *Inverse Permutation* action and then the *Masking* action. The actions of the block mapping operation are executed as specified by Fig. 14. The actions in Fig. 14 whose names match those in Fig. 7 have identical functionality. The rest such as Inverse Permutation have slightly different functionality and thus we will discuss them.

Let $x_1x_2...x_n$ a sequence of directives and $t_1t_2...t_n$ be a sequence of 8-place tuples, where each tuple t_i is used to produce the directive x_i. The Inverse Permutation action (which reverses the effect of Permutation action), processes a sequence of tuples backwards: from the last tuple t_n down to t_1. For each tuple t_i, the action uses this tuple to access the appropriate entry in level 3. The retrieved value from level 3, say r_i, is used to move the directive at the position r_i (x_{r_i}) to the current position i.

Once all the directives sequence correctly permutated, the masking action fires. The masking action processes the sequence of directives in usual way: from left to right. In particular, the left 6 bits of the tuple t_i are used to access one of the flag masking entries of Level 2 and the left 7 bits are used to access one of the distance masking entries. The retrieved symbol from flag masking entries is used to unmask the flag and the retrieved value from distance masking entries is used to unmask the distance of the directive.

The Block Mapping operation is executed next. The output of the previous operation, say $y_1y_2...y_n$, is processed in natural order from left to right. Therefore, for the tuple t_i, the *Initial Move* action determines the three parameters of the initial move in exactly the same way as in the encoding process. After

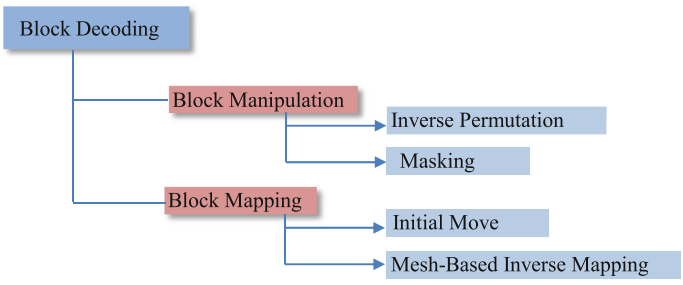


Fig. 14. The decoding process.

fridaya@uef.fi

the initial move action performed the initial move within the mesh, the *Mesh-Based Inverse Mapping* action uses the mapping dimension determined by the initial move action to recover the symbol encoded by the directive y_i. The inverse mapping does this by starting from the current point (determined by the initial move) and moving along the mapping dimension in a direction specified by the directive flag ("+" or "–"). The corresponding symbol in the mapping dimension is looked up, which is the original symbol encoded by the directive.

8 Key-Controlled Directive Substitution

The key-controlled substitution replaces a directive with a symbol. The directives $\pm d$ are resulted from mapping the plaintext symbols to the $N \times N$ mesh. Therefore, the range of the distance part of a directive (d) is from 0 to N–1. Since each directive is flagged with "+" or "–", the total number of directives is $2N$–1. For instance, if $N = 256$, the number of the directives flagged with "+" is 256 and the directives flagged with "–" is 256 (512 in total). As such, each directive $\pm d$ can be represented by 9 bits—8 bits for the distance (d) and one for the flag (+ or –).

We define our key-dependent directive substitution as a routine S: $D \mapsto N$, where D (the input) is a set of directives and N (the output) is a set of symbols 0...$2N$–1. To define how the routine S maps directives to the specified symbols, we utilize a $K \times L$ lookup table (called DIR–SUB). In our discussion, we assume, without losing the generality, $N = 256$. Under this assumption, the lookup table DIR–SUB can be represented by 32×16 array. We populate the table with the symbols from 0 to 511. These symbols are randomly scattered within the table. The random scattering employs a set of $2N$–1 random numbers obtained from the random generator. The scattering is performed by swapping the entry at the index i with the entry at the index r_i (a random number). In addition, to complete the definition of the routine S, we specify the range of symbols 0...511 that correspond to the plus-flagged directives and to the minus-flagged directives. To do this we split these symbols into two halves: 0...255 and 256...511. We let the "+" flagged directives correspond to the integers 0...255 and the "–" flagged directives correspond to the symbols 256...511. Observe, when we assign the +x directives to the symbols 0...255, we implicitly represent the "+" with the bit 0.

The substitution using the routine S is performed as follows. For any directive $\pm d \in D$, the left 5 bits of the directive index one of the 32 rows of the DIR–SUB table and the right 4 bits index one of its 16 columns.

Since it is required to recover the original directives during the decryption process, we define the Inverse Substitution. The Inverse Substation is a routine S^{-1}: $N \mapsto D$. To define how the routine S^{-1} maps a symbol $n \in N$ to a directive, we utilize a $K \times L$ lookup inverse table (called DIR–SUB^{-1}. This inverse table is created from the corresponding DIR–SUB table. The entries of DIR–SUB^{-1} are created as follows. Suppose that the symbols n is in the cell DIR–SUB$[r, c]$. First, we form a symbol rc from the row index (r) and the column index (c) by concatenating the five bits that represent r and the four bits that represent c.

fridaya@uef.fi

Second, the 9-bit representation of n is obtained and the left five bits become the row index r' and the right four bits become the column index c'. We then insert the symbol rc in the cell DIR–SUB$^{-1}[r', c']$. For instance, the symbols "234" ("ê") is at DIR-SUB [8, 2]. The five bits representation of 8 is "01000" and the four bits representation for 2 is "0010". The concatenation is "010000010", which is the decimal 130. On the other hand, the 9 bits representation for 234 is "011101010". Now, the left five bits "01110" become the row index and the right four bits "1010" become the column index. Thus, we place the value 130 in the cell DIR–SUB$^{-1}[4, 10]$ of the inverse substitute table.

The substitution using the routine S^{-1} is performed as follows. For any symbol $n \in N$, the left 5 bits of n index one of the 32 rows of the inverse substitute table and the right 4 bits index one of the 16 columns.

9 The Cipher Technique

This section presents the proposed cipher. Namely it provides the details of the encryption process (Subsect. 9.1) and the decryption process (Subsect. 9.2).

9.1 The Encryption Process

The encryption process encrypts plaintext blocks $b_1 b_2 ... b_n$ using a key $k_1 k_2 ... k_m$. We impose no specific restrictions on the length of the block or the key. The encryption process consists of two stages: the *initialization* stage and the *encryption* stage. Figure 15 shows the main performed tasks during the initialization stage. The key expansion process extends the key into 64 bytes, which are used as a seed to the random generator. The random generator provides sequences of random numbers for performing the following tasks: (1) reorder the symbols of the mesh's two dimensions, (2) populate and reorder the entries of the binary tree in Fig. 8, and (3) reorder the directive substitution table (DIR-SUB). In more details, a sequence of $2N$ random numbers is used to reorder the symbols of the vertical and horizontal dimensions (N random numbers for each). A sequence of L random numbers is used to populate and reorder the entries of the binary tree (Fig. 8). Finally, a sequence of $2N$ random numbers is used to reorder the entries

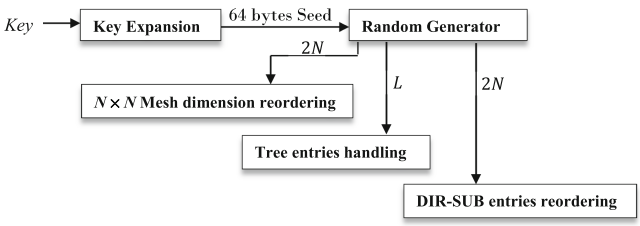


Fig. 15. The activities of the initialization stage.

fridaya@uef.fi

of the DIR-SUB table. All the symbol reordering is performed in the same way. The symbol at entry i is moved to the location r_i (r_i is random number).

The encryption stage handles its input as shown in Fig. 16. For each plaintext block $b_1b_2...b_n$, the key expansion operation extends the key to produce the sequences K^1, K^2, ..., K^{k+1}, where each K^i consists of n symbols (n is the length of the block). The **Diffusion** process handles the input block using k iterations. As previously discussed, the diffusion process ensures a high avalanche effect by amplifying change in the input block and propagating this change to impact every symbol in the output block. The output of the diffusion process is passed to the **Encoding** process, which applies two operations to the input according to the specified order: **Mesh-Based Mapping** and **Block Manipulation**. These two operations consume the key sequences K^i as shown in Fig. 17 (left part). The mesh mapping uses the sequence K^1 to map the symbols of the block to a sequence of directives. The block manipulation handles its input (a sequence of directives) in k iterations each of which (iteration) consumes a key sequence K^i. In any iteration, the block manipulation applies two actions: **Directive Masking** and **Directive Permutation**—according to the order specified in Fig. 16. The output of the block manipulation is a sequence of manipulated directives $y_1y_2...y_n$. The Directive Substitute confuses further the directives $y_1y_2...y_n$ by substituting each directive with a symbol as discussed in Sect. 8. The outcome of this operation is the ciphered block $c_1c_2...c_n$.

For every subsequent plaintext block (if any), the last sequence K^{k+1} in the extended key is used by the key expansion operation to create new sequences $K^1, K^2, ..., K^{k+1}$ for encrypting the new block. In this way, every plaintext block is encrypted using a new and a very different version of the key.

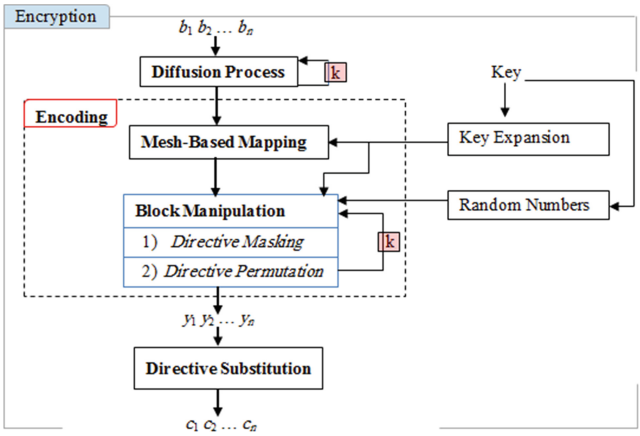


Fig. 16. The encryption process.

fridaya@uef.fi

9.2 The Decryption Process

The decryption process takes a ciphered block $c_1c_2...c_n$ and a key as an input and returns the corresponding plaintext block $b_1b_2...b_n$ as an output. The decryption process is shown in Fig. 18. Like the encryption process, the decryption process consists of two stages: the initialization stage and the decryption stage. The initialization stage is identical to that of the encryption process. It performs all the initializations in Fig. 15 in addition to creating the inverse table DIR-SUB^{-1}. DIR-SUB^{-1} is used by the **Inverse Substitution** operation (Sect. 8).

The decryption stage is similar to the encryption stage except that it (1) uses the inverse operations and (2) executes these inverse operations in reverse order. Therefore, as Fig. 18 shows, the decryption process starts from the **Inverse Directive Substitution** operation, the **Decoding** operation, and the **Inverse-Diffusion** process. Observe that the operations prefixed with the word "Inverse" reverse the effect of the corresponding operations in the encryption process. For instance, the operation **Inverse Directive Substitution** cancels the effect of the operation **Directive Substitution** and extracts the directives y_i's from the symbols c_i's.

Furthermore, the expanded key sequences $K^1, K^2, ..., K^{k+1}$ are used backwards (see the right part of Fig. 17). Therefore, the sequence K^{k+1} is used by the **Inverse Permutation** action, the sequence K^k is used by the **Directive Masking** action, and so on.

10 Performance Analysis

We evaluate our proposed technique in this section. Before we study the performance of the number generator and the encryption technique, we introduce some terminologies. The randomness hypotheses are: H_0: `the tested data is random` and H_1: `the tested data is not random`. Accepting H_0 or H_1 depends on a computed value called p-value and a specified value called the significance level α. The p-value is computed by the applied statistical test based on an input sequence. The significance level α is specified by the tester (e.g. 0.00001, 0.001, 0.01, 0.05 are typical values for α). In particular if p-value $\geq \alpha$, H_0 is accepted (H_1 is rejected); otherwise H_0 is rejected (H_1 is accepted). Finally, in all our tests, we assume the significance level α is 0.05.

10.1 Key-Based Number Generator

We report in this section our preliminary evaluation of the number generator. We are specifically interested in studying its randomness properties. This preliminary evaluation includes a set of 19,024 keys of size 16 bytes (128 bits). We started with 128 keys each consists of 16 identical symbols (e.g. all the symbols are zeros or a's). We then extracted 128 different keys from each key by changing only the bit i of the original key. The total number of keys obtained in this way is 16,512. Additionally, we used 2000 keys, which were generated using the online

fridaya@uef.fi

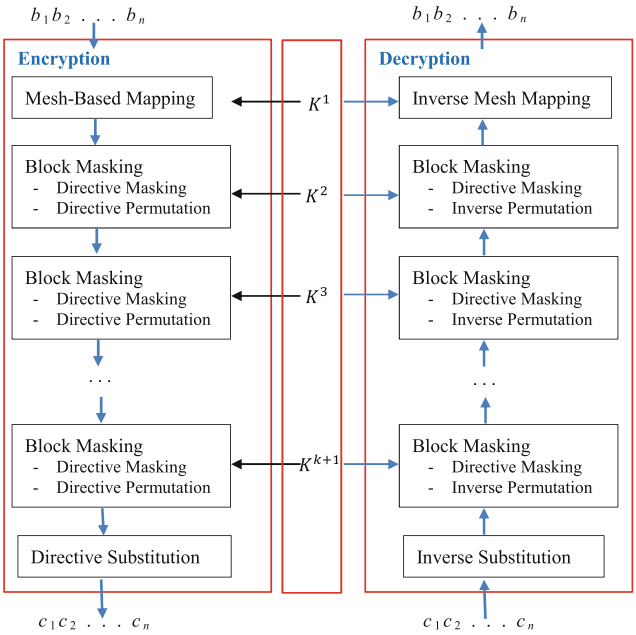


Fig. 17. The key sequence consumption by the encryption/decryption operations.

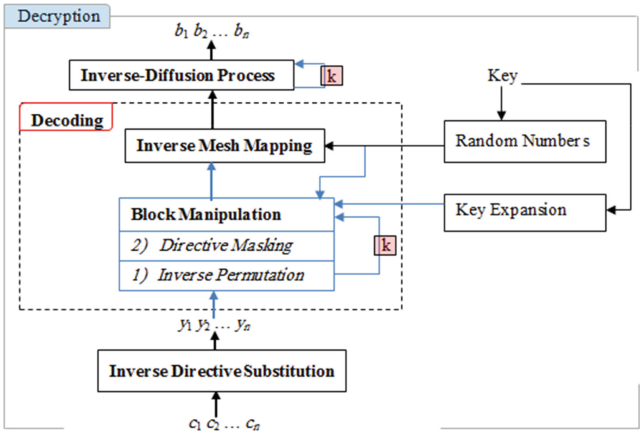


Fig. 18. The decryption process.

service [14]. Finally, we created 512 handcrafted keys. The total number of keys that we used in this evaluation is 19,024 different keys.

These keys are used by the number generator to produces sequences of numbers. Each sequence is composed of 64,000 integer numbers within the range [0, 255]. We tested the randomness of these sequences using two randomness tests recommended by NIST [16]: **Runs Test** and **Monobit Test**.

fridaya@uef.fi

Table 2. The randomness tests results

Runs test				Frequency test (Monobit)		
Sequence	P-Value			P-Value		
	Average	Min	Max	Average	Min	Max
19,024	0.62	0.16	0.99	0.49	0.11	0.89

Table 3. Key avalanche test

Randomness test	Successes	Failures	Success rate	Upper limit of CI (0.05)
Runs test	688	12	98.2%	52.3
Monobit test	685	15	97.8%	52.3
Spectral test	663	37	94.7%	52.3

To apply these two tests, we converted each sequence of numbers to a binary string of zeros and ones. Since all the numbers are integers from the range [0, 255], the binary sequences are created by finding the 8 bits representation of each number (e.g. the binary representation for 12 is 00001100).

Table 2 shows the results of the randomness tests. According to the figures in Table 2, all the 19,024 sequences of numbers passed the two randomness tests. That is because the minimum computed p-value is 0.16 and 0.11 for respectively Runs and Frequency tests, which all exceed the pre-specified significance level 0.05. As such, none of the generated sequences deviates from randomness.

10.2 The Security Analysis

The testing data were prepared according to [15]. We use–without losing the generality–the unicode symbols from 0 to 256. This enables for representing the distance part of each directive by 8 bits. Testing the proposed encryption technique is conducted using the following data sets: (1) **Key Avalanche Data Set**, (2) **Plaintext Avalanche Data Set**, and (3) **Plaintext/Ciphertext Correlation Data Set**. The former two examine the reactivity of our technique to respectively key's changes and plaintext's changes while the latter examines the correlation between plaintext/ciphertext pairs.

Table 4. Plaintext avalanche test

Randomness test	Successes	Failures	Success rate	Upper limit of CI (0.05)
Runs test	686	14	98.0%	52.3
Monobit test	683	17	97.5%	52.3
Spectral test	653	47	93.3%	52.3

fridaya@uef.fi

Firstly, to study the responsiveness of our technique to the key change, we created and analyzed 700 sequences of size 65,536 bits each. We used a 512-bit plaintext of all zeros and 700 random keys each of size 128 bits. Each sequence was built by concatenating 128 derived blocks created as follows. Each derived block is constructed by XORing the ciphertext created using the fixed plaintext and the 128-bit key with the ciphertext created using the fixed plaintext and the perturbed random 128-bit key with the i^{th} bit modified, for $1 \leq i \leq 128$.

Secondly, to analyze the sensitivity to the plaintext change, we created and analyzed 700 sequences of size 65,536 bits each. We used 700 random plaintexts of size 512 bits and a fixed 128-bit key of all zeros. Each sequence was created by concatenating 128 derived blocks constructed as follows. Each derived block is created by XORing the ciphertext created using the 128-bit key and the 512-bit plaintext with the ciphertext created using the 128-bit key and the perturbed random 512-bit plaintext with the i^{th} bit changed, for $1 \leq i \leq 512$.

Thirdly, to study the correlation of plaintext/ciphertext pairs, we constructed 700 sequences of size 358,400 bits per a sequence. Each sequence is created as follows. Given a random 128-bit key and 700 random plaintext blocks (the block's size is 512 bits), a binary sequence was constructed by concatenating 700 derived blocks. A derived block is created by XORing the plaintext block and its corresponding ciphertext block. Using the 700 (previously selected) plaintext blocks, the process is repeated 699 times (one time for every additional 128-bit key).

We applied three randomness tests to these three data sets. The *Runs* and *Monobit* tests, which were described in Sect. 10.1 and additional test called *Discrete Fourier Transform Test* (*Spectral*). The latter test detects periodic features (i.e. repetitive patterns that are near each other) in the tested sequence that would indicate a deviation from the assumption of randomness. Furthermore, we set–in an ad hoc way–the number of rounds K to 12. (We set K to this high number because more rounds typically yield better randomness in the output, but demand more execution time. It is left for our future work to estimate a reasonable value for K that achieves both high randomness in the output and lower execution time.)

Tables 3, 4, and 5 show the results of the randomness tests. The tables present the randomness tests, the number of sequences that passed the respective test (Successes), the number of failed sequences (Failures), and the Success Rate. The significance level was fixed to 0.05 for all the randomness tests. This 0.05 significance level implies that, ideally, no more than 5 out of 100 binary sequences

Table 5. Plaintext/ciphertext correlation test

Randomness test	Successes	Failures	Success rate	Upper limit of CI (0.05)
Runs test	682	18	97.4%	52.3
Monobit test	681	19	97.3%	52.3
Spectral test	657	43	93.8%	52.3

fridaya@uef.fi

may fail the corresponding test. However, in all likelihood, any given data set will deviate from this ideal case. For a more realistic interpretation, we use a confidence interval (CI) for the proportion of the binary sequences that may fail a randomness test at 0.05. We therefore computed the maximum number of binary sequences that are expected to fail the corresponding test at significance level of 0.05. (The maximum number is shown in the rightmost column of each table.) For instance, a maximum of 52.3 (or 52) binary sequences are expected to fail each of the three tests.[2]

The success rate in the three tables is remarkably high. As a matter of fact, none of the success rates is below 93%. For instance, 98.2% of the sequences passed the Runs test for the key avalanche test. Furthermore, the numbers of sequences that failed the randomness tests are less than the maximum expected number of sequences that may fail the tests.

11 Conclusions and Future Work

We proposed an innovative encryption technique, which is significant extension to technique [1]. The technique uses non-linear mesh-based mapping to substitute plaintext symbols in terms of directives. The technique proposed many other distortion and diffusion methods that drastically increase diffusion, confusion, and the avalanche effect. Furthermore, the paper proposed a very effective key expansion method that allows the key to expand to any arbitrary length. The key-based number generator provides pseudo-random numbers that embed random noises to the ciphertext.

The performance evaluations for the number generator and the encryption technique are very promising. In fact, the randomness tests for both the number generator and encryption technique output indicate that their outputs do not deviate from randomness. We believe, however, that more randomness tests are required. This is left for future work.

References

1. Al-Muhammed, M.J., Abuzitar, R.: Dynamic text encryption. Int. J. Secur. Appl. (IJSIA) **11**(11), 13–30 (2017)
2. Bogdanov, A., Mendel, F., Regazzoni, F., Rijmen, V.: ALE: AES-based lightweight authenticated encryption.In: Moriai S. (ed.) Fast Software Encryption. FSE 2013. Lecture Notes in Computer Science, vol. 8424. Springer, Heidelberg (2014)
3. Knuden, L.R.: Dynamic encryption. J. Cyber Secur. Mobility **3**, 357–370 (2015)
4. Mathur, N., Bansode, R.: AES based text encryption using 12 rounds with dynamic key selection. Procedia Comput. Sci. **79**, 1036–1043 (2016)
5. Daemen, J., Rijmen, V.: The Design of RIJNDAEL: AES-The Advanced Encryption Standard. Springer, Berlin (2002)

[2] The maximum number of binary sequences that are expected to fail at the level of significance α is computed using the following formula [17]: $S.(\alpha + 3.\sqrt{\frac{\alpha(1-\alpha)}{S}})$, where S is the total number of sequences and α is the level of significance.

fridaya@uef.fi

6. Nie, T., Zhang, T.: A study of DES and blowfish encryption algorithm. In: Proceedings of IEEE Region 10th Conference, Singapore, January 2009
7. AL-Muhammed, M.J., Abuzitar, R.: κ-lookback random-based text encryption technique. J. King Saud Univ. Comput. Inf. Sci. (2017). https://doi.org/10.1016/j.jksuci.2017.10.002
8. Patil, P., Narayankar, P., Narayan, D.G., Meena, S.M.: A comprehensive evaluation of cryptographic algorithms: DES, 3DES, AES, RSA and blowfish. Procedia Comput. Sci. **78**, 617–624 (2016)
9. NIST Special Publication 800-67 Recommendation for the Triple Data Encryption Algorithm (TDEA) Block Cipher Revision 1, Gaithersburg, MD, USA, January 2012
10. Bogdanov, A., Mendel, F., Regazzoni, F., Rijmen, V., Tischhauser, E.: ALE: AES-based lightweight authenticated encryption. In: Moriai, S. (ed.) FSE 2013. LNCS, vol. 8424, pp. 447–466. Springer, Heidelberg (2014)
11. Stallings, W.: Cryptography and Network Security: Principles and Practice. 7th edn. Pearson (2016)
12. Anderson, R., Biham, E., Knudsen, L.: Serpent: a proposal for the advanced encryption standard. http://www.cl.cam.ac.uk/~rja14/Papers/serpent.pdf. Accessed Feb 2018
13. Burwick, C., Coppersmith, D., D'Avignon, E., Gennaro, R., Halevi, S., Jutla, C., Zunic, N.: The MARS encryption algorithm. IBM, August 1999
14. Online Random Key Generator Service. https://randomkeygen.com
15. Soto, J.J.: Randomness testing of the AES candidate algorithms. http://csrc.nist.gov/archive/aes/round1/r1-rand.pdf. Accessed May 2018
16. Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A., Dray, J., Vo, S.: A statistical test suite for random and pseudorandom number generators for cryptographic applications. NIST special publication 800-22, National Institute of Standards and Technology (NIST), Gaithersburg, MD (2001)
17. Soto, J.: Randomness testing of the advanced encryption standard candidate algorithms. NIST IR 6390 (1999)

fridaya@uef.fi

Photos and Tags: A Method to Evaluate Privacy Behavior

Roba Darwish[1(✉)] and Kambiz Ghazinour[2]

[1] Computer Science and Engineering Department,
Colleges and Institutes Sector of Royal Commission Yanbu,
Yanbu University College, Yanbu, Kingdom of Saudi Arabia
darwishr@rcyci.edu.sa
[2] Department of Computer Science, Kent State University, Kent, OH, USA
kghazino@kent.edu

Abstract. Online Social Networking Sites attracted a massive number of users over the past decade but also raised privacy concerns with the amount of personal information disclosed. Studies have shown that 25% of the users are not aware of privacy settings provided by these sites or do not know how to change them. This paper investigates an approach towards understanding users' privacy behavior on social media, e.g. Facebook, through studying faces, tags and photo privacy settings. It classifies users based on their privacy selections and proposes a system for monitoring and recommending stronger privacy settings. An application is developed, and our case study examines the effectiveness of our model.

Keywords: Social media · Privacy · Face · Tag · Classification

1 Introduction

Social networking sites have become very important in our lives since their inception. They have revolutionized the world of communication, as they allow individuals to communicate with peers all over the world. However, the popularity of social networking sites presents the growing dilemma of preserving users' privacy.

Each social media site provides users with privacy settings. However, studies have shown [1, 2] that around 25% of social media users lean toward not to change, or are not by any means mindful of the service's privacy settings. As these services have become very popular over the past few years, more privacy risks are faced by users.

As of the first quarter of 2017, Facebook had 1.94 billion monthly active users, making it the largest social media service in terms of active users, and this number continues to grow [3]. Such prolific use was a prime motivation to investigate privacy issues and concerns on Facebook.

Users disclose their personal information, photos of themselves, their family and friends, ignoring the consequences of such behavior on their privacy.

This work inspects the use of faces and the presence of tags as a measurement tool to evaluate users' privacy behavior. Subsequently, new privacy categories are introduced and added to the existing categories defined by Westin [4]. Moreover, we propose an application for monitoring and recommending better privacy settings for

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 797–816, 2019.
https://doi.org/10.1007/978-3-030-22868-2_55

fridaya@uef.fi

users. Machine learning techniques are also utilized in order to comprehend privacy settings of various users and suggest stronger settings for them.

This research contributes to the area of privacy in social media by:

- By using face detection, tags and their location on photos posted by individuals, a new method is proposed to measure behavior of privacy of the individuals.
- Based on the proposed new method, new privacy categories (besides existing Westin's privacy categories) are introduced.
- An application that monitors privacy settings for users and screens the privacy risks in their profiles. It then educates social media users of privacy-related issues, helping users to avoid them when using social media networks.

This recommender system provides benefits for both the user and the researcher. It suggests a better privacy setting and provides the user with the necessary steps on how to set up stronger privacy controls. It also has the ability to aid researchers by helping with data collection and the study of user demographics. Such information would then be used to design a social media network in which users are more aware of privacy settings, as well as future research in the field.

2 Background

2.1 Privacy on Social Media

Social networking sites have been developed significantly in recent times. Users aim to create many friends and connections. However, social media users are risking their privacy when utilizing such sites by disclosing a huge amount of personal information, images, messages and others, thus raising many privacy-related concerns.

By using social networking sites [6, 7] users expose themselves to many risks that significantly affect their privacy. Studies [6, 7] find that privacy could be attacked in a few ways if personal information was not provided reasonably and reliably.

Eecke and Truyens [7] discuss that there is an evidence that once published, removing a "defamatory" or unwanted data from the Internet is nearly impossible. Moreover, they found that while different social network sites provide information protection tools to their users, the average user lacks an understanding of them, let alone properly using them. As mentioned earlier, most users tend not to change the default configuration set to make information publicly accessible. Subsequently, identity theft, stalking, physical harm, and other risks are increased.

2.2 Privacy on Facebook

Using Facebook, millions of individuals can create online accounts and share information with an enormous network that includes both friends and even strangers [8]. Facebook users usually disclose personal information, photos, and details about themselves and others. As the information disclosed increases, the risk of privacy violation increases as well [9]. Facebook content that is marked as "public" or with a little earth symbol can be seen by almost everyone over the internet, including

fridaya@uef.fi

criminals, strangers and unwanted people [10]. They can discover many things about you, from what you like to your everyday activities and places you go. For instance, posting daily routes on Facebook or checking into nearby location of places you visit makes you an easy target for stalkers [11]. Other privacy risks result from offering your information to promoters. Moreover, this can be extended to creating a database gathered from individuals' profiles; this is done by third-parties who then sell the data to others. Passwords or even the whole database might also be stolen by intruders.

2.2.1 Tagging. Facebook allows users to add tags to their photos. Tags can be added to particular areas of your photo. For example, users can tag friends on their faces in a photo taken in a restaurant, tagging the restaurant with its name. According to the Facebook Help Center [12], the post that someone is tagged in it, can also be seen in that person's timeline. Hence, anyone who sees the posts can go to that person's profile to gain more information. To be more accurate, if the person tagged has a Facebook account, the tag is displayed as a hyperlink to that person's profile. If the tagged person does not have an account, then the tag is shown in plain text, not linking to any profile. So even if a user is concerned about privacy and has taken steps to protect his or her data by setting the privacy album visible to friends only, someone who is not a friend can still access that tagged photo and display the picture. Besides, friends of friends can also view that photo and find more information about the user. For example, it is easy to predict the age of a person through searching tagged Facebook photos and then checking posted birthday photos [13].

2.2.2 Facebook Privacy Settings. Facebook provides different privacy settings for users. Users can determine who can see their personal information. They can also set to whom their photo albums can be visible (e.g. Public, Friends, Friends of Friends, Custom, and Only Me). Table 1 shows privacy settings in Facebook.

Table 1. Privacy settings available for Facebook

Setting	Description
Public	When Facebook content is public, anyone can see it, even people you are not friends with, on or off Facebook
Friends	Only those who are your friends on Facebook can access the content provided
Friends of Friends	Your friends on Facebook and their friends can access the content provided
Custom	Selective audience, specific people and networks can access the content provided
Only Me	When you select the only me setting, only you can see the content you post

Unfortunately, most users do not know how to use these privacy settings [14]. They do not know how the settings work or simply do not have time to read them. This may account for most users leaving their privacy controls to the default setting, set to be visible to the public. According to Facebook, if the content is public, anyone over the internet can view it, including people who are not your friends and even those who do

fridaya@uef.fi

not have a Facebook account. In a survey of 200 Facebook users that aims to measure the disparity between the desired and actual privacy settings [14], researchers find that 36% of Facebook content remains shared with the default privacy setting. Likewise, they find that privacy settings match the users' desire only 37% of the time, and when incorrect, almost always expose content to a bigger number of users than anticipated.

When it comes to tagged photos, the default setting is that the tag is added to the Timeline and shared with the audience with no option available to the tagged person to review the content first [15]. This actually raises many privacy concerns.

It is worth noting that even if the selected setting for visibility is at "friends" level and the user has tagged someone in the post, the tagged person and their friends can view it. Similarly, the tagged person and their friends can see the post when the "only me" option is selected. In this work, the only tags considered are those found in images visible to the "public" and "friends of friends". Some of the user's information is always public on Facebook and users have no choice to change its setting. This information is listed in Table 2, along with the reason why Facebook makes it publicly visible. These reasons are provided by Facebook [12].

Table 2. Public information on Facebook

Public information	Reason
Name, Profile picture, cover photo	Enables others to identify you
Gender	Helps Facebook when referring to you (e.g. "add him as a friend")
Networks	Enables others to find you (e.g. suggest friends who went to same school as yours)
Username and user ID	Displayed in the URL of any user account
Age	Helps to give the user information that is appropriate for user's age
Language and country	Enables Facebook to give the user a proper material

In this work we do not use face recognition methods to recognize celebrities and discard the photos that have pictures of people whose information is widely and publically available and posting them online would not necessarily violate their privacy (e.g. posting a picture of an ad in which a singer is promoting his/her concert).

3 Related Work

3.1 Recommender Systems

Many researchers have recognized the problem of controlling privacy settings in social networking sites. They notice that people disclose a lot of information through social sites with little to no privacy context [16]. As a result, they propose recommendation systems in order to assist users to easily configure privacy settings. This section discusses some of recommender systems currently existing in the field.

fridaya@uef.fi

In [17], the work presents a system which allows users to display the current privacy settings on their Facebook profile. It also detects possible privacy risks. The system monitors the privacy settings of user profiles and then recommends a setting for the user by using machine learning techniques to look at the similarity of the preferences chosen by the user who desires to set the privacy setting with other users who share common preferences.

Ghazinour et al. [5] are interested in the users' personal profiles, users' interests and users' privacy settings on photo albums to see whether they are visible to the public, friends, friends of friends, or on a custom setting. By observing how different users choose their settings on photo albums, the researchers classify them into one of three privacy categories. These privacy categories are specified by Alan Westin as Fundamentalists, Pragmatic and Unconcerned [4]. A detailed discussion is given in section B. Later, the Decision Tree is used to find different profile types. When it comes to recommendations, the K-Nearest Neighbor (KNN) algorithm was applied to predict the class of a new Facebook user by finding the similarities between their profile and others. Based on this, a recommendation is given by the system. The use of KNN makes this classifier like a collaborative filtering recommender system [18].

Ghazinour et al. [5] analyze users' data in order to understand their behavior in terms of how they choose their privacy settings. They found that most of the users shared information about their age, gender and education. However, when it comes to religion, political views and degree, users were more conservative. This study also shows a relationship between the users' interest and how they choose their own privacy settings. For instance, if the user's age is less than 21 years old, he/she usually belongs to the unconcerned category.

Mehatre and Chopde [16] propose a Privacy Policy Prediction system to assist users to compose privacy settings for their shared images. Their system relies on an image classification framework for image categories that might be linked with similar policies, and on a policy prediction algorithm to generate a policy for each newly uploaded image. This is done based on the user's social features. In summary, their proposed methodology is as follows: (a) User uploads an image which has both objects and background. (b) The object will be extracted and the background is suppressed to improve classification accuracy using foreground features. A saliency map, which is a kind of image segmentation, is used to help extracting image features. (c) The extracted features are then compared with the database features of images. (d) KNN is used to classify the class of the newly uploaded image. (e) A policy comparison from the database is done by using a linear matching technique. (f) Policy is accepted by the user.

Li et al. [8] present a trust-based privacy assignment system for social sharing that uses resources in social object networks. The presented system helps people select the privacy preference of the information being shared. This system, called the Personal Social Screen (PerCial), assists in assigning a privacy preference by automatically generating topic-sensitive communities users are interested in. It detects a two-level topic-sensitive community hierarchy before assigning a privacy preference to the users, depending on their personalized trust networks.

Ginjala et al. [19] introduce an intelligent semantics-based privacy configuration system (SPAC). This system automatically recommends privacy settings for social network users. SPAC uses machine learning techniques on both the privacy setting

fridaya@uef.fi

history and users' profiles to learn configuration patterns and then make predictions. In this system, semantics are integrated into the KNN classification to increase the accuracy of recommendations, such as semantics information in users' profiles.

Shripad and Vaidya [20] introduce a framework for handling trust in social networks based on a reputation mechanism. This mechanism works by capturing the implicit and explicit connections between the network users. It analyzes the semantics and dynamics of these connections. The system then provides personalized positive and negative user recommendations to another network user. The positive recommendations assist in connecting trustworthy users while the negatives alert users to not connect to untrustworthy users.

3.2 Westin's Privacy Categories

In our work users are classified into one of three privacy categories based on the setting chosen for their photo albums. These privacy categories are specified by Alan Westin, a researcher who conducted over 30 privacy surveys in the 1970s, as Fundamentalists, Pragmatic and Unconcerned. Privacy researchers around the world have used Westin's privacy index to measure attitudes and categorize people into these three groups. Descriptions of Westin's groups are provided below [21, 22].

Privacy Fundamentalist: Privacy Fundamentalists are highly concerned about their information. They are unwilling to provide any data or reveal information about themselves on websites when requested. People of this category tend to be worried about the precision of automated data and the additional uses made from it. They are agreeable to new laws that clearly explain an individual's rights and privacy policies. Privacy Fundamentalists form about 25% of the public.

Privacy Unconcerned: Privacy unconcerned people are willing to reveal and disclose any information upon request. They are the least protective of their privacy. Moreover, they do not favor expanded regulation to protect privacy. About 18% of the public are unconcerned.

Privacy Pragmatists: Privacy pragmatists are willing to disclose their information if they gain some benefits in return. Initially, they measure the potential advantages and disadvantages of sharing data with the organization. Next, they measure what protections are available for them and their trust level in the organization. Afterwards, they make a decision on whether or not to reveal their information to them and if revealing that information is actually worth it. 57% of people belong to this category.

4 Our Approach

Facebook currently owns the largest archive of personal photos. Thus, sharing and tagging images is actually built around real identities [23]. It is clear that the presence of faces and tags have a wide impact on the privacy of users. A photo showing a face of someone with a tag placed on the face could display the person's name, add a link to their profile for even more detailed information.

fridaya@uef.fi

4.1 Privacy Levels

In this study, all conceivable instances of human faces and tags on user's photos are covered with the aim of categorizing them into various classes based on privacy categories explained in the next section. The following are different cases considered:

4.1.1 CASE 1: No Faces, No Tags. Photos that do not include faces and tags are considered to be the cases in which least privacy violation occurs. At the point when there is neither a face uncovering the identity of the individual nor a tag connecting directly to the user profile, no privacy violation is implied. However, cases exist that reveal information about the user whether a photo contains faces and tags or not. For instance, a user may post a photo of a landscape, in spite of the fact that there are neither faces nor tags, that photo may uncover other information about the user's present location.

4.1.2 CASE 2: Some Faces, No Tags. Having human faces in a photo may violate the privacy of an individual even if no tags are found. If the person is identified, information regarding the identity of the person is revealed and data with respect to the character of the individual is uncovered. This case is regarded as more privacy-revealing than Case 1.

4.1.3 CASE 3: No Faces, Some Tags. As mentioned earlier, tags link your profile to the post shared, making it easier for others to access your profile, view your information and discover more about you. This case is considered to disclose more private information than Case 2 (faces without tags) because through a profile, additional information can be uncovered, making it more valuable for the adversary than just knowing the person's face. For example, a friend of a friend can click on the tag link and view personal information such as location, hometown, relationships, and other information. Moreover, that user can view photos as long as they are visible to public or friends of friends. Keep it in mind that the profile picture, cover photos, and basic personal information are always visible to the public as clarified in Table 2.

4.1.4 CASE 4: Some Faces, Some Tags. The fourth case is when both faces and tags exist in a photo. In this case, whether the tag was placed on the face or not is examined. Answering this question will help in determining the privacy category of the user. Thus, if the tag is on the face, this facilitates in linking the tag name directly to the user, identify the user and get additional information from the user profile. If not, extra efforts were required in order to map the tag to the correct face in the photo which can sometimes be difficult. For instance, a photo that has a large number of faces and tags that are not placed on these faces makes it difficult for others to map the tag to the right face. However, this is still considered a violation of privacy.

This case incorporates three different cases. These cases are listed below:

(a) No. of Tags < No. of Faces
(b) No. of Tags = No. of Faces
(c) No. of Tags > No. of Faces

Those cases (with the same order in which they are mentioned above) and the tag not placed on the face, come next in privacy levels. The reasons for the order chosen are mentioned below with examples given to illustrate each case.

fridaya@uef.fi

Tag Not on Face

No. of Tags < No. of Faces. When the number of faces is greater than the number of tags, this might mean that there are faces existing on the photo that have no relation to any of the tags located in the photo. There can be faces of individuals with no names attached to it or links to their accounts. This refers to Case 2 (faces with no tags). Additionally, because these tags are not placed on the faces, there is minimal privacy risk involved compared to the case when the tags are positioned exactly on faces. To simplify this, it can be said that the tag does not directly relate the face to the person's name or the person's profile. For example, a photo which has three faces and one tag not placed on the person's face is less privacy-revealing than a photo having three faces and three tags placed on the faces. Figures 1, 2, 3, 4, 5, and 6 illustrates all the possible cases.

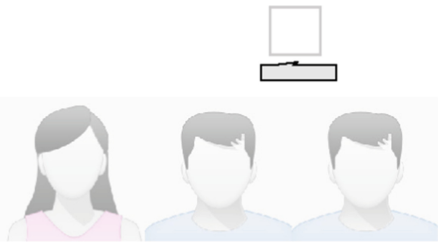


Fig. 1. A photo with less no. of tags than faces with tag not on face.

No. of Tags = No. of Faces. When the number of tags is equal to the number of faces, assigning a tag to its related face would be much easier in case those tags belong to the existing faces on the photo. It would be much easier for adversary to map the person's face with the tag name and the attached user account. Figure 2 shows an example of having a photo with 3 faces and 3 tags with tags not placed on the faces.

No. of Tags > No. of Faces. Some tags may not belong to any of the faces found in the picture. It is difficult to map tags to their corresponding faces, minimizing privacy violation. Most tags not placed on faces usually do not belong to the faces in the picture. However, in this case a user can discover more about other people from the tags available (as mentioned earlier, having more tags than faces violate more privacy).

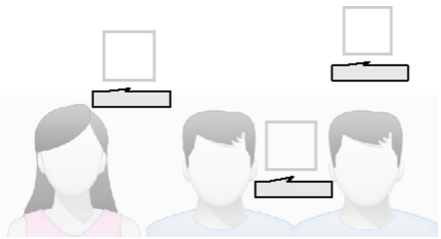


Fig. 2. A photo with equal no. of tags and faces with tag not on face.

fridaya@uef.fi

For instance, if you have three faces and four tags in a photo as illustrated in Fig. 3, the chance of finding personal information of others through their profile is higher.

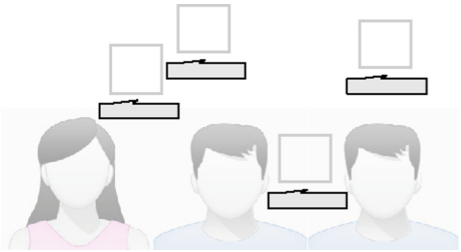


Fig. 3. A photo with greater no. of tags than faces with tag not on face.

Tag on Face

(a) ***No. of Tags < No. of Faces.*** As discussed earlier, having tags located on the person's face reveals a lot. By accessing a profile, adversary can easily identify a person's identity, appearance, and other revealing information such as personal information, photos, activities and much more.
Figure 4 shows three faces with one tag located on one of the faces. By looking at the picture, it is easy to figure out who the person is, as well as what he or she looks like. When it comes to other faces existing on the picture, privacy is also violated even if no tags are attached to them, albeit with a lower risk.

Fig. 4. A photo with less tags than faces with tag located on face.

(b) ***No. of Tags = No. of Faces.*** In this case, the tags directly relate the person's face with his/her name and profile account. This reveals much about people found in the photo with no effort needed to map the tags to the correct person. In other words, information about every person found in the photo is easily available. Figure 5 shows three faces with three tags located on them to illustrate the idea.

Fig. 5. A photo with equal number of tags and faces with tag located on face.

fridaya@uef.fi

(c) ***No. of Tags > No. of Faces***. The most privacy revealing case compared to all other cases is having more tags than faces, with these tags located on the faces. In addition, since these tags are located on the face, the privacy risks are higher. Figure 6 shows a picture with three faces having tags located directly on them. Other tags also exist that reveal information about other people. Figure 7 concludes the cases as the higher up the case, the less private information can be revealed.

Fig. 6. A photo with more tags than faces with tag located on face.

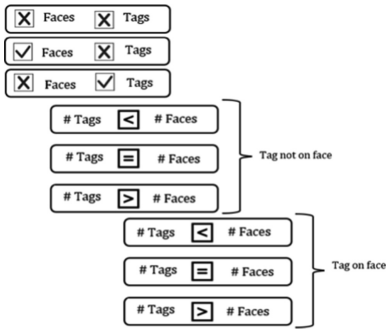


Fig. 7. Privacy levels of all cases covering faces and tags.

4.2 User Classification, New Privacy Categories

As mentioned in Sect. 3.2, Westin categorizes people into three different groups: Fundamentalists, Pragmatics and Unconcerned. Ghazinour et al. [5] studied users' attitudes towards sharing their photo albums to determine the privacy category they belong to. The authors mention: "We also use photo albums since users treat them as a very personal and tangible type of personal identifiable information. Furthermore, it is one of the data items that Facebook allows us to check its privacy settings using the Facebook API functions [5]".

fridaya@uef.fi

The authors use the following rules to determine the privacy category of each user for the profiling phase:

If # of photos shared == 0 then:
    The user's privacy_category = **Fundamentalist**.
Else if ratio of photos visible to Public or Friends of Friends < %50 then:
    The user's privacy_category = **Pragmatic**.
Else:
    The user's privacy_category = **Unconcerned**.

In their work [5], Ghazinour et al. acknowledged that using the setting of photo albums as an indicator of unconcerned users may not be enough. They mentioned that having pictures of nature or art works in albums, which are set to be visible to public, do not imply any privacy violation which led us to take faces and tags in the pictures into account. In order to examine the model used in this study, as a first step, users were classified into one of five categories based on **faces** found in their photos. Three are the main categories derived by Westin. In addition, two new categories were introduced:

Fundamentalist-Pragmatic (FP): Fundamentalist leaning toward pragmatic. This group of people shares little data about themselves. The revealed information is less likely to violate any privacy. For instance, users who set their photo albums to be visible to public ensure that those photos shared do not contain any faces.

Pragmatic-Unconcerned (PU): pragmatic leaning toward unconcerned. This group of people shares some data about themselves which also shows highly revealing information. An example of this is a user having only 10 photos in his profile, but with each photo containing many faces including the user and his friends.

The following rules are used to determine the privacy category of each user:

If # of photos shared == 0 then:
    The user's privacy_category = **F**
Else if all photos have faces then:
    The user's privacy_category = **U**
Else if # of faces in all photos == 0 then:
    The user's privacy_category = **FP**
Else if # faces < %50 of photos then:
    The user's privacy_category = **P**
Else if # faces >= %50 of photos then:
    The user's privacy_category = **PU**

Second, users were classified based on both **faces** and **tags** into one of the seven categories using the following rules. New classes introduced here are P+ , P and P− where P+ is more privacy preserving due to lack of tag use compared to P and P−.

fridaya@uef.fi

```
If # of photos shared == 0 then:
    The user's privacy_category = F
Else if (# of faces == 0) AND (# of tags == 0) then:
    The user's privacy_category = FP
Else if (# of faces != 0) AND (# of tags == 0) then:
    The user's privacy_category = P+
Else if (# of faces == 0) AND (# of tags != 0) then:
    The user's privacy_category = P
Else if (# of faces != 0) AND (# of tags != 0) then:
     If (# of tags < # of faces) Then:
         The user's privacy_category = P-
     Else if (# of tags == # of faces) then:
          The user's privacy_category = PU
     Else if (# of tags > # of faces) then:
          The user's privacy_category = U
```

The difference between having the tag placed on the face or not was acknowledged; for simplicity, when classifying users, they were treated equally in the above rules.

There are many other ways possible to label the instances such as sorting based on the number of tags, etc. However, this method was chosen in order to show the importance of combining faces and tags in studying privacy behavior. Interested researchers are invited to select their desired method for labeling.

4.3 Implementation

In this study, we present a Facebook application enables researchers to collect information from Facebook users, creating a dataset. The application is hosted on a virtual machine located in our lab. This system is designed as a Facebook application in order to collect information of user's personal profile, user's interests or likes and user's privacy settings on photo albums.

The Facebook app is written in JavaScript and PHP to access the Facebook user`s profile and settings. The system uses JavaScript SDK which provides a rich set of client-side functionality to access Facebook's server-side API calls.

4.3.1 Face Detection. Microsoft Cognitive Services (formerly Project Oxford) is utilized to detect human faces in user images visible to either public or friends of friends. The API detects human faces in an image as input, returning face rectangles for where in the image the faces are in the output.

4.3.2 Tag Location. Facebook API provides the following: (a) The time the tag was created. (b) Tagging person, representing the user who added the tag. (c) X and Y coordinates in the photo where the tag is. (d) Names of friends tagged in each photo.

In our application, getting the x and y coordinates of the tags was of interest so that the coordinates returned by Microsoft Face API for the face could be compared with the coordinates returned by Facebook API for the tags. This is done in order to

fridaya@uef.fi

determine if the tag is on the face or not. For educational purposes, the app also displays the tagged names to the user.

4.3.3 System Database. User data collected through Facebook API is stored in a secure database. In this work, phpMyAdmin, one of the most well-known applications for MySQL databases administration, is used. In this database, a table was created to store personal information such as age, birthday, gender, hometown, location, political view, relationship status, religion, education, degrees, etc. Lookup tables were designed for maintaining data integrity in our database environment. For instance, if a user is entering his/her relationship into a data item, the User Data table containing the relationship item can reference a lookup table to verify that only one of the specified values is inserted. However, in tables such as location, the values are inserted in lookup tables in case they did not exist before; this is done by comparing the location ID with the IDs existing in the table.

4.3.4 Users' interests. Such as music, movies, TV shows, and books they like are stored, as well as information related to user's photos, like the number of faces and tags in each image and whether the tag is placed on the face (if any). These information where collected to help us with the classification process in which users with the same interests would be grouped together.

4.3.5 Application Interface. Users run the study's application and give permission to it to access their personal information disclosed on their profile. That user will then receive a report from the application that assists him/her in selecting a better privacy setting. This report summarizes the information provided to Facebook and, in a nutshell, notifies the user of his/her profile's privacy risks, suggesting a tighter setting. Figure 8 shows the report generated for the user where personal information, interests, publicly visible albums, number of faces detected, number of tags are displayed.

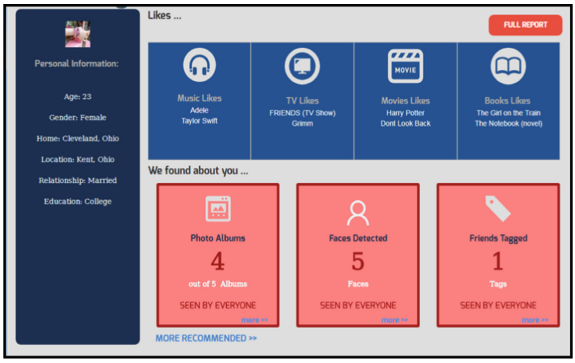


Fig. 8. The app interface where user's report displays personal information, interests and photo albums results.

fridaya@uef.fi

5 Data Collection and Analysis

5.1 Research Approvals

Before beginning data collection, the following approvals were obtained:

Institutional Review Board (IRB). Since the research involves collecting data from human subjects, it was necessary to submit the research proposal and supporting documents to IRB for approval. The IRB is responsible for protecting the rights and privacy of participants. Research conducted in this study using human participants was approved as a Level II/Expedited, category 7 project.

Facebook approval. Before releasing an application, Facebook needs to approve it. During the approval process, permissions required are submitted to Facebook through the App Dashboard. Some of the permissions submitted for review are: public_profile, user_birthday, user_education_history, user_hometown, user_location, user_relationships, user_religion_politics, user_likes and user_photos. Detailed instructions on how these permissions are used on the app, as well as screenshots, is also required.

5.2 Facebook User Data

In this study, it was of interest to collect the following user's data:

- *User's personal profile*: The attributes were collected from the user profile are: gender, birthday, education history, hometown, location, political view, relationship and religion.
- *User's interests*: This included books that users like to read, music they like to listen to, TV shows, and movies they like to watch.
- *User's privacy settings on photo albums*: Names, privacy settings, and cover photos of each album were collected.
- *User's photos*: URLs of photos in albums visible to public or friends of friends.
- *User's tags*: User's tagged friends in each photo of the publicly visible albums and friends of friends were collected.

The above parameters assisted in building a set of user data. When collecting user's photos, only recent photos are added. Using Facebook Graph API, the researchers were able to get the most recent ten photos of five albums posted within one year.

5.3 Data Harvesting and Participating Users

Users were recruited using Amazon Mechanical Turk (AMT). AMT is a service offered by Amazon that enables researchers to hire workers to perform human intelligence task. Workers were offered 1 dollar for running the app. 200 participants from 15 different countries with different educational backgrounds (median age 33; range 14–73) were asked to run the app at: [URL MASKED FOR BLIND REVIEW]

Participants were informed that the information they share will remain strictly confidential and will be used solely for the purposes of this research. Photos and images will not be stored. The data collected from them may be used verbatim in

fridaya@uef.fi

presentations and publications but the users themselves will not be identified. Results will be published in a pooled (aggregate) format. Moreover, the collected data is stored in a password-protected database at our lab where only the Principal Investigator and the co-Investigator have access to that database. Participants were informed that they needed to only run the app once with no need to answer any questions.

5.4 Data Preprocessing

Prior to running machine-learning algorithm for classification, it was necessary to prepare the collected data for further processing, including some preprocessing which was performed by the two Authors manually. A few issues faced and addressed were the following: Having multiple values referring to the same input. For instance, in the religion attribute, having values such as (Islam, islam, MUSLIM, Muslim, مسلم, مسلمة) referring to the same thing needed to be manually fixed so they all referred to one value "e.g. Muslim". This also applies for entries from different languages. For education attribute, values such as (Grad, Graduate School) needed to be changed to refer to the same value "e.g. Graduate". Changing location and hometown attributes from (city, state) or (city, country) format to (country) only limited the values we have (e.g. changing "Boston, MA" to "USA"). Entries such as (Christianity, Christian-Catholic) are also changed to their root value.

5.5 Data Analysis

After the preprocessing phase, data collected from participants are analyzed. When it comes to personal information such as gender, relationship, education, degree, location, hometown, religion and political view, users are comfortable disclosing some attributes but not others. The red color in the illustrated bar charts from Figs. 9 and 10 indicates the percentage of missing (not provided) values for the corresponding attribute. All participants provided their Gender with 0% of missing values. On the other hand, the highest percentage of missing values belonged to the Degree attribute. Table 3 shows the percentage of missing values for each attribute.

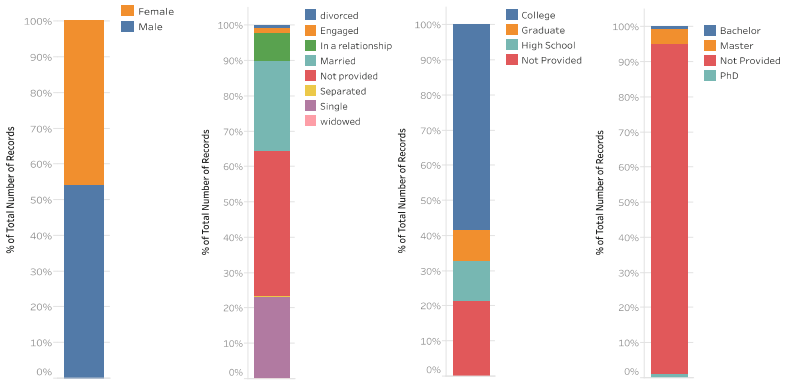


Fig. 9. Percentage of total number of records for gender, relationship, education and degree attributes (left to right).

fridaya@uef.fi

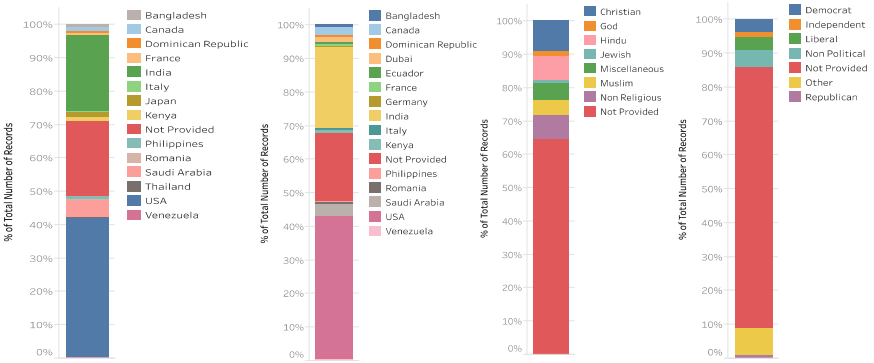


Fig. 10. Percentage of total number of records for hometown, location, religion and political view attributes (from left to right).

Table 3. Percentage of values users did not disclose in this study

Attribute	% Missing
Degree	94%
Political view	77%
Religion	65%
Relationship	41%
Hometown	23%
Location	20%
Education	21%
Gender	0%
Age	0%

Concerning albums results collected from participants, the maximum number of albums is 25, while the average number visible to public is 10. Participants with an age range from 40–70 have the maximum number of 11 albums as the total number of albums they share. Moreover, participants with the age range of 14–29 have the maximum number of 50 faces as the total number of faces they share. When it comes to the total number of tags, the maximum is 23. The maximum number of faces and tags for participants with the age range: 30–70, are 28 and 10, respectively.

While both genders tend to share similar proportions in relation to the number of faces, albums, and public albums, there is a significant difference when it comes to the number of faces: males tend to share more faces than females (see Fig. 11).

Almost 32% of the users belong to the P+ class. This means that most of the users shared faces with no tags. 24% of the users belonged to the fundamentalist pragmatic category, thus sharing pictures with no faces or tags. 15% of participants were fundamentalists, while only 4% fell into the unconcerned category.

fridaya@uef.fi

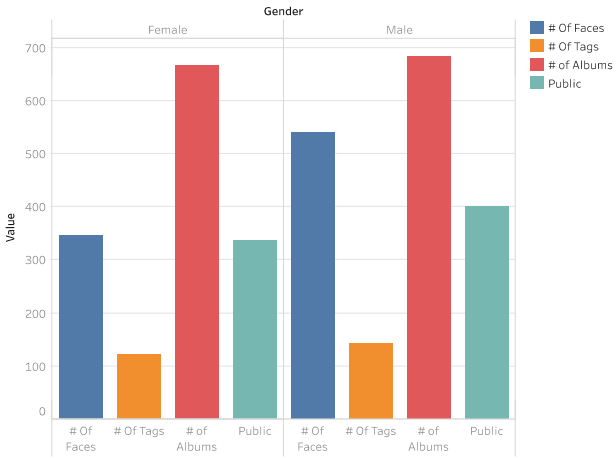


Fig. 11. Total no. of faces, tags, albums and public albums for both genders.

5.6 User Classification

Using the rules discussed in Sect. 4.2, the Decision Tree algorithm in WEKA (J48) was used to categorize users into the following classes after labeling them:

- *3 Classes*: F, P and U.
- *5 Classes*: F, FP, P, PU and U.
- *7 Classes*: F, FP, P+ , P, P−, PU and U.

When running Decision Tree algorithm on seven classes, the percentage of correctly classified instances is 87.85% using percentage split.

5.7 Recommender System

The KNN classifier is used to recommend a better privacy setting for the user. KNN ran on the 3 classes, 5 classes and then 7 classes of privacy behavior explained earlier. The goal is to compare the results in order to find the one with the highest prediction rate and thus the highest number of correctly classified instances. In all three different types of classification, participants are classified based on personal information disclosed on their profile. These are age, gender, hometown, location, relationship, religion, political views, education and degree. See [5] to understand how the recommender system works and suggests a less privacy revealing setting.

KNN is used with K = 4 for three different attributes. Those are education, location and relationship. A 66% Percentage split (66% training data and 34% test data with random split of dataset) is used as a test option. Results for different 3 classes, 5 classes and 7 classes of privacy behavior are shown in Tables 4, 5, and 6, respectively.

From the results shown above, the percentage of correctly classified instances for 7 classes of privacy behavior is not as strong as using 5 classes.

fridaya@uef.fi

Table 4. Three classes of privacy behavior for three attributes

Attribute	Education	Location	Relationship status
Correctly classified instances	77.61%	79.10%	70.15%

Table 5. Five classes of privacy behavior for three attributes

Attribute	Education	Location	Relationship status
Correctly classified instances	83.85%	79.10%	59.70%

Table 6. Seven classes of privacy behavior for three attributes

Attribute	Education	Location	Relationship status
Correctly classified instances	80.59%	83.58%	77.61%

5.8 Results and Main Findings

A study is conducted using the data collected from participants. The main results of the study are displayed below:

Users are comfortable disclosing their age and gender. They are more conservative when it came to sharing their location and hometown. More than 60% of participants are not willing to share their religions and political views. 94% of participants do not disclose their degree. Participants under the age of 30 reveal twice as many faces and tags in their posted photos than others. When the participants are classified into 7 categories, a large percentage of users are found to share many faces but no tags.

When recommending a better (less revealing) setting using KNN, the three and five Privacy classes give a better overall prediction rate than the seven classes.

6 Conclusion and Future Work

The rapid growth of social networking sites has a negative impact on the privacy of the individual. Although privacy settings are available at these sites, they are not used to serve the user because the user is unaware of them, they are difficult to control, or the user simply does not care about them.

In this work, we introduced an application where privacy settings set by Facebook users are monitored and better settings are recommended. The application displays a report to the user where major privacy holes detected in their profile are shown. A new method for evaluating privacy issues by considering existence of faces and tags on publicly visible photos posted by Facebook users is introduced. Based on the new method, new privacy categories are introduced to Westin's privacy categories.

In this study, data of 200 participants from 15 different countries are collected. Participants are classified into three, five and seven privacy categories using a decision tree algorithm. Finally, KNN (with K = 4) is used to recommend a tighter setting.

Limiting the number of photos (recent photos only) to evaluate for faces and tags has advantages such as reducing the expected waiting time to load the user report.

fridaya@uef.fi

However, this might affect the accuracy of measurement when classifying the user into various categories, since other photos might contain more faces and tags.

For future direction, processing all users' photos using a faster face detection method/algorithm is suggested. Furthermore, faces of famous people should be excluded from the face detection report. Another suggestion for future work is to analyze comments on posts since they reveal user information that may affect user privacy. For example, posting a beach image might not reveal anything, but a comment saying, "Relaxing in South Beach, Miami" can reveal where the user is her home is empty.

Finally, future work will concentrate on conducting further experiments considering a larger number of users. This is also needed in order to have a bigger training set, eventually resulting in better predictions. Moreover, in terms of interface design the aim is to display faces that have been detected in the user's photos; this can be used to warn the user and draw attention to the revealed people. Additionally, studying the impact of the application in terms of the amount of awareness raised is also necessary. It is of interest to monitor the main privacy settings that they prefer to change to both before and after running the application.

References

1. Widerlund, J.: Social Media: The Privacy and Security Repercussions. Search Engine Watch, 19 June 2010
2. Consumer Reports Magazine: Facebook and your privacy. CR Consumer Reports, NY
3. Constine, J.: Facebook now has 2 billion monthly users … and responsibility, 27 June 2017
4. Kumaraguru, P., Cranor, L.F.: Privacy Indexes: A survey of Westin's Studies. CMU-ISRI-5-138, Pittsburgh, PA, December 2005
5. Ghazinour, K., Matwin, S., Sokolova, M.: YOURPRIVACYPROTECTOR, A recommender system for privacy settings in social networks. Int. J. Secur. Priv. Trust Manag. (IJSPTM) **2**(4), 11–25 (2013)
6. Senthil Kumar, N., Saravanakumar, K., Deepa, K.: On privacy and security in social media – a comprehensive study. Procedia Comput. Sci. **78**, 114–119 (2016)
7. Van Eecke, P., Truyens, M.: Privacy and social networks. Comput. Law Secur. Rev. **26**(5), 535–546 (2010)
8. Li, L., Sun, T., Li, T.: Personal Social Screen—A dynamic privacy assignment system for social sharing in complex social object networks. In: 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, pp. 1403–1408 (2011)
9. Debatin, B., Lovejoy, J.P., Horn, A.K., Hughes, B.N.: Facebook and online privacy: attitudes, behaviors, and unintended consequences. J. Comput. Mediated Commun. **15**(1), 83–108 (2009)
10. Facebook Privacy Issues | Avoid Biggest Risk - Consumer Reports. http://www.consumerreports.org/cro/news/2014/03/how-to-beat-facebook-s-biggest-privacy-risk/index.htm
11. Poh, M.: Facebook & Your Privacy: Why It Matters. HKDC. http://www.hongkiat.com/blog/facebook-privacy-matters/. Accessed 17 Oct 2017
12. Facebook Help Center. https://www.facebook.com/help/
13. Albarran, A.B.: The Social Media Industries, pp. 147–148. Routledge, New York (2013)

fridaya@uef.fi

14. Liu, Y., Gummadi, K.P., Krishnamurthy, B., Mislove, A.: Analyzing Facebook privacy settings: user expectations vs. reality. In: Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, pp. 61–70, New York, NY, USA (2011)
15. Sieber, T.: 3 Things You Need To Know About Photo Tagging In Facebook. MUO, 13 June 2012. http://www.makeuseof.com/tag/3-things-you-need-to-know-about-photo-tagging-in-facebook/. Accessed 1 Dec 2016
16. Meharte, S.J., Chopde, N.R.: Review paper on preserving privacy policy of user uploaded images on data sharing sites. Int. Res. J. Eng. Technol. (IRJET), **3**(3) 2016
17. Ghazinour, K., Matwin, S., Sokolova, M.: Monitoring and recommending privacy settings in social networks. In EDBT/ICDT Workshops, New York (2013)
18. Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative filtering recommender systems. In: The Adaptive Web, pp. 291–324. Springer, Berlin (2007)
19. Ginjala, A., Li, Q., Wang, H., Li, J.: Semantics-enhanced privacy recommendation for social networking sites. In: 2013 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (2011)
20. Shripad, K.V., Vaidya, A.S.: Privacy preserving profile matching system for trust - aware personalized user recommendations in social networks. Int. J. Comput. Appl. **122**(11) (2015)
21. Krane, D., Light, L., Gravitch, D.: Privacy On and Off the Internet: What Consumers Want. Harris Interactive (2002)
22. Woodruff, A., Pihur, V., Consolvo, S., Schmidt, L., Brandimarte, L., Acquisti, A.: Would a privacy fundamentalist sell their DNA for $1000… if nothing bad happened as a result? The Westin categories, behavioral intentions, and consequences. In: Symposium on Usable Privacy and Security (SOUPS), vol. 4, p. 2 (2014)
23. Hatmaker, T.: Facebook will tag every photo ever taken of you—whether you like it or not. The Daily Dot, 7 February 2015. https://www.dailydot.com/debug/facebook-deepface-science/

fridaya@uef.fi

The Analysis of the Socio-Technical Environment (STE) of Online Sextortion Using Case Studies and Published Reports from a Cybersecurity Perspective

Alex Cadzow(✉)

Cadzow Communications Consulting,
73–75 High Street Stevenage, Hertfordshire SG1 3HR, UK
alex@cadzow.com

Abstract. This paper examines the Socio-Technical Environment (STE) of online Sextortion in the Digital Society. Sextortion refers to: "the broad category of sexual exploitation in which abuse of power is the means of coercion, as well as to the category of sexual exploitation in which threatened release of sexual images or information is the means of coercion". STE can be defined as he complex interrelationships between people, information, processes and technology and online networks. It addresses issues that are central to the domains of online sextortion and the interactivities between tasks, information and social. The methods chosen for this research project to address the sextortion problems and online presence have been identified in the research and developed here to address the STE of online sextortion from a cybersecurity perspective using Checklands Soft System Methodology (SSM) and a modified version of the Event Analysis of Systemic Team-work (EAST) model to illustrate it. The results of this project discuss the impact of sextortion on our socio-technical environment, and then how and where, law enforcement, governments, NGOs and technology companies are tackling these online threats. This paper further investigates they could take to prevent online sextortion from happening, protecting potential victims and better risk managing this environment. The research suggests that any cybersecurity measures taken, need to be considered in a holistic light. To prevent online sextortion because through the STE shows that the platforms, services and wide-ranging geographic nature of the groups and individuals of the victims and perpetrators because no single measure could mitigate all the vulnerabilities that lead to the proliferation of sextortion.

Keywords: Cybersecurity · Human-factors · Socio-Technical Systems

1 Introduction

Extortion is an old crime that has taken on a new dimension with developments in technology that led to four men in 2016 taking their own lives [13]. With the advent of modern communication technology, there is now the potential to affect anyone who is

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 817–833, 2019.
https://doi.org/10.1007/978-3-030-22868-2_56

fridaya@uef.fi

targeted and becomes trapped in an online sextortion racket irrespective of the geographic and social distance between extortioner[1] and extortionee.[2] Online sextortion is on the rise in the UK and many other major countries. The Federal Bureau of Investigation (FBI) [6] describe sextortion as *"a serious crime that occurs when someone threatens to distribute a person's private and sensitive material if they don't provide the perpetrator with images of a sexual nature, sexual favours, or money"*. The perpetrator may also threaten to harm a targeted individual's friends or relatives by using the information they have obtained from that person's electronic devices or online profile unless the target complies with their demands. The FBI state that online predators work to gain the trust of the targeted victim by pretending to be someone they are not. Predators look for their prey by lurking in chat rooms and record users who post or live-stream sexually explicit images and videos of themselves, or they may hack into a target victims' electronic devices using malware to gain access to their files and control their web camera and microphone without them knowing it. The advice the FBI gives includes:

1. Never send compromising images of yourself to anyone, no matter who they are—or who they say they are.
2. Do not open attachments from people you do not know.
3. Turn off your electronic devices and web cameras when you are not using them.

These pieces of advice are considered good security practice, but the problem is that not enough people follow these simple pieces of advice to lead to good security behaviour. Furthermore, for communication by friends, work or family a smartphone or other personal device cannot simply be turned off when not in use, so the advice does not account for the always connected always available nature of 21st-century living. There are other reasons for failure to act safely online, and these include naivety due to age (young or old), lack of knowledge about online security threats or failure to consider cybersecurity to be that important or to not consider security when interacting with services online. The thesis shows how education and awareness schemes can be used to improve peoples' attitudes and behaviour to cybersecurity through the social-technical environment.

This paper asserts that online sextortion is a social-technical issue from a cybersecurity perspective by showing the STE of online sextortion. STE can be defined as the complex interrelationships between people, information, processes and technology and online networks. It expands on the definition for Socio-technical Systems (STS). It is about people using their devices to access different social media and/or messaging services. A malicious user uses these same devices and services to socially engineer and/or attack a target person(s), in order, to obtain information or data which the malicious user can then use against the target, in order, to extort from them. Classically much of the initial social engineering is a side-channel attack that is required to achieve

[1] **Extortioner** *is a criminal who extorts money from someone by threatening to expose embarrassing information about them. Synonyms: blackmailer, extortionist. Type of criminal, crook, felon, malefactor, outlaw. Someone who has committed a crime or has been legally convicted of a crime.*

[2] Extortionee is somebody who has been subject to insidious extortion.

fridaya@uef.fi

the end goal. In themselves, in isolation, the social engineering may be impossible to distinguish from conventional social contact. The social aspect is that when people are affected, it affects their emotional state, dignity, may cause a financial harm and in some cases, may destroy their reputation. The social response revolves around education and awareness. The technology aspect is both the means that by which it happens and ways in which it may offer counters to online sextortion.

The scale of the problem of online sextortion has been discussed in the Guardian newspaper [12, 30] in its series "The Facebook files". In the report which analysed the content of many leaked documents, they show how Facebook moderators struggle with the mammoth task of policing content ranging from nudity to sexual abuse. Facebook had to assess nearly 54,000 potential cases of revenge pornography and "sextortion" on the site in a single month. Figures shared with Guardian staff reveal that in January 2017 Facebook had to disable more than 14,000 accounts related to these types of sexual abuse – and 33 of the cases reviewed involved children. (In this thesis, a child is below the age of legal consent for sexual activity i.e. in the UK this means 16 years of age) Because Facebook relies on users to report most abusive content, it means the real scale of the problem could be much higher.

There have been trends identified in The National Center for Missing and Exploited Children's CyberTipline sextortion reports [20], and their findings are broadly supported by the Child Exploitation and Online Protection Centre (CEOP) threat assessment of child sexual Exploitation and abuse report [20]. Online sextortion has most commonly occurred via phone/tablet messaging applications, social networking sites and video chats. Often the perpetrator uses multiple mediums and profiles. It ignores international borders leading to transnational child abuse. Also, online sextortion can lead to the perpetrator aiming to meet their victim as contact child sexual abuse. This is summarised in Fig. 1.

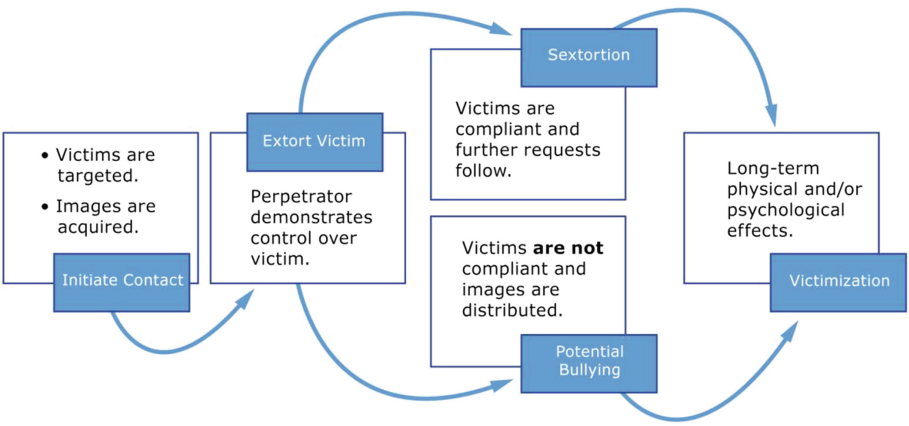


Fig. 1. Process of online sextortion from start to end

fridaya@uef.fi

2 The Problem Space

Previous studies have focused on young adults (18 to 25) who have been victims themselves [20]. This study aims for a broader examination of the age range to include minors (under 18) and adults in general (over 25) because online sextortion doesn't limit itself to affecting one age group.

Due to the methods that malicious actors use to carry out sextortion. Using social engineering and malware, they target users through social media, messaging and streaming platforms. There can be a waiting period between the images or videos that are obtained by the online perpetrator beginning the campaign of online sextortion. It is this ease of capturing and sharing data that has led to the increase in online child sexploitation and the proliferation of indecent images of children (IIOC). The targeted victims of online sextortion can experience a wide range of adverse effects. These include anxiety and depression with the worst outcome being death through suicide. These can manifest themselves at any time through the ordeal or even many years after. These adverse effects are often the reason the victims seek help or go to the authorities.

Shinayakana Systems Approach [23]
The research used the principles Shinayakana Systems Approach by making use of the Knowledge Pentagram System or the i-System for knowledge integration and creation. This is an approach which uses the pentagram system for problem-solving. The five ontological elements of the pentagram system are intelligence, involvement, imagination, intervention and integration. These will serve to provide the background information to the thesis and to develop the STE.

Intervention
In aiming to prevent online sextortion, education and awareness play an important role from interviews with novice Internet users about their knowledge and experience with security. Furnell et al. [8] report that users recognise responsibility for their protection online but they ignore the potential impact of failing to take adequate steps. This where the role of education and awareness campaigns aim to provide people the knowledge to stay safe online. By helping people, understand the threats and how attacks happen thus enabling awareness. On the other side, there are processes to teach people how to spot or identify when something doesn't appear right online, for example, a fake online social media profile. Education includes how to use technology safely. The goal is ensuring that people know and recognise that actions online have consequences. There is now a needed a long-term view of cybersecurity. The end result of education and awareness in cybersecurity is good security behaviour.

Ensuring that users recognise the threats and respond if an incident happens has been going on for a long time. A study by Leach [18] looking at improving user security behaviour showed opportunities to make significant security gains by creating a strong security culture. This advice has not changed since it is still about encouraging cybersecurity best practise. The importance has been stated by Atkinson et al. (2009) due to the increasing number of internet-enabled devices that people use. With the age of when people get their first device decreasing from late teens to sometimes before starting primary school. Therefore, developing their security awareness and

fridaya@uef.fi

behaviour is vital to them staying safe online. But an information security and people's behaviour study from Pham et al. [24] shows that there are still gaps between end-user's compliance in following security best practises. And businesses expectations of how they want cybersecurity implemented still has not been satisfactorily solved, but the trend of good security behaviour is becoming ingrained though slowly. There are now, plans to move towards cyberherd immunity.

The role of technology in cybersecurity has been to protect against malicious code and unauthorised access. It is about protecting software and services from vulnerabilities in their design. This relevance to online sextortion relates to threats from webcam hacking and malicious actors gaining access to victims' online accounts through social engineering to inject malware onto devices. As well a holistic approach to educating about cybersecurity by connecting it to the evolving technology threats has the potential to lead to more conscientious cyber citizens who are prepared and are aware of dealing with online threats [22]. This approach is vital to tackling online sextortion because it doesn't fit neatly into a single threat category because it can involve multiple types of attacks from social engineering, malware and hacking that a single approach to protect users is inadequate. Therefore, a holistic approach is needed.

Intelligence

As online 'sextortion' against children has grown, Federal Officials have urged back-to-school awareness. Donna [3] has investigated as part of public service announcement from the U.S. Department of Justice of how Federal officials and advocates have observed the increasing prevalence of online sextortion and have attempted to study and understand the scale of the problem. The report to Congress on The National Strategy for Child Exploitation Prevention and Interdiction [15] noted a 2015 analysis of 43 sextortion cases with child victims found 28% of cases examined there were suicides or attempted suicides due to the psychological toll on the children. The report observes that online sextortion is an evolving threat due to its online context this is due to the evolving means of sexploitation is that mobile smart devices have fundamentally changed the way offenders can abuse children. Apps on these devices can be used to target, recruit or groom and coerce children to engage in sexual activity. This again due to the ability for a single perpetrator to efficiently target many victims than before the internet didn't exist. But online sextortion is only part of a broader problem of online sexploitation with many reported instances an individual has attempted to groom or lure a child for sexual purposes during online communication. Requiring a need to understand how online actions an impact in the real world since the adverse effects have do not just stop when devices are switched off, they continue.

In cybersecurity, the role of behaviour and motivation must be considered. The researcher Aikans has studied these and how users' actions online can be greatly magnified when compared to offline actions which have relevance to online sextortion. A study of user modelling perspective [33] that social computing applications (SCA) which include services like Facebook and Twitter by allowing users to discuss, share and connect with other people. These have been designed to motivate users to participate in an online community. It is this design philosophy which unintentionally produces the security and safety flaws that allow users to be the socially engineered leading to online sextortion. This behaviour is recognised as the Online Disinhibition

fridaya@uef.fi

Effect [29] where while online, some people self-disclose or act out more frequently or intensely than they would in person. Six factors create this effect: dissociative anonymity, invisibility, asynchronicity, solipsistic introjection, dissociative imagination and minimisation of authority. In online sextortion, it is these behavioural traits that lead people to carry out malicious acts online because they believe they cannot be caught, and their actions give power over their victims.

In the STE of online sextortion, there needs to be an understanding of how people interact with technology and how perpetrators take advantage of these online social interactions to carry out malicious acts. Explaining the behaviour of perpetrators, by the way, they think they operate behind by a veil of anonymity, and false personas are what law enforcement can pierce.

Imagination

There needs to be a way to integrate the information and knowledge from the education, awareness and technology case studies. By the STS approach which incorporates two different types of structures and processes: technical systems which are engineered to identify interactions between users and systems [7]. STS does not provide fixed solutions but provides a framework which the designers and users can contribute to the conceptual framework of meta-design. A research project from Carmien et al. [2] on Socio-Technical Environments (STE) Supporting People with Cognitive Disabilities Using Public Transportation looked at this problem of integrating human-centred usability, context-aware paired with ubiquitous and pervasive computing, with the public transport system. The relation to the proposed STE for online sextortion is due to human interaction with computers and smart devices which often have a design focus on being readily usable. Through these devices, the platforms that sextortion occurs through happen on them and finally the internet provides the means of communication which felicitates the acts of sextortion.

To give a cohesive structure to the STE for online sextortion the Event Analysis of Systemic Team-work (EAST) as an integration of methods for analysing complex sociotechnical systems (Stanton 2016). It aims to represent behaviours within sociotechnical systems with three network elements. The first, task networks describe the relationship between tasks, their sequence and interdependences. Second, social networks analyse the organisation of the system and the communications taking place between the actors and agents within the system. Finally, information networks describe the information that the different actors and agents use and communicate with (see Fig. 2). The networks can be related to the STE of online sextortion with tasks being the acts and processes of online sextortion; the social networks are the communication and social engineering process that occurs between the perpetrator and the victim. Within the information networks are messages, images and video streams exchanged between the perpetrator and victim.

A report from the UN [32] on releasing children's potential and minimising risks recognising's the risks and benefits of social and technology interactions when they are only a click away through the internet. The report stresses there is a need to understand the online risks and harm in order them to be recognised so technology, education and awareness can be used to protect children. The technology elements and social elements cannot be considered in isolation instead need to be tackled together. These

fridaya@uef.fi

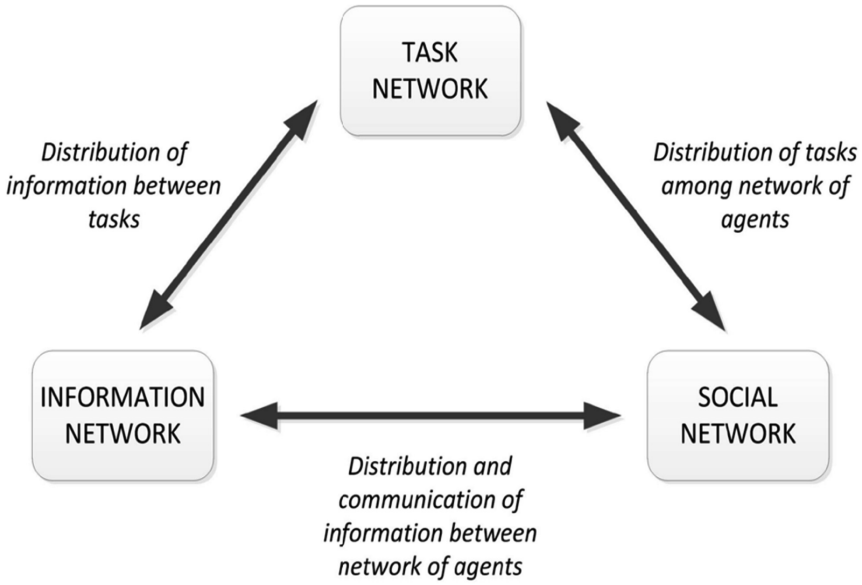


Fig. 2. Networks of Networks approach used by EAST

themes are repeated in a webinar from the Office of Juvenile Justice and Delinquency Prevention [24] and in a report from the Wittes et al. [34]. This interconnectivity of society through digital technologies [9] allows for potential victims to be always accessible to perpetrators due to the nature of how social networking sites and always online communication applications. Reports of the negative effects of victims include depression, anxiety, substance abuse, and suicidal ideation. Finally, as with the UN report only by recognising the interaction between technology and social elements can just begin to prevent or stop. First, by following cybersecurity best practise and accepting awareness of information sharing online should be carried out sensibly to prevent that information from being used against an individual.

Involvement

To better understand the STE of online sextortion makes use of statistical incidents of it occurring but due to the nature of the subject, this research project was not able to carry out its' own original research. Therefore, studies done previously will be used to give context to the STE of online sextortion.

While there are variations in the rate of incidents between countries about online sextortion the statistics tend to follow the same pattern. An analysis of online sextortion that affected Czech children is that affects approximately six to eight percent of them. Nearly, all cases of online sextortion had very similar techniques that were used to blackmail children. These focused on gaining confidence, luring out intimate material and subsequent blackmailing. Also, to obtain the child's victim attention, the perpetrator positively responds to all materials the victim's shares or posts. During the period of the campaign of online sextortion, one hundred percent of the child victims during

fridaya@uef.fi

the time they were blackmailed did not contact an adult or officially organization [17]. These kinds of statistics and observations are not just relevant to other western countries. In India, even though it doesn't have the same internet penetration rate as western countries it has experienced the same malicious problems. In India, around 134 million children have access to the internet mainly through mobile devices. While there no statistics about online sextortion (they know that cyberbullying in Indian schools is a serious problem) affecting children, it has been recognised as a problem which they have begun to tackle. In India, online risks have exposed children to these online crimes and abuses are due to the lack of digital literacy and online safety measures [14]. Girls were more likely to be affected than boys when they are younger than fourteen years old with older children if they are affected the gender split narrows. This type of information can inform areas where education and awareness campaigns need to focus on and could show what kind of technologies might be useful in preventing or mitigating online sextortion.

In an online survey of 1631 persons aged 18 to 25 who had been affected by online sextortion, the respondents were primarily female (83%), and the majority were under 20. It should be noted that survey findings might not reflect the number of real-life incidents. Overall the sextortion incidents being reported were diverse, but incidents broadly fell into two groups: (a) In the wake of face-to-face romantic or sexual relationships during which sexual images were taken or shared, an aggrieved partner threatened to disseminate images to force submission or humiliate the respondent. (b) A perpetrator who met a respondent online used a sexual image obtained from the respondent or some other source to demand more images or sexual interactions. But only 1 in 5 sort helps from the website or apps they had been targeted through. Just 16% of respondents reported episodes to police, but police involvement was considerably more common among those who disclosed sextortion incidents to family or friends, were victims of violence or threats of violence in addition to the sextortion, or who saw a doctor or mental health professional because of the incident. An online survey of Australian adults that examined technology-facilitated sexual violence victimisation recorded results from 2,956 adults (aged 18 to 54 years). The survey results showed that 18 to 24 women were more likely to be affected but over the total age range, there was no significant difference [26]. Also, women were more likely to report sexual incidents of revenge porn and men were more likely to report sexual blackmail.

Integration

The Event Analysis of Systemic Teamwork (EAST) model refined by Stanton et al. (2005) to analysis and observe the distributed situation awareness in dynamic systems. Has been applied to the analysis of socio-technical systems of submarines to better understand and improve the cognitive workflow [28]. This method has also been applied to cybersecurity in a study of the distributed nature of cyber situation awareness [31]. They showed that the cognitive work of cybersecurity happens across individuals, technological agents and functional domains. They noted that the degree to which the cybersecurity system can respond to a threat is a function of the awareness of individuals; the intrusion detection analyst, the system administrator, the forensic analysts, the policymaker and the threat landscape analysts. The EAST model has also been used to understand and the impact of team collaboration on cybersecurity situational

fridaya@uef.fi

awareness to improve it in the cybersecurity domain [27]. This applies to the STE of online sextortion because it involves the interaction of many different platforms, the users and participants in the system and the threats agents that form part of online sextortion. By understanding how these induvial elements interact and relate to each other ways to counter and prevent online sextortion can be better understood and proposed. The role of cyber security will be integrated into the EAST model is presented in Fig. 3 and is developed in the method section.

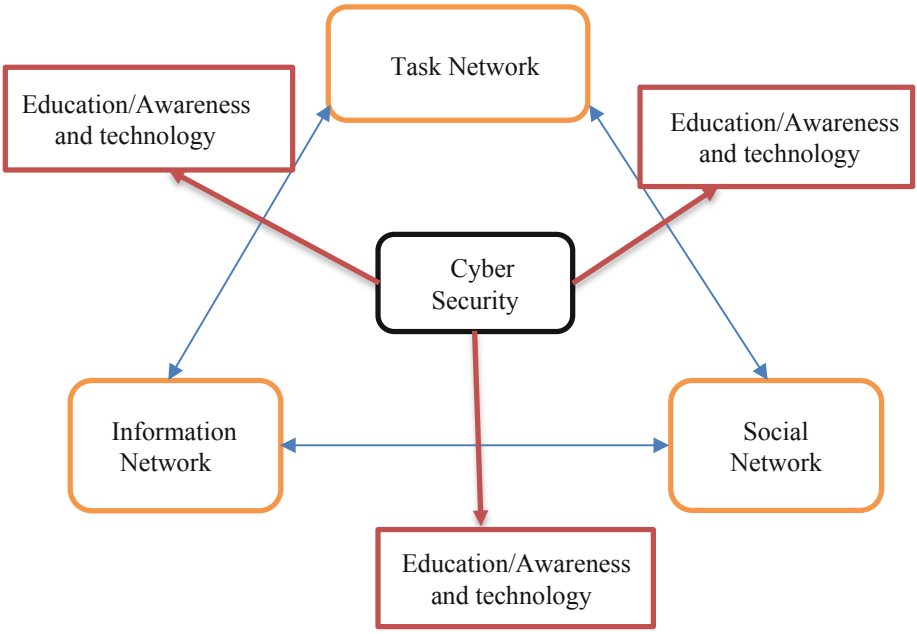


Fig. 3. STE concept integrated with cybersecurity solutions

Checklands Soft-Systems-Methodology (SSM)

Checklands SSM is a way to apply systems engineering approaches to solve a problem. In this paper is that online sextortion can be represented by the STE from a cybersecurity perspective, in order, to gain insight into how it can be prevented. Here only three parts of Checklands SSM where be used to model and develop the STE. These are rich text, root definitions (structured description of a system) which will use the three elements of what, why and how using the BATWOVE (changing benefactor to the perpetrator) to formulate the root definitions. The decision to change benefactor to the perpetrator is that identifying an individual that undertakes malicious behaviour as benefactor would be in abysmal taste. Also, it reflects the actions and process that occur during online sextortion by describing who initiates the actions within the system (perpetrator, actors, transformation, worldview, owner, victim, environmental constraints).

fridaya@uef.fi

A EUROPOL report from the law enforcement perspective of online sexual coercion and extortion form of crimes affecting children. The report aimed to raise awareness and to contribute to the public debate on what are the effective responses to it. The report determined two broad perpetrator profiles.

Offender Profile Sexual Motivation:

- Male.
- Operates alone but trades the acquired content.
- May act on both international or national level.
- Activity-driven by knowledge of languages.
- Targets female victims.
- May know the victim in person.
- Goal to obtain sexual material.

Offender Profile Financial Motivation:

- Both genders.
- Members of an organised criminal enterprise.
- Operates in teams based in developing countries.
- May act at both international and national level.
- Targets male victims in countries linked by language.
- Does not know the victim in person.
- Main goal: to obtain money.

The report noted there is no 'typical' victim of sexual coercion and extortion. There are no determined patterns of causation as to how and why people become victims. They do note from studies that females are more likely to be affected than males and when minors are affected there is a correlation between higher risk of sexual coercion as result of unsupervised internet access and failure to perceive risks to online actions.

Victim Profile:

- Any person whose sexual material could be acquired by a perpetrator.
- Usually female in case of sexually motivated perpetrators.
- Usually male in case of financially motivated perpetrators.

The characteristics in cases of online sexual coercion and extortion affecting children:

- The naivety of the victims, either on a relational level or on a technical level.
- The absence of parental control.
- Willingness to share self-generated sexual content.
- A significant amount of time spent online each day.
- Use of social networks and other ways of online communication, especially through mobile devices.
- Befriending strangers.
- Sexualised conversions with strangers.
- Lack of technical knowledge.

fridaya@uef.fi

The EUROPOL report [5] summarises what response needs to be implemented to counter or prevent online sextortion:

- Guidelines for IT industry.
- Prevention software.
- Safer internet policies.
- Studies on offenders in the virtual world/grooming.
- Virtual – versus real-world studies.
- Evaluation of national strategies.
- Law enforcement-oriented actions:
- Websites for the online reporting of online sextortion.
- Police advise/warning in newspaper.
- Education and awareness raising.

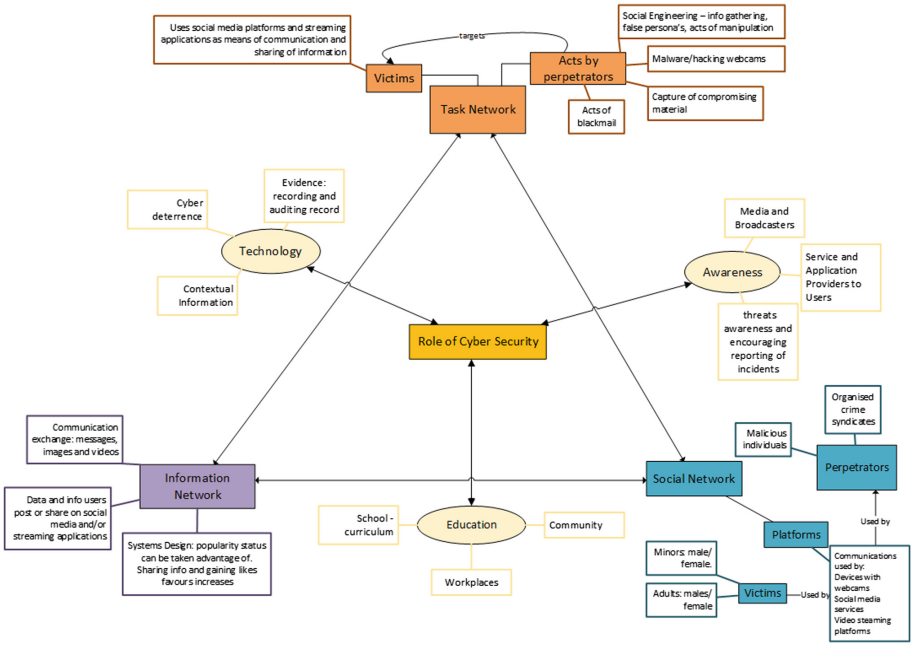


Fig. 4. East model illustrating the STE

Hybrid EAST model for Online Sextortion

The EAST model was originally developed to analysis work patterns within defence units for example submarines and warships. It uses social, task and information networks to classify the connections between people and the actions they undertake. Instead of analysing work patterns the model has been adapted to develop the STE (Fig. 4) by using the networks to show the social links between victims and perpetrators. Also, the tasks that occur during the process of online sextortion. Finally, what information is exchanged between the victims and perpetrators during online sextortion.

fridaya@uef.fi

The development of the socio-technical environment using the EAST model pulls in data and ideas that have been generated through the SSM and the literature review. The EAST model comprises three networks models' task, social and information. Task networks describe the relationships between tasks, their sequence and interdependences. Social networks analyse the organisation of the system (i.e. communications structure) and the communications taking place between the actors and agents working in the team. Finally, information networks describe the information that the different actors and agents use and communicate during task performance (i.e. distributed situation awareness).

3 Discussion

Discussion of the STE of online sextortion will be from a cybersecurity perspective from four domains: law enforcement, political, institutional and personal (or intercommunal).

A report from Bracy [1] shows the conversation about online sextortion is becoming more open by discussing why online sextortion forms part of the issues surrounding privacy and cybersecurity. The application of the philosophy of Louis Brandeis, "sunlight is the greatest disinfectant", in which exposure of deviant behaviour reduces the likelihood or incidence of such behaviour would be expected to be a positive way forward. Increased general awareness about sextortion and the harm it causes helps victims feel more confident about coming forward to prosecute their tormentors. Helping get police the budgets, tools, and skills they need to go after perpetrators; starts parents and teachers down the path of educating children about digital hygiene and privacy; and gets businesses thinking about ways they can build in protections and education into their products and services. This connection to STE of online sextortion from a cybersecurity perspective is that many of the resources and legislation that gives people protection and the right to justice applies to this area as well and it will be argued that progress is being made in tackling these problems and that the situation is improving.

From the law enforcement area, there are local, national and international (Europol and Interpol) organisations that must deal with reports and crimes that come under the banner of online sextortion. In the report 'Police Effectiveness in a Changing World' [16], it is explained that police balance the change from a traditional approach to law enforcement through visual deterrence (police in the neighbourhood being visible symbols of arrest and conviction of crime). This can deal with visible crime, towards an approach of intelligence-led and problem-oriented policing to deal with the unseen crimes represented by cybercrime and similar less visible crimes. Law enforcement has had success in prosecuting perpetrators who carry out acts of online sextortion. A pedophile named as Paul Leighton when investigated by the police carrying out a search warrant examined his phone, they found evidence that he had targeted hundreds of children. In Seaham [32], County Durham, one of his victims informed the police which led to him being arrested and sentenced to prison [34]. But not all cases are this straightforward, and many have traumatic endings [4]. In June 2015, a Northern Irish teen committed suicide after being blackmailed. In this instance, his tormentor shared

fridaya@uef.fi

intimate photos of the teen when he couldn't pay the ransom, i.e. the ransomer or tormentor acted on his threat. Jailing the tormentor took two years. Apprehending the perpetrator only followed from an extensive joint investigation by the Police Service of Northern Ireland who [25] identified the perpetrator as a resident of Romania. They contacted the Romanian authorities to include them in the investigation. In the end, it spanned several countries, including Australia and Belgium, and enlisted the help of Europol and Britain's National Crime Agency, in order gather evidence and liaise to build a case across spanning Europe, before the tormentor was arrested and sentenced in Romania.

For law enforcement role to tackle online sextortion deterrence is one of the primaries means to deter perpetrators from committing malicious acts. In policing, deterrence of crime relies on two key principles for it to be effective [11]. Theses are:

1. Police deter crime by increasing the perception that criminals will be caught and punished.
2. Incarceration can act as a deterrent when individual fears incarceration before they commit a crime and thus refrain from committing future crimes. Though for incarceration to act as a deterrent the severity of punishment must be sufficient.

For cyber deterrence to work there three key components: a credible defence; the ability to retaliate; and, the will to retaliate [19]. It is possible for these components to be adapted to law enforcement since from a defence perspective cyber deterrence seeks to dissuade the attacker from acting for fear of retaliation. For law enforcement, cyber deterrence needs to be able to dissuade the perpetrator from committing acts of online sextortion for fear of being caught and incarcerated in prison. The problem is at this moment there is no single doctrine from a law enforcement perspective which has yet to be developed. Concerning the STE of online sextortion, there are only partial solutions from government, NGOs and technology companies which only when brought together as part of a holistic solution might fill the role of a cyber deterrence.

The political sphere is what governments and international organisations that include the EU and UN are doing to mitigate or prevent online sextortion. The role of these government and political bodies in tackling online sextortion come through the policies they propose and their actions in implementing those policies. There is a broad consensus across of the political spectrum to tackle online sextortion though there is a difference of opinion whether it should be through enforced measures or voluntary actions. The UK government has an open approach to cybersecurity policy [10]. For example, they instigated measures to keep children safer online at school and home by ensuring that schools filter inappropriate online content and teach pupils about staying safe online and ensure this is achievable by also making sure internet providers provide the option of filters (Gov.uk 2015). The UK policy is also moving towards adopting a more active posture in defending the UK from cyber threat through the closer partnership between government, industry and law enforcement (ncsc.gov.uk 2016). This involves taking an active approach to blocking IP addresses as the host/node address of a source of malicious content to block phishing attacks, DDOS and malware sites. But these are technical solutions which have limited application to preventing online sextortion which is largely a social engineering attack. The area where the blocking of IP addresses could be useful is in tackling organised crime syndicates. This would aim

fridaya@uef.fi

to block the means of communication of organised criminal groups who carry out online sextortion from a known area but in doing so there is a chance legitimate user would also be blocked. The Digital Economy Act 2017 aimed to better protect minors online by implementing better age-verification on pornographic sites [21]. The original intent was to apply such age verification technology to a wider type of site, including social media sites, but due to concerns of correctly determining a user's age without leading to an increase in the sharing of sensitive data led it being limited to pornographic sites. This highlights that proposals to tackle online malicious acts must carefully balance freedom the internet provides with the privacy of peoples' data.

Though it will be a while before cases of online sextortion decrease because as more people are aware of online sextortion and know who to contact about it. It will lead to increase in the number of the reported cases of online sextortion before the number of cases starts to decrease.

Cyber herd immunity initiatives are a possible way to lead to better prevention of online sextortion. There are key ideas that form the concept of cyber herd immunity. There foundations come from the gov.uk (2014) cybersecurity behavioural insights report:

1. Security by default.
2. Maintain confidentiality of data and information by using encryption.
3. Use of access protection.
4. Application whitelisting.
5. Malicious code detection/prevention.
6. Always install software updates.

4 Conclusion

The analysis of the socio-technical environment of online sextortion from a cybersecurity perspective has shown that it cannot be treated separately as a purely social problem (social engineering) or the technical problem (vulnerabilities of devices and software) that can cause online sextortion to happen. Instead due to the wide-ranging and overlapping nature of the issues surrounding online sextortion, a holistic view must be taken. It showed why online sextortion is a socio-technical issue from a cybersecurity perspective.

From a law enforcement domain, they have had success against organised crime groups and individuals who carry out online sextortion by being to trace and identify who the perpetrators are. To do more law enforcement could develop a cyber deterrence strategy to deter individuals and groups from carrying out these kinds of online malicious acts at all. The Governments and other international bodies use education policy and collective action to tackle online sextortion by aligning responses to it. Also, Government departments and agencies act to deal with fraudulent messengers to prevent phishing by blocking IP addresses used for malicious actions. They could expand this to block IP addresses being used by organised crime groups who are carrying out online sextortion. This would disrupt crime groups means to socially engineer people thus depriving them of income. NGOs and charities often take the lead in education and

fridaya@uef.fi

awareness campaigns to improve children's and adults' internet safety and ensure that they can recognise online malicious attacks, for example, phishing attacks and malware downloads. By working more closely with Government bodies and technology companies, they could further improve their education and awareness campaigns and be able to reach more people than they could alone.

Technology companies already have in places measures to block and remove users who break the terms of the agreement of the services they provide which often involves the false profiles that can be used to carry out online sextortion and they also have policies in place to co-operate with law enforcement warrants and investigations. Now, this is a passive and reactive way to deal with online sextortion. Technology companies could take a more active role by using software and AI to determine whether a user is using false profiles which could be used to carry out malicious acts.

Families and individuals are now being encouraged to take a more active role in their online security. Parents could be encouraged to use applications like Oyoty an app that helps their children learn and improve their security behaviour when online. While it doesn't prevent online sextortion directly but by enhancing users' security behaviour, they should have the knowledge and awareness to avoid being caught out in a social engineering scam which online sextortion is a specific type of social engineering.

List of recommended future actions are:

1 Greater use of holistic planning.
2 Law enforcement should develop a cyber-deterrence strategy.
3 Improved recording and auditing of digital evidence.
4 Expanded use of IP blocking of known malicious addresses.
5 Greater co-operation between Government, NGOs and technology organisations to improve cyber safety and security education and awareness.
6 Technology companies should develop software and AI to remove false profiles before they can be used for malicious acts.
7 Encourage better personal security behaviour through feedback applications like Oyoty aim to achieve.
8 Expanding cyber insurance to cover and protect individuals.

References

1. Bracy, J., iapp.org: Why 'sextortion' is part of a larger privacy and cybersecurity issue. Portsmouth (2016). https://iapp.org/news/a/why-sextortion-is-part-of-a-larger-privacy-and-cybersecurity-issue/. Accessed 01 Nov 2017
2. Carmien, S., Dawe, M., Fischer, G., Gorman, A., Kintsch, A., Sullivan, J.F.: Socio-technical environments supporting people with cognitive disabilities using public transportation. ACM Trans. Hum. Comput. Interact. **12**(2), 232–262 (2005)
3. Donna, Washingtonpost.com: As online 'sextortion' against children grows, feds urge back-to-school awareness. Washington, DC (2016). https://www.washingtonpost.com/local/education/online-sextortion-against-children-growing-feds-urge-back-to-school-awareness/2016/09/19/395a6cbe-7b5b-11e6-beac-57a4a412e93a_story.html?utm_term=.2c7a32c506af. Accessed 28 Aug 2017

fridaya@uef.fi

4. Edwards, R.: How police caught online blackmailer whose plot led to suicide of Irish teen. Dublin (2017). https://www.irishtimes.com/news/crime-and-law/how-police-caught-online-blackmailer-whose-plot-led-to-suicide-of-irish-teen-1.3207653. Accessed 02 Nov 2017
5. European Cybercrime Centre: Online sexual coercion and extortion as a form of crime affecting children – Law Enforcement Perspective. Europol, The Hague (2017)
6. FBI.gov: What is sextortion? (2017). https://www.fbi.gov/video-repository/newss-what-is-sextortion/view. Accessed 11 July 17
7. Fischer, G., Herrmann, T.: Meta-design: transforming and enriching the design and use of socio-technical systems. In: Wulf, V., Schmidt, K., Randall, D. (eds.) Designing Socially Embedded Technologies in the Real-World. Computer Supported Cooperative Work. Springer, London (2015)
8. Furnell, S., Tsaganidi, V., Phippen, A.: Security beliefs and barriers for novice Internet users. Comput. Secur. **27**, 235–240 (2008)
9. García-Peñalvo, F.J.: The WYRED project: a technological platform for a generative research and dialogue about youth perspectives and interests in digital society. J. Inf. Technol. Res. **9**(4), vi–x (2016)
10. Gov.uk: Using behavioural insights to improve the public's use of cyber security best practices. Government Office for Science, London (2014). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/309652/14-835-cyber-security-behavioural-insights.pdf. Accessed 27 Sept 2017
11. Gov.uk: Cyber security. London (2017). https://www.gov.uk/government/policies/cyber-security. Accessed 06 Nov 2017
12. Guardian.co.uk: UK child protection officers receive one sexting-related case every day. The Guardian, London (2016). https://www.theguardian.com/uk-news/2015/jun/15/uk-child-protection-officers-one-sexting-related-case-every-day. Accessed 06 June 2017
13. Independent.co.uk, Massey, N.: Sextortion: rise in blackmail-related suicides over sexual images shared online. Independent Digital News and Media, London (2016). https://www.independent.co.uk/news/uk/home-news/sextortion-rise-suicides-blackmailing-sexual-images-sharing-social-media-a7446776.html. Accessed 06 June 2017
14. Jain, R.: Sexual predation: protecting children in the era of Internet – The Indian Perspective. Fiat Iustitia **11**(1), 141–152 (2017)
15. Justice.gov: The National Strategy for Child Exploitation Prevention and Interdiction – A Report to Congress, U.S. Department of Justice (2016). https://www.justice.gov/psc/file/842411/download. Accessed 20 Oct 2017
16. Karn, J.: Police and Crime Reduction – The evidence and its implications for practice. The Police Foundation, London (2013)
17. Kopecky, K.: Online blackmail of Czech children focused on so-called "sextortion" (analysis of culprit and victim behaviours). Telematics Inform. **34**(1), 11–19 (2017)
18. Leach, J.: Improving user security behaviour. Comput. Secur. **22**(8), 685–692 (2003)
19. Wei, L.H.: The Challenges of Cyber Deterrence. Pointer J. Singap. Armed Forces **31**(1), 12–22
20. missingkids.com: Sextortion – Trends identified in CyberTipline sextortion reports (2017). http://www.missingkids.com/Sextortion. Accessed 11 Aug 2017
21. Muffett, A.: Digital Economy Bill – Written evidence submitted by Alec Muffett. Parliament, Westminster (2016). https://publications.parliament.uk/pa/cm201617/cmpublic/digitaleconomy/memo/DEB39.htm. Accessed 06 Nov 2017
22. Nationalcrimeagency.gov.uk: Help available for webcam blackmail victims: don't panic and don't pay. London (2016). http://www.nationalcrimeagency.gov.uk/news/960-help-available-for-webcam-blackmail-victims-don-t-panic-and-don-t-pay. Accessed 23 Oct 2017

fridaya@uef.fi

23. Nakmori, Y., Wierzbicki, A.P: A methodology for knowledge synthesis. In: IEEE International Conference on Systems, Man and Cybernetics (2008); Pham, H.C., Pham, D.D., Brennan, L., Richardson, J.: Information security and people: a conundrum for compliance. Australas. J. Inf. Syst. **21**, 1–16 (2017)
24. Psni.police.uk: Prison Sentence for Blackmailer of Coalisland Tennager. Police Headquarters, Belfast (2017). https://www.psni.police.uk/news/Latest-News/290817-prison-sentence-for-blackmailer-of-coalisland-teenager/. Accessed 13 Nov 2017
25. Powell, A., Henry, N.: Technology-facilitated sexual violence victimization: results from an online survey of Australian adults. J. Interpersonal Violence, 1–29 (2016)
26. Rajivan, P., Cooke, N.: Impact of team collaboration on cybersecurity situational awareness. In: Liu, P., Jajodia, S., Wang, C. (eds.) Theory and Models for Cyber Situation Awareness, Lecture Notes in Computer Science, vol. 10030. Springer, Cham (2017)
27. Stanton, N.A., Bessell, K.: How a submarine returns to periscope depth: analysing complex socio-technical systems using Cognitive Work Analysis. Appl. Ergon. **45**(1), 110–125 (2014)
28. Suler, J.: The online disinhibition effect. CyberPsychology Behav. **7**(3), 321–326 (2004)
29. theguardian.com: Facebook flooded with 'sextortion' and revenge porn, files reveal (2017). https://www.theguardian.com/news/2017/may/22/facebook-flooded-with-sextortion-and-revenge-porn-files-reveal. Accessed 11 July 17
30. Tyworth, M., Giacobe, N.A., Mancuso, V., Dancy, C.: The distributed nature of cyber situation awareness. In: IEEE International Multi-Disciplinary Conference Methods in Situation Awareness and Decision Support, New Orleans, LA (2012)
31. UN: Releasing children's potential and minimizing risks – ICTs, the Internet and Violence against Children. Office of UN Special Representative of Secretary-General on Violence against Children, New York (2014)
32. Unwin, B.: Seaham internet pervert jailed after threatening girl with gang rape and selling her into prostitution. High Wycombe, Buckinghamshire (2017). http://www.thenorthernecho.co.uk/news/local/northdurham/peterlee/15398454.Internet_pervert_jailed_after_threatening_girl_with_gang_rape_and_selling_her_into_prostitution/. Accessed 02 Nov 2017
33. Vassileva, J.: Motivating participation in social computing applications: a user modelling perspective. User Model. User Adap. Inter. **22**(1–2), 177–201 (2012)
34. Wittes, B., Poplin, C., Jurecic, Q., Spera, C.: Sextortion: cybersecurity, teenagers, and remote sexual assault. Center for Technology Innovation at Brookings (2016)

fridaya@uef.fi

Lightweight Datapath Implementation of ANU Cipher for Resource-Constrained Environments

Vijay Dahiphale(✉), Gaurav Bansod, and Ankur Zambare

Electronics and Telecommunication,
Pune Institute of Computer Technology, Pune, India
vijaydahiphale96@gmail.com, ankur.zambare@gmail.com,
gvbansod@pict.edu

Abstract. With the advent of technologies like IoT, the need for lightweight designs and architectures are in focus. The lightweight cryptography is the frontrunner for resolving the issues related to the memory size, power consumption and lesser gate counts while implementing the designs in resource-constrained environments. This paper focuses on lightweight datapath implementation of cipher ANU for environments like IoT. The 8-bit datapath design for ANU cipher is proposed and implemented in this paper for small, tiny 8-bit microcontrollers. The previous datapath implementation of ANU cipher results in higher gate counts and thus consumes more footprint area. The design proposed in this paper is lightweight and suitable for the environments where designs metrics like Gate Counts and power dissipation plays an important role. The paper presents detailed analysis and implementation of 8-bit datapath design of ANU cipher and compares with other existing datapaths designs on various design metrics.

Keywords: Lightweight cryptography · IoT · Embedded security · Encryption · FPGA implementation · Datapath

1 Introduction

In recent years, the size of electronic devices is reducing and they are becoming smarter day by day. Due to the evolution of Internet of Things (IoT), a lot of devices like security cameras, smart TVs, refrigerators etc. are capable of accessing the internet and passing a large amount of data to the server. Most of these devices are placed in a hostile environment where the data can be accessed by the attacker. Security plays an important role in IoT. AES [1] (Advanced Encryption Standard) and DES [2] (Data Encryption Standard) allow us to implement a secure environment by encrypting data at the transmitter and decrypting data at the receiver. IoT devices are normally battery powered devices and use 8-bit microcontrollers. These microcontrollers have low computational power. Use of algorithms such as AES and DES prove to be more costly in such constrained environments, as they require a lot of memory, GE and power consumption for their implementation [3]. IoT devices are vulnerable to physical

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 834–846, 2019.
https://doi.org/10.1007/978-3-030-22868-2_57

fridaya@uef.fi

attacks, hence the field of lightweight cryptography has grabbed the interest of the cryptographic community.

In an Embedded System, Lightweight cipher can be implemented in software but the devices like RFID cards do not have a software stack or processors. Hence providing security by hardware implementation of lightweight ciphers is very important in such devices. Thus hardware based security is also very useful for different application in IoT. The efficiency of hardware implementation relies on various design metrics like area of the chip, power consumption, and execution speed.

ANU is 25 round cipher and supports 64-bit Plaintext and 80/128-bit key. It beats many standard lightweight ciphers in terms of performance [4, 5]. An ISO certified cipher PRESENT [6] needs around 1560 GE for its ASIC level hardware implementation while ANU cipher needs only 1015 GE making it much more lightweight than PRESENT. ANU needs only 22 mW of dynamic power while PRESENT cipher consumes 39mW of power [4, 5]. Hence, because of lower power consumption and lower memory requirement, ANU is well suited for battery operated as well as resource-constrained devices.

Achieving efficiency in every metrics of the hardware implementation is a challenging task and hence we have to make a trade-off between the different metrics [7]. By considering this, we have implemented 8-bit serialized architecture of ANU. Also, the same architecture performs differently on different platforms. Hence we have evaluated the performance of proposed architecture on four different FPGA platforms and compared based on various metrics such as power consumption, throughput, number of GE required, number of LUT's required, efficiency, clock cycles, latency, etc. Also, proposed architecture performance is compared with the implementation standard lightweight ciphers.

1.1 Notations and Acronyms

P_MSB_i	→	Most Significant 32 Bits of Plaintext for i^{th} Round
P_LSB_i	→	Least Significant 32 Bits of Plaintext for i^{th} Round
<<<	→	Left Circular Shift
>>>	→	Right Circular Shift
⊕	→	Exclusive OR
ASIC	→	Application-Specific Integrated Circuit
GE	→	Gate Equivalent
LUT	→	Look Up Table
EXOR	→	Exclusive OR
S-Box	→	Substitution Box
P-Layer	→	Permutation Layer
LSB	→	Least Significant Bit
MSB	→	Most Significant Bit
FSM	→	Finite State Machine

fridaya@uef.fi

2 ANU Block Diagram

ANU [4, 5] is 25 round, feistal structured lightweight block cipher. Figure 1 shows the structure of ANU for a single round, which consists of P_MSB_i and P_LSB_i, which are MSB 32-bits and LSB 32-bits for the i^{th} round respectively. To provide more non-linearity in the Ciphertext, ANU cipher uses 'F' function which has S-Box and Circular shift block. Bit Permutation (BP) is used in this cipher to increase randomness in the Ciphertext.

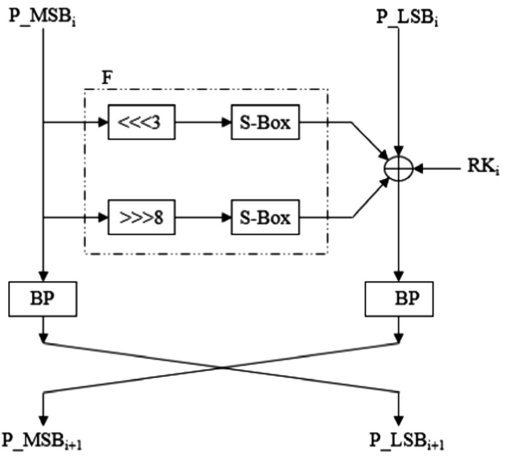


Fig. 1. Block diagram of ANU

2.1 Substitution Box

As S-Box is the only non-linear component in the cipher, it decides the strength of the cipher against different attacks [6]. Therefore, values in the S-Box are chosen in such a way that all S-Box properties are fulfilled. ANU uses 4-bit Substitution Box (S-Box) i.e. 4-bit input and 4-bit output. This 4-bit S-Box is used repetitively for substitution of 32-bits. S-Box for ANU is shown in Table 1.

Table 1. S-Box used for ANU

X	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
S(X)	2	9	7	E	1	C	A	0	4	3	8	D	F	6	5	B

2.2 Permutation Layer

Permutation Layer works on individual bits to produce more randomness in Ciphertext. It shuffles the bits coming from the input and produces the output. ANU uses 32-bit P-Layer. It is basically used to produce the good avalanche effect and to increase the complexity of the cipher. 32-bit P-Layer of ANU is shown in Table 2.

fridaya@uef.fi

Table 2. 32-bit P-layer used in ANU

i	00	01	02	03	04	05	06	07
BP(i)	20	16	28	24	17	21	25	29
i	08	09	10	11	12	13	14	15
BP(i)	22	18	30	26	19	23	27	31
i	16	17	18	19	20	21	22	23
BP(i)	11	15	03	07	14	10	06	02
i	24	25	26	27	28	29	30	31
BP(i)	09	13	01	05	12	08	04	00

2.3 Circular Shifts

ANU cipher uses two different circular shifts inside 'F' function.

- Left Circular Shift By 3 (<<<3)
- Right Circular Shift By 8 (>>>8)

2.4 EXOR and Swapping

The output of 'F' function, LSB 32-bits of 64-bit data and LSB 32-Bits of 128 bit round key are EXORed in each round of ANU Cipher. EXOR operation can be easily noticed from Fig. 1. At the end of each round, swapping of MSB 32-bits and LSB 32-bits is done to maintain the strict avalanche criterion.

2.5 Encryption Algorithm

Input: Plaintext, Key

$key = K_{127}\ K_{126}\ K_{125}$ ..... $K_2\ K_1\ K_0$.

For *i=0 to 24 do*

$RK_i = K_{31}\ K_{30}$ ……..$K_2\ K_1\ K_0$.

temp1 = sbox(msb<<<3)

temp2 = sbox(msb>>>8)

$Lsb = temp1 \oplus temp2 \oplus lsb \oplus RK_i$.

msb = player(msb)
lsb = player(lsb)

swap(msb, lsb)

keyschedule(key).

End

Output*: Ciphertext*

Here 'i' specifies the round number.

2.6 Key Scheduling Algorithm

Key scheduling of ANU is very robust and it is motivated by the PRESENT [6] cipher. Key scheduling algorithm of ANU performs following operations.

fridaya@uef.fi

128-bit key scheduling algorithm-

```
keyschedule(key)
  {
    key = (key <<< 13)

    key[K3 K2 K1 K0] = Sbox[K3 K2 K1 K0].
    key[K7 K6 K5 K4] = Sbox[K7 K6 K5 K4].

    key[K63 K62 K61 K60 K59] = [K63 K62 K61 K60 K59] ⊕ i.
  }
```

3 Hardware Implementation

FPGA is a reprogrammable device which can be configured by the user in different ways. It has different components like LUT's, Flip-Flops, FPGA slices, logic elements and many more [8, 9]. The performance of any FPGA device is dependent on all these parameters [10, 11] and the devices which has more number of LUT's, Flip-Flops, FPGA slices etc. are generally preferred for the implementation complex circuits. Circuit performance is calculated with respect to different parameters such as throughput, power consumption, latency, efficiency, the area required, the maximum frequency that can be applied to the circuit, propagation delays etc.

The FPGA implementation of lightweight cipher is application dependent [11] and achieving the efficiency in every parameter is the challenging task and finally it's a tradeoff between important design metrics [7].

By considering the trade-off between the parameters, we have implemented 8-bit datapath architecture of ANU cipher. The proposed architecture is mainly designed to have efficiency in throughput, area and power consumption.

Figure 2 shows the top level layout for proposed architectures of ANU.

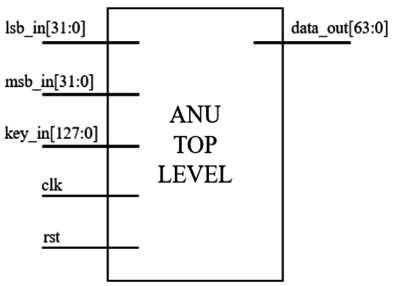


Fig. 2. A Top level model of ANU

3.1 Proposed 8-Bit Datapath Architecture

Figure 3 shows 8-bit datapath serialized for ANU cipher. This architecture consists of two 32-bit registers i.e. LSB and MSB which are used as shift registers. To reduce the number of clock cycles and speed up the operations, the architecture uses two parallel S-Box simultaneously.

fridaya@uef.fi

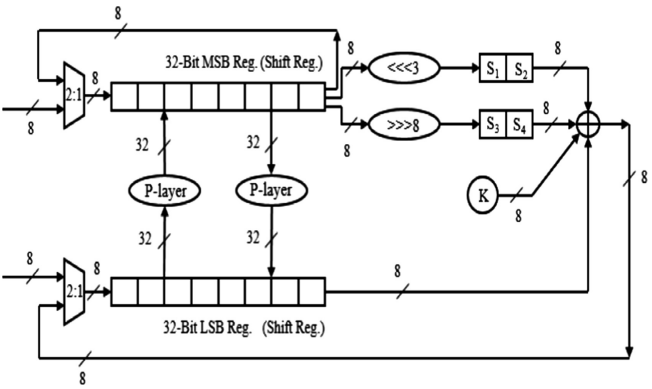


Fig. 3. 8-bit datapath architecture of ANU

Total 4 states of FSM are required to implement this architecture. Initially, the 64-bit data is loaded into both of these 32-bit shift registers i.e. in state 0. After the successful loading of the data, the encryption process starts from state 1. In this state, the 8-bits corresponding to <<<3 and >>>8 will be selected. After selecting these bits, substitution and EXOR operations with LSB 8-bits of a 128-bit key will be performed on these 8-bits in one clock cycle. To perform this operation on 32-bit, it requires a total of 4 clock cycles. Note that, after performing the substitution and EXOR operation on each 8-bits, data from both registers as well as in key register will be right circularly shifted by 8 (>>>8) and hence MSB, LSB and key registers act as a shift register in state 1. After that, bit permutation of MSB and LSB along with key scheduling will be performed in state 2. Hence for completing one round, this architecture requires 4 + 1 = 5 clock cycles.

Since ANU has a total of 25 rounds, to encrypt the block of 64 bits, it requires a total of 5*25 = 125 clock cycles. The flowchart of the circuit operation is shown in Fig. 4.

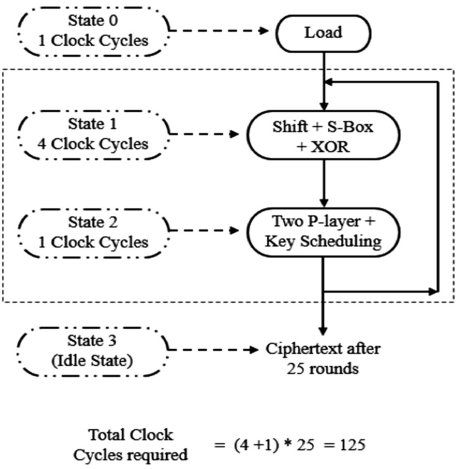


Fig. 4. Clock cycles required per state for 8-bit datapath architecture.

fridaya@uef.fi

Key Scheduling

Architecture for key scheduling is shown in Fig. 5. Two separate S-Box are used for the key scheduling to reduce the FSM overhead in the implementation. Initially, the key is loaded in a 128-bit key register and the key scheduling is performed in state 2 of the data encryption process to avoid extra clock cycles required for the key scheduling.

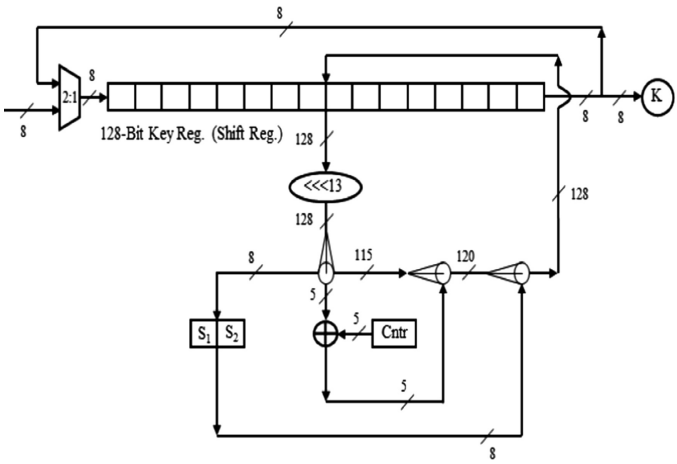


Fig. 5. Key scheduling architecture for ANU

Gate Equivalents

Gate Equivalents (GE) is also an important parameter for implementing lightweight cipher in a constrained environment. GE required for ANU are calculated using standard ASIC library IBM 8RF (0.130 μ). Standard library values for the different gates are shown in Table 3 [12].

Table 3. GE required for each logic circuit

Gate	EXOR	MUX	DFF	AND	OR
GE	2	2.25	4.25	1.25	1.25

Tables 4 and 5 show the GE calculation for this architecture. The control logic for this architecture implementation requires 5 flip-flops for the round counter, 2 flip-flops for state machine and 2 flip-flops for state 1 control. Hence control logic requires a total of 9 flip-flops. Shifts and permutation layers don't require any GE as these layers are implemented using wires ideally.

Total GE = Data-Layer + Key Schedule = 490.25 + 620 = 1110.25 GE. Hence for implementation of the 8-bit architecture of ANU cipher requires around 1110 GE.

fridaya@uef.fi

Table 4. GE calculation for D1

Data-layer	GE required
2 32-bit registers	2*32*4.25 = 272
3 4-bit EXOR	3*8*2 = 48
4 S-Box	4*24 = 96
2 8-bit MUX	2*8*2.25 = 36
Control Logic	9*4.25 = 38.25
P-Layer	0
Shift	0
Total	490.25

Table 5. GE calculation for key scheduling

Key schedule	GE required
128-bit register	128*4.25 = 544
5-bit EXOR	5*2 = 10
2 S-Box	24 + 24 = 48
8-bit MUX	8*2.25 = 9
Total	620

3.2 Performance Metrics

(1) *Platform*

Xilinx FPGA board with ISE Design Suite 14.7 is used for the implementation. ANU is implemented on LUT-4 as well as LUT-6 based devices to get a clear idea about its performance in different environment. Spartan-3 (xc3s700an-5fgg484) and Virtex-4 (xc4vlx25-12ff668) devices are used for LUT-4 based implementation with speed grade of −5 and −12 respectively. Spartan-6 (xc6slx45t-3fgg484) and Virtex-5 (xc5vlx50t-3ff1136) devices are used for LUT-6 based implementation with speed grade of −3. Since 13.56 MHz frequency is the ISO standardized for a smart card [13], all the throughput results are calculated for the frequency of 13.56 MHz.

(2) *Area*

The area metric includes LUT's, flip-flops and FPGA slices required to implement the cipher on FPGA. High-performance FPGA with LUT-6 gives the most compact solution for hardware implementation as compared to LUT-4 based FPGA's. Spartan-6 devices are mainly designed for the resource-constrained environment. The architecture should be designed in such a way that it will require less area to implement.

(3) *Speed of Operations*

Speed includes main parameters as latency, maximum frequency, and throughput. Latency is the number of clock cycles required to encrypt one block of 64-bit data. Since ANU is a 25 rounds lightweight block cipher, latency is the total number of clock

fridaya@uef.fi

cycles required for the completion of 25 round of ANU. It purely depends on the architecture, hence architecture should be designed in such a way that it has less latency. The architecture structure is also responsible for the maximum frequency of the circuit. It is the inverse of the longest path delay.

$$Maximum\ Throughput\ (Thr) = \frac{Maximum\ frequency * Data\ block\ Size}{\text{Latency}}$$

$$Throughput\ at\ 13.56\ Mhz\ (Thr*) = \frac{13.56\ Mhz * Data\ block\ Size}{\text{Latency}}$$

$$Throughput\ per\ slice = \frac{Throughput\ at\ 13.56\ Mhz\ (Thr*)}{\text{Slices}}$$

$$Maximum\ Throughput\ per\ slice = \frac{Maximum\ Throughput\ (Thr)}{\text{Slices}}$$

(4) *Power Consumption*

Power consumption is also an important metric for battery operated devices. It depends on both block size and the frequency. Data block size for all the architecture is the same and frequency is set to 13.56 MHz as this frequency is ISO standardized for a smart card [13].

(i) Static Power
It is the power when the circuit is in an idle or non-switching condition.
(ii) Dynamic Power
It is the power when the circuit is in operating or in switching condition.

(5) *Energy Utilization*

Energy utilization is the total energy required to encrypt one block of the data. It depends on the total power dissipation, the latency of the architecture and the operating frequency of the circuit. For evaluation, it is calculated at a standard frequency of 13.56 MHz [13].

$$Energy = \frac{Total\ Power * Latency}{13.56\ \text{MHz}}$$
$$Energy\ per\ bit = \frac{Energy}{\text{Data block size}}$$

fridaya@uef.fi

4 Results and Evaluation

The proposed architecture is evaluated on four different FPGA platforms with different metrics like throughput, area, power consumption, energy, latency etc. Table 6 shows the performance of proposed 8-bit datapath architecture on different FPGA platforms while Table 7 compares the proposed architecture with the serialized implementation of standard lightweight ciphers. From Tables 6 and 7, it we can be observed that all the performance metrics highly depend on the platform used for the implementation and Spartan-6 platform requires less area while Virtex-5 platform produces the highest throughput as compared to other platforms. Note that '*' in Tables 6 and 7 represents the parameter which is calculated at a fixed frequency of 13.56 MHz.

Table 6. Performance of proposed architecture on different platforms over area and throughput

Architecture	Device	Data Size (bit)	Key Size (bit)	Flip-Flops	LUT	Slices	Latency	Max. Freq. (MHz)	Thr (F_{max}) (Mbps)	Thr* (13.56) (Mbps)	Thr* per slice (Kbps)
Proposed	Spartan-3 xc3s700an-5fgg484	64	128	210	460	265	125	243.253	124.54	6.94	26.188
Proposed	Spartan-6 xc6slx45t-3fgg484	64	128	216	357	93	125	266.766	136.58	6.94	74.623
Proposed	Virtex-4 xc4vlx25-12ff668	64	128	205	668	419	125	329.937	168.92	6.94	16.563
Proposed	Virtex-5 xc5vlx50t-3ff1136	64	128	204	417	178	125	496.475	254.19	6.94	38.988

Table 7. Comparison of proposed architecture of ANU with the serialized implementation of standard ciphers

Device	Architecture/Cipher	Datapath size (bit)	Data size (bit)	Key size (bit)	Flip-Flops	LUT	Slices	Latency	Max. Freq. (MHz)	Thr (F_{max}) (Mbps)	Max Thr per slice (Kbps)
Spartan-3 xc3s700an-5fgg484	Proposed	8	64	128	210	460	265	125	243.253	124.54	469.96
Spartan-6 xc6slx45t-3fgg484	Proposed	8	64	128	216	357	93	125	266.766	136.58	1468.60
Spartan-3 xc3s50-5	LED $(X)^2$ [14] Sect. 3.2	8	64	128	219	388	203	912	142.01	9.97	50
	LED $(X)^4$ [14] Sect. 3.2	16	64	128	227	414	219	528	128.73	15.6	70
Spartan-3 xc3s50-5	LED $(X)^2$ [14] Sect. 3.3	8	64	128	48	167	86	1152	120.75	6.71	80
	LED $(X)^4$ [14] Sect. 3.3	16	64	128	70	248	127	384	117.87	19.65	150
Spartan-3 xc3s200-5	PRESENT (C4) [10]	16	64	80	153	215	124	133	213.81	102.89	829.75
Spartan-3 xc3s200-5	PRESENT (C5) [10]	16	64	128	201	264	151	136	194.63	91.59	606.55

(*continued*)

fridaya@uef.fi

Table 7. (*continued*)

Device	Architecture/Cipher	Datapath size (bit)	Data size (bit)	Key size (bit)	Flip-Flops	LUT	Slices	Latency	Max. Freq. (MHz)	Thr (F_{max}) (Mbps)	Max Thr per slice (Kbps)
Spartan-3 xc3s50-5	PRESENT [15]		64	128	–	–	117	256	114.8	28.46	240
Spartan-3 xc3s50-5	HIGHT [15]		64	128	–	–	91	160	163.7	65.48	720
Spartan-3 xc3s50-5	xTEA [16]		64	128	–	–	254	112	62.6	35.78	140
Spartan-3E xc3s500	SIMON [17]		128	128	–	–	36	–	136	3.60	100
Spartan-3 xc3s50-5	AES [18]		128	128	–	–	184	160	45.6	36.5	200
Spartan-3 xc3s50-5	AES [19]		128	128	–	–	393	534	–	16.86	40

4.1 Power and Energy Consumption

The graphs mentioned in this section from Figs. 6 and 7 shows the energy and power consumption of proposed architecture on different platforms. Different platforms used for FPGA implementation are Spartan-6, Spartan-3, Virtex-5, and Virtex-4. The frequency used while calculating all the performance metrics is 13.56 MHz which is ISO standard and most appropriate for RF applications in IoT [12].

From Figs. 6 and 7, The Spartan-6 platform requires less power and energy consumption as compared to other FPGA platforms.

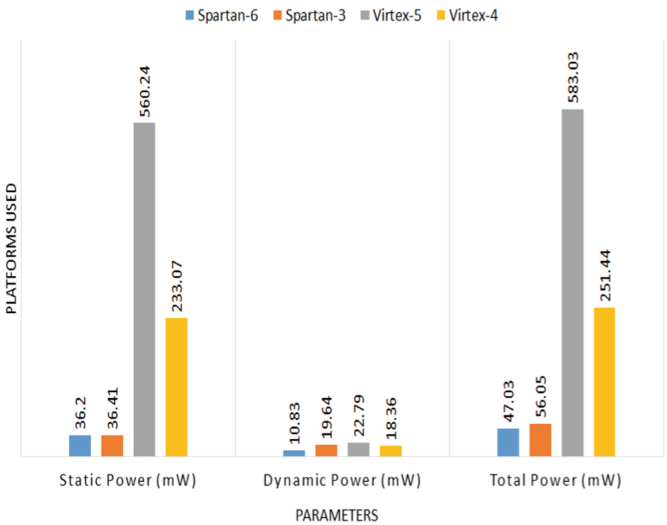


Fig. 6. Power Consumption of proposed architecture on different platforms

fridaya@uef.fi

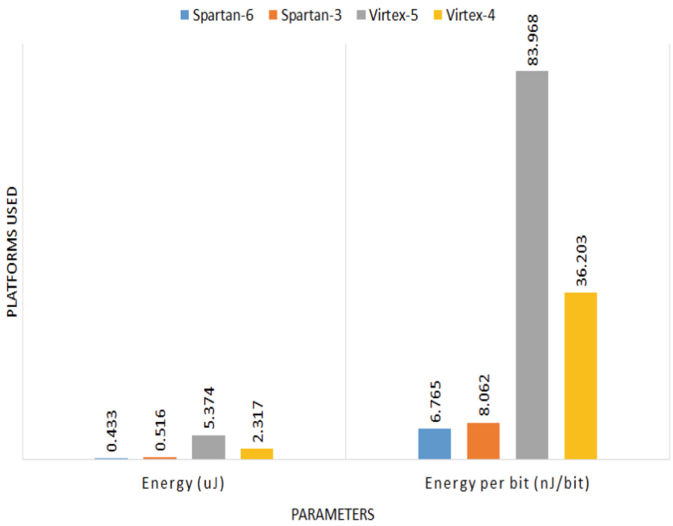


Fig. 7. Energy Consumption of proposed architecture on different platforms

5 Conclusion

This paper presented a detailed implementation of 8-bit datapath design of ANU cipher. The proposed architecture is implemented and tested on four different FPGA platforms. The results vary significantly among these platforms. Spartan platform allows us to implement the datapath designs for the resource-constrained environment with lower throughput requirements and less Gate Counts whereas Virtex boards support the implementation of faster architecture which results in more power consumption.

The 8-bit datapath design of ANU cipher is implemented on various FPGA platforms and compared based on the different design metrics. This paper gives reader the holistic approach of designing smaller datapath architectures for resource-constrained environments. One can choose wisely the suitable FPGA platform for the specific applications.

References

1. National Institute of Standards and Technology. Advanced Encryption Standard (AES). Federal Information Processing Standards Publications – FIPS 197, November 2001. http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf
2. National Institute of Standards and Technology. Data Encryption Standard (DES). Federal Information Processing Standards Publications – FIPS 46-3, October 1999. http://csrc.nist.gov/publications/fips/fips46-3/fips46-3.pdf
3. Poschmann, A.: Lightweight Cryptography - Cryptographic Engineering for a Pervasive World. Number 8 in IT Security. Europ¨aischer Universit¨atsverlag, 2009. Ph.D. Thesis, Ruhr University Bochum (2009)

fridaya@uef.fi

4. Bansod, G., Patil, A., Sutar, S., Pisharoty, N.: ANU: an ultra lightweight cipher design for security in IoT. SCN-15-0848.R1, Security and Communication Networks, DO. https://doi.org/10.1002/sec.169
5. Bansod, G., Patil, A., Sutar, S., Pisharoty, N.: An ultra-lightweight encryption design for security in pervasive computing. In: 2nd IEEE International Conference on Big Data Security on Cloud, pp. 79–84, 2016. Columbia University, New York, USA (2016)
6. Bogdanov, A., Leander, G., Knudsen, L.R., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT - an ultra-lightweight block cipher. In: Paillier, P., Verbauwhede, I. (eds.) Cryptographic Hardware and Embedded Systems, vol. 4727, pp. 450–466, Springer, Berlin, Heidelberg (2007)
7. Xu, T., Wendt, J.B., Potkonjak, M.: Security of IoT systems: design challenges and opportunities. In Proceedings of the 2014 IEEE/ACM International Conference on Computer-Aided Design, pp. 417–423, Piscataway, NJ, USA (2014)
8. Xilinx. Spartan3A FPGA Family, Data Sheet. http://www.xilinx.com/support/documentation/data_sheets/ds529.pdf
9. Xilinx. Spartan6 FPGA Configuration. http://www.xilinx.com/support/documentation/user_guides/ug380.pdf
10. Lara-Nino, C.A. Diaz-Perez, A., Morales-Sandoval, M.: Lightweight hardware architecture for the PRESENT cipher in FPGA. IEEE Trans. Circ. Syst. I: Regul. Pap. vol. 64, 1–12 (2017)
11. Okabe, T.: FPGA Implementation and Evaluation of lightweight block cipher – BORON. Int. J. Eng. Dev. Res. **3**(4), 2017. ISSN: 2321–9939
12. Dahiphale, V., Bansod, G., Patil, J.: ANU-II: a fast and efficient lightweight encryption design for security in IoT. In: 2017 International Conference on Big Data, IoT and Data Science (BID), pp. 130–137 (2017)
13. Identification Cards—Contactless Integrated Circuit Cards—Proximity Cards—Part 2: Radio Frequency Power and Signal Interface, document ISO/IEC 14443-2, August 2010
14. Anandakumar, N.,N., Peyrin, T., Poschmann, A.: A very compact FPGA implementation of LED and PHOTON. In: INDOCRYPT, Lecture Notes in Computer Science, vol. 8885, pp. 304–321 (2014)
15. Yalla, P., Kaps, J.-P.: Lightweight cryptography for FPGAs. In: International Conference on Reconfigurable Computing and FPGAs, pp. 225–230. IEEE (2009)
16. Kaps, J.-P.: Chai-Tea, cryptographic hardware implementations of xTEA. In: Chowdhury, D.R., Rijmen, V., Das, A. (eds.) INDOCRYPT, vol. 5365, pp. 363–375. Springer, Heidelberg (2008)
17. Aysu, A., Gulcan, E., Schaumont, P.: SIMON says, break the area records for symmetric key block ciphers on FPGAs. In: IACR Cryptology ePrint Archive (2014). http://eprint.iacr.org/2014/237
18. Chu, J., Benaissa, M.: Low area memory-free FPGA implementation of the AES algorithm. In: 2012 22nd International Conference on Field Programmable Logic and Applications (FPL), pp. 623–626. IEEE (2012)
19. Kaps, J.-P., Sunar, B.: Energy comparison of AES and SHA-1 for ubiquitous computing. In: Zhou, X., Sokolsky, O., Yan, L., Jung, E.-S., Shao, Z., Mu, Y., Lee, D.C., Kim, D.Y., Jeong, Y.-S., Xu, C.-Z. (eds.) EUC Workshops 2006. LNCS, vol. 4097, pp. 372–381. Springer, Heidelberg (2006)

fridaya@uef.fi

Check for updates

An Evolutionary Mutation Testing System for Java Programs: eMuJava

Muhammad Bilal Bashir[1(✉)] and Aamer Nadeem[2]

[1] IQRA University, Karachi, Pakistan
bilal.bashir@iqraisb.edu.pk
[2] Capital University of Science & Technology, Islamabad, Pakistan

Abstract. Mutation-based testing is costly but we can bring it down with combination of evolutionary testing approaches. Search-based mutation testing combines mutation and evolutionary testing to exploit the advantages offered by both of these techniques. The evolutionary techniques like genetic algorithm supports automating the test case generation during software testing that can reduce a lot of testing resources. The processes of mutation testing can also be automated to further save the testing cost. In this research, we present a testing tool, eMuJava that can perform mutation-based testing of Java-based programs automatically. eMuJava is fully automated system that performs all the activities of mutation testing and test case generation automatically. For mutation testing, it supports conventional as well as object-oriented mutation operators. eMuJava is implementing four testing techniques including three different genetic algorithms and random testing. The system offers complete control to the tester to perform testing of Java programs, allows monitoring of all the steps that tool performs, and all intermediary outputs that the tool generates. We have evaluated the tool by performing extensive experiments on Java programs. The outcomes of an experiment include test case set, mutation score, and statistical information about the experiment. We have also statistically evaluated the experimental results to prove their effectiveness. The generated test cases are further analyzed on different set of mutation operators to evaluate their strength.

Keywords: Adaptable mutation · Control flow · Object-oriented paradigm · Search-based mutation testing · Genetic algorithm · Equivalent mutant · Suspicious mutations · Two-way crossover

1 Introduction

Evolutionary testing takes the responsibility of generating test cases using various techniques like Genetic Algorithms [1]. Genetic algorithm can generate set of test cases to achieve set of testing goals automatically. There are many activities that genetic algorithm performs during its execution but fitness evaluation is core

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 847–865, 2019.
https://doi.org/10.1007/978-3-030-22868-2_58

fridaya@uef.fi

activity among all. The fitness function evaluates a given test case on the basis of some desired testing goal and guides the search to achieve the goals quickly. We find some variations of fitness functions [7,12,14,18,20,21,23,26,28,38] proposed by researchers in the literature.

Mutation testing intentionally injects faults in the code and then tries to generate test cases that can uncover them. This particular type of testing technique helps generating a set of potential test cases that can find out real software bugs from the code. Mutation testing requires execution of a lot of test cases on original and mutant (programs with faults) programs that make it computationally expensive. Another inherited problem with this approach is generation of equivalent mutants. Such mutants cannot be detected. This is another interesting area for researchers as the literature survey [4,5,9–11,24,36] shows.

Search-based mutation testing automates process of test case generation during mutation testing using evolutionary testing approaches like genetic algorithm. The whole process merges the steps involved in mutation and evolutionary testing. Initially mutant programs are produced using two inputs; one of them is the program under test whereas the other is mutation operators. Then genetic algorithm tries to detect all non-equivalent mutants with the help of suitable test cases that it produces. To catch a mutant, it has to satisfy three conditions (reachability [6], necessity [2], and sufficiency [2]). If a test case satisfies these conditions, it uncovers the injected fault in the code. Literature witnesses some research [6,30,31,34,35,37,39] but this area is relatively new and open for research.

Testing can be laborious if it is performed manually. Research shows only test case generation part can consume up to 70% or even more of the total testing time. These issues bring in the need of automation. Besides good testing practices, it is equally important to automate the testing process and activities involved in it. Evolutionary mutation testing has inherited problem of being computationally expensive and more effort demanding due to mutation testing hence testers may feel the need of an automated solution for aid. With the aid of automated solution the whole evolutionary mutation testing process can be sped up.

In this research we present eMuJava; **e**volutionary **Mu**tation tool for **Java** programs. eMuJava provides four different approaches for automatic test case generation It is a complete automated solution as all the activities involved in the testing process are done by the tool with hardly any effort required from the tester. The tester gets full control over the testing process and can monitor the activities performed by the tool. Tester can also see the intermediary outputs generated by the tool and at the end a brief analysis of the testing results are presented. eMuJava shows all the generated test cases, their fitness, and effective test cases.

2 Related Work

We have found some related tools in literature, which are implemented by researchers as proof of concept. Mutation testing tools [8,11,15,19,22,27,32,41,42] are in large number as compared to evolutionary testing tools [14,18,25,28].

fridaya@uef.fi

Although we find a lot of tools in the literature but only few of them are actually available for download and use (for example [14,15,19,28] and few more).

The literature survey shows, EvoSuite [34] is the only available tool at the moment that can perform search-based mutation testing. It also supports branch-coverage, which is one of the coverage criteria to perform control-flow testing. EvoSuite can test Java-based programs but it can handle one class at a time. That means EvoSuite only supports unit level testing though tester can feed large packages and projects into the tool as input. EvoSuite supports a small set of mutation operators [16,29] for structured paradigm. After it receives input, it generates test case set for specified coverage criteria using various techniques including [13,17,33]. The generated test cases uses JUnit [43] format, which is a testing tool of Java programs. JUnit is used to generate test cases to perform unit level testing.

3 eMuJava

The eMuJava system is designed to perform evolutionary mutation testing of Java-based programs The tool is implemented in Java programming language in version 7 while the recent updates are done in version 8. Our tool supports unit level (one class) as well as integration level (more than one class) testing. It is a complete solution for the testers as it does not need any other component to function and perform testing of the programs. eMuJava is a multi-platform tool so it is capable of running on any operating system. The complete source code of the tool is available online and researchers can extend or change it as per need.

3.1 Mutation Operators

One of the important operations of eMuJava system is mutant generation. eMuJava inserts one fault per mutant and faults are inserted using predefined mutation operators. There is no limit on number of mutants to be generated so eMuJava generates all possible mutants from original program. It supports 10 operators to introduce mutations; five of the operators are conventional operators, while five operators are object-oriented. The conventional mutation operators are selected from [3]. Table 1 lists the limited set of mutation operators.

Offutt, Ma, and Kwon [24] propose a number of object-oriented and feature based mutation operators for Java language. From the aforementioned research, we have selected five mutation operators and have implemented them in eMuJava. Table 2 presents list of those mutation operators along with their description.

We have chosen these operators on the basis of their feature coverage and effectiveness. Some mutation operators mostly generate equivalent mutants and some generate easy to kill mutants [10] so it should not have much impact leaving them out from the tool.

fridaya@uef.fi

Table 1. List of mutation operators for structured paradigm

Name	Description
ABS	Insert an absolute value
AOR	Replace the arithmetic operator with another one
LCR	Replace the logical connector in condition
ROR	Replace the relational operator with another one
UOI	Insert the unary operator

Table 2. List of mutation operators for object-oriented paradigm

Name	Description
IOD	Delete the overriding method
PNC	Change constructor call with the subclass type
OMD	Deletie the overloading method
JID	The initialization of a data member
EOC	Replace == operator with equals() and vice versa

3.2 Automated Test Case Generation

We have implemented four different techniques for automated test case generation. Tester can select among these approaches through tool interface while beginning the experiments. The main reason of implementing four techniques is to make a comparison of the results produced by each of them, which will help the tester to determine, which technique performs best in case of a specific type of program.

1. *Random Testing:* It generates the test case suite absolutely randomly. All the iterations use randomly generated tests and input data so it does not repair the test cases by any means. Random testing can be very useful in cases when severely random input values are required.
2. *Standard Genetic Algorithm:* This technique uses genetic algorithm to produce the test cases. For first iteration the test cases are produced randomly and executed on the mutant programs. The execution traces are recorded and analyzed on the basis of three conditions required to detect a mutant [6]. This evaluation process produces fitness information, which is used later on to repair the test cases through crossover or biological mutation.
3. *Genetic Algorithm with Improved Fitness Function* [37]*:* This technique uses genetic algorithm and quite similar to the one described in (2). The main difference in this technique is the fitness function [37]. It includes state of the object and control flow of the program as important segment of the fitness. The later phases then use the provided fitness information to improve and repair the test cases so they can achieve the targets (kill the mutants) quickly.
4. *Improved Genetic Algorithm* [40]*:* This technique not only uses novel and better fitness function [37] but it also includes a novel crossover method that

fridaya@uef.fi

performs two way crossover of test cases. Besides these inclusions, it has a novel and better adaptable mutation method for repairing the test cases.

3.3 Configuration Options

eMuJava provides many useful configuration options to the testers to perform the experiments more efficiently and effectively. The tester has to input Java code files for testing including some configurations; mutation operators that tester wants to apply on code under test, testing technique to generate test cases with, size of the initial population, allowed iterations, type of the crossover, and rate of mutation. eMuJava offers some additional configurations that aids the testing process. These customization options and their details are presented in Table 3.

Table 3. Configuration options in eMuJava tool

Option	Description
Range of mathematical integers	This range specifies the limit for mathematical integers that may be required by certain method invocations as arguments in a test case. eMuJava produces an integer that remains between 0 and the defined boundary by this parameter
Range of ASCII characters	This configuration parameter deals with character input generation. If the program needs a single character, it is generated randomly
Timeout for genetic algorithm	eMuJava generates driver classes. This is because each test case needs to be executed on original as well as mutant program. Due to some syntax error in the original or instrumented code, the Java compiler will fail to generate .class files. In this case, eMuJava may go on and wait for indefinite amount of time. To handle this situation, we have introduced a timeout value that allows the tool to carry on with next phase without further waiting for the compiled code
Maximum GA threads	We have implemented multi-threading in eMuJava for concurrent test case generation. By default just a single thread is run assuming an average speed computer. But if the computer on which tests are run is extremely powerful then this feature will help to take full advantage of powerful resources. Multi-threading will reduce amount of time to generate test cases
Population selection percentage	After every iteration eMuJava selects a specific percentage of fitter test cases for the next iteration if target remains alive. This parameter allows the user to change the percentage of test cases to be used for next iteration
Method class sequence count	This parameter allows the tester to change the maximum number of method calls in a test case. The actual length for methods call sequence in a test case is generated randomly between 0 and this parameter

fridaya@uef.fi

3.4 Features of eMuJava

eMuJava is a GUI based testing tool that makes testing easier for the tester. Tester can learn to use the tool and with minimum of effort test cases can be generated.

There are many salient features that make eMuJava a useful testing tool. These features allows the tester to prepare the testing environment, load the program under test, generate mutants, filter mutants, view the generated test cases along with effective test cases and statistics towards the end of test case generation process. Figure 1 presents two screenshots of eMuJava and next we provide brief details of the main features in eMuJava.

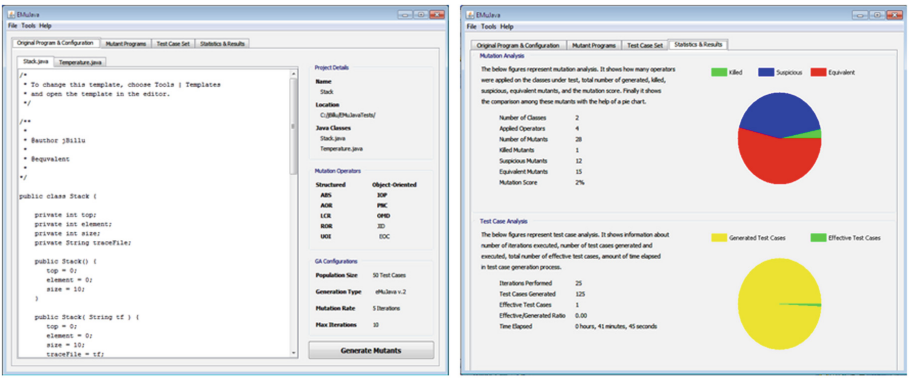


Fig. 1. Screenshots of eMuJava

After the tool is run, tester creates a new project and provides information about the program under test in shape of Java classes, mutation operators, and test case generation technique and its configuration.

The first tab in eMuJava interface is **Original Program & Configuration** that allows the tester to view the Java classes loaded in the tool for testing purpose. Figure 1(a) presents the screenshot of this feature. In the right panel of the GUI tester can view details about the program under test including name of the classes, mutation operators, and some useful configuration information. The configuration includes name of the technique, population size, mutation rate, and maximum number of iterations. The tester then begins the process by clicking on Generate Mutants button that generates all possible mutants.

In the **Mutant Programs** tab, tester can see all the generated mutants by the tool. Complete source code of the mutant is displayed along with the highlighted statement where mutation is applied. Initially all the mutants are shown *Alive*. Once the tool completes the test case generation process, Mutant Programs tab shows three categories of the mutants including *Alive*, *Suspicious*, and *Killed*. Tester can go through the mutants to determine if there are any

fridaya@uef.fi

equivalent mutants there. In case some equivalent mutants are found, tester can use **Filter Mutants** interface to filter out equivalent mutants so when eMuJava generates test cases, it does not consider them.

Now tester is all set to begin the process. The **Test Case Set** tab is updated dynamically at runtime. There are two sections of this tab, first one shows the generated test cases to kill a mutant along with their fitness values. In the second section, tool lists effective test cases that have killed mutants.

After completion of the test case generation process, eMuJava presents detailed results in **Statistics & Results** tab as shown in Fig. 1(b). The information it presents about the code under test includes number of classes, applied mutation operators, total number of generated, killed, suspicious, and alive mutants. The most important information is about the experiment including mutation score, iterations, and total number of test cases.

3.5 Availability

We have made the source code of eMuJava available online for researchers so they can extend it and change it as per need. We are also open for suggestions and feedback from the users in case they experience any problem or want us to add a new feature in it, we will be happy to help them. The URL to download the source code of the tool and programs that are used for experiments is https://github.com/bilalbezar/eMuJava. All the required information to compile and run the tool is also there for the users.

4 eMuJava Architecture

eMuJava is complex system, which is further divided into three large sub-systems including Mutant Generator, Code Instrumentation & Compiler, and Test Case Generator. Each sub-system has certain inputs, computations (responsibilities), and one or more outputs. All the sub-systems are linked together in such a way that output of one becomes input of the next. Figure 2 presents architecture of eMuJava as a block diagram and in coming subsections, we present details about each major sub-system of the tool along with their inputs, outputs, and responsibilities.

4.1 eMuJava Sub-systems

The implemented system is complex in nature that comprises of three large modules and they are further decomposed into smaller modules for simplicity and for ease of implementation. The details about each module are presented in next paragraphs.

The **mutant generator** receives three inputs from the tester including source code (Java classes), mutation operators, and test configurations. *Project Manager* allows the tester to create new project, load Java classes, select mutation operators, and test configuration. Then Java classes are sent to *Statement*

fridaya@uef.fi

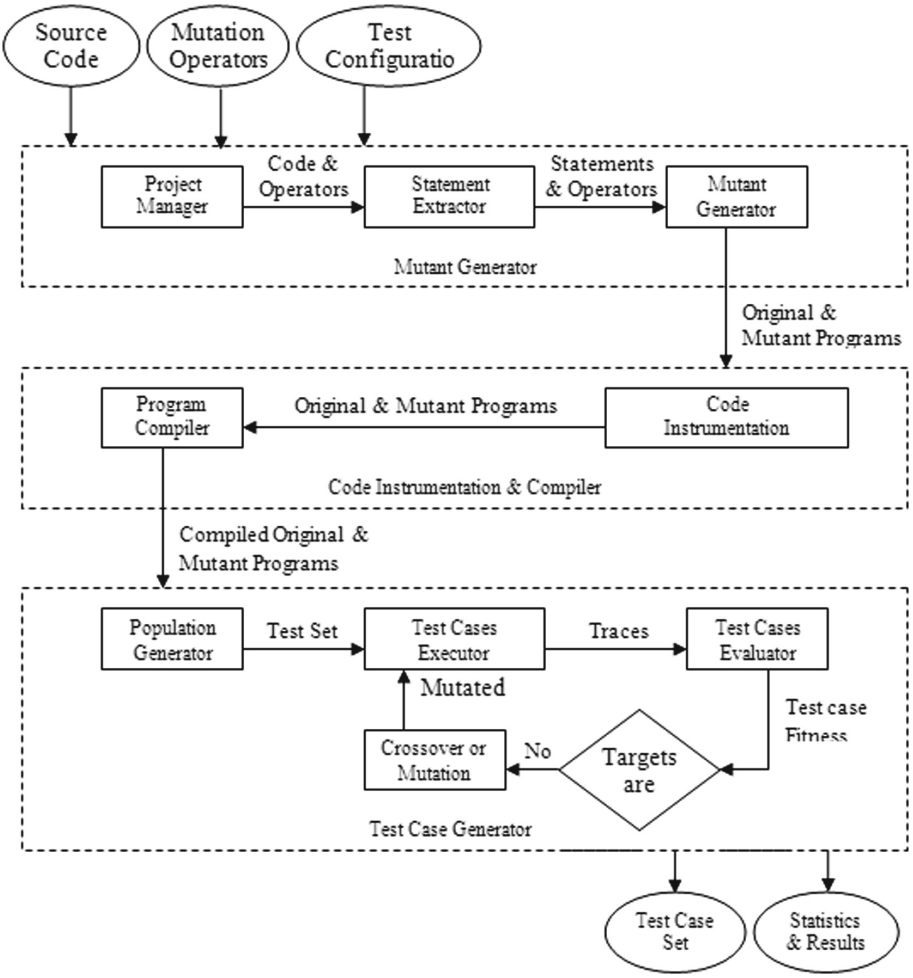


Fig. 2. The eMuJava tool architecture

Extractor that scans, parses, and combines the tokens to form supported program statements (assignment statement, method call, return statement, object creation, while statement and so on) in Java. These statements and mutation operators then become the input of *Mutant Generator* that applies all the operators to the statements and generates mutants, which contain complete code and one mutation each.

The **code instrumentation & compiler** prepares the environment for test case generation process. *Code Instrumentation* receives original and mutant programs as input and it instrument all the instances with additional code so when tests are run on them, traces are generated for analysis. The instrumented programs are sent to *Program Compiler* that uses standard Java compiler to generate .class files for all original and mutant programs.

fridaya@uef.fi

The **test case generator** receives compiled programs as input and generates test cases. This module has been designed to support four test case generation techniques (for details please read Sect. 3.2). This module can generate test cases automatically to detect the mutants. The output includes effective test case suite and some statistical information about the experiment.

4.2 eMuJava Activities

Now we present various different activities that eMuJava performs while generating test cases. After receiving inputs from the user the tool performs following activities till the generation of test case set.

Mutant Generation: Once eMuJava completes identification of program statements from the input code, it starts generating mutant programs. Until now eMuJava is implementing 10 mutation operators. Five of the total ten operators handle features of structured paradigm and remaining five handle object-oriented features like inheritance, or polymorphism.

Population Generation: After generation of mutants, eMuJava goes on to generate test cases, which begins with the generation of random population in first iteration. The tool produces a test case having format proposed by Tonella [14] and its anatomy is presented in Fig. 3. By default eMuJava generates set of 50 test cases, which tester can change before beginning the test case generation. Each test case comprises of object instantiation, series of method calls, method being tested, and input parameters for the methods. Input values are generated random using the specified ranges in configuration (Sect. 3.3).

Fitness Evaluation: Test cases are run on the original and mutant programs and if the mutant remains undetected, fitness evaluation is performed. eMuJava uses execution traces that original and mutant programs generate when test cases are run to evaluate a test case. For random testing, eMuJava does not perform any evaluation at all. If tester chooses to standard genetic algorithm then the tool computes fitness information using the method of Bottaci [6]. For the remaining two techniques, it uses improved fitness function [37] for the same purpose.

Crossover: After evaluating the test cases, eMuJava performs crossover, which is a standard genetic algorithm operation to produce new chromosomes (test cases) for next iteration. To perform the tool picks fitter test cases, which it selects through tournament selection. The selected test cases are grouped in pairs. There are two types of crossover methods implemented in the tool; single-point crossover, and two-way crossover [40]. The single-point crossover is more like conventional method that picks a random point in both test cases to cross them over. On the other hand, the two-way crossover uses state fitness to decide upon the point of crossover.

Biological Mutation: After certain number of iterations when crossover fails to form the required test cases, biological mutation helps the cause. Just like

fridaya@uef.fi

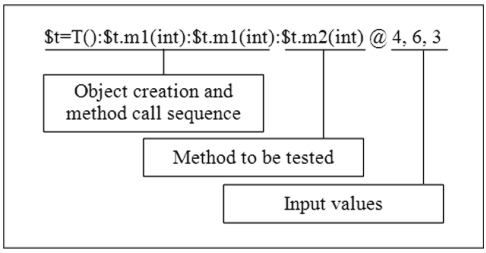


Fig. 3. Structure of a test-case [28]

crossover, eMuJava supports two different biological mutations. The type of mutation to be performed is decided on the basis of value of state fitness. If the state fitness for a test case is evaluated as "0.0" that means the method call sequence is correct and all it needs is a different combination of input parameters for method under test. In such conditions, crossover will not help because it just changes the method call sequence so the tool increases mutation rate by 1. But on the other hand, if value of the state fitness in non-zero that shows not only input parameters but the method call sequence also needs some repair. So eMuJava changes methods call sequence as well as input parameters.

5 eMuJava Algorithms

There are six main activities that the tool performs while generating test cases including mutant generation, population generation, test case execution, test case evaluation, performing two-way crossover, and applying biological mutation on test cases. The first three are quite straight forward so we are not presenting their algorithms here. We present the algorithms for last three of the six activities below.

5.1 Test Case Evaluation

To perform evaluation of a given test case, eMuJava has implemented the algorithm shown in Fig. 4. Input of this algorithm includes target (the alive mutant) to be achieved, trace of execution, and test case set. After that the evaluation begins on execution traces. The algorithm processes the input and computes three different costs for every single test case. These costs include reachability cost, necessity cost, and sufficiency cost. Besides the traces, algorithm uses control flow information and state of the object. The output of this algorithm is fitness values computed for all the input test cases.

5.2 Two-Way Crossover

The algorithm presented in Fig. 5 explains how the two-way crossover works. It performs crossover in pair of test cases, which are selected though tournament

fridaya@uef.fi

selection. Input of this algorithm is not only the selected test cases in pairs but also the fitness values they posses. Those test cases, which do not get selected, are ignored. Every pair of test cases shares similar *state_fitness*. A pair with 0.0 *state_fitness*, goes through single-point crossover and rest of the pairs go through two-way crossover. On successful completion, this algorithm generates new offsprings that are ready to be fed into next iteration. generates after performing crossover.

5.3 Adaptable Biological Mutation

eMuJava implements biological mutation algorithm presented in Fig. 6. This algorithm comes into play when crossover fails to produce the required test case after certain number of iterations. This algorithm checks all the test case one by one and mutates method call sequences or input parameters. It can also modify mutation rate on the basis of *state_fitness*.

6 Experiment Results

In this section, we present subject programs, procedure of experiments, experimental results, statistical analysis, and analysis on the results. eMuJava supports multi-threading and allows the tester to run multiple threads that can generate test cases in parallel. By default the tool runs single thread and experiments have revealed that eMuJava can produce 180–200 test cases per minute after going through all the phases of genetic algorithm in multiple iterations.

6.1 Selected Programs

We have chosen 20 sample programs for experimentation from various different domains. These programs are available online along with the source code of eMuJava, which testers can use and modify for further experiments. Table 4 lists the programs that are used to perform extensive experiments using our tool. The table also presents other related information about the programs like total number of class and methods in them, lines of code and total number of mutated that are generated for experiments.

6.2 Experiments Procedure

eMuJava is used to evaluate testing techniques (see Sect. 3.2) with extensive experiments. We have tried to test them under two questions; which testing technique is capable of detecting maximum mutants and among all the testing techniques, which one consumes less amount of time to detect more mutants. For first phase of experiments, we restricted eMuJava to certain number of attempts for a given mutant. It iterates 10 times to detect a mutant and check which of the four test generation techniques scores highest mutation score. For second phase we let the tool to run until all the non-equivalent mutants are not detected.

fridaya@uef.fi

```
Algorithm:    EvalTestCases
Input:        test case set TCSet
              target mutant Traget
              execution traces Traces
Output:       test case evaluation results Fitness
Declare:      test case fitness set Fitness
              test case fitness F as String
              approach_level as integer
              state_fitness as double
              coverage_fitness as double
              reachability_cost as String
              necessity_cost as double
              sufficiency_cost as String
Begin: EvalTestCases
 1.  for each T in TCSet do
 2.    read traces of T from Traces
 3.    if T fails to execute mutation then
 4.      evaluate approach_level
 5.      if predicate contains state variables then
 6.        evaluate branch_distance of state variables as state_fitness
 7.      end if
 8.      if predicate contains local variables then
 9.        evaluate branch_distance of local variables as coverage_fitness
10.      end if
11.      assign approach_level, state_fitness, coverage_fitness as reachability_cost
12.      assign a constant C to necessity_cost
13.      assign a constant C to sufficiency_cost
14.    else if T fails to introduce infection then
15.      assign reachability_cost as 0
16.      if mutation involves object then
17.        evaluate branch_distance of object's state variables as necessity_cost
18.      else if mutation does not involve object then
19.        evaluate branch_distance as necessity_cost
20.      end if-else
21.      assign a constant C to sufficiency_cost
22.    else if T fails to propagate infection in output then
23.      assign reachability_cost as 0
24.      assign necessity_cost as 0
25.      count the nodes having same values after mutation as sufficiency_cost
26.      if execution path differs in original and mutated programs then
27.        Add flag with T as suspicious
28.      end if
29.    else do
30.      assign reachability_cost = necessity_cost = sufficiency_cost = 0
31.    end if-else-if
32.    F = [reachability_cost ; necessity_cost ; sufficiency_cost]
33.    add F to Fitness set
34. end for
35. return Fitness
    end EvalTestCases
```

Fig. 4. Algorithm for test case evaluation

fridaya@uef.fi

```
Algorithm:    TWCrossoverTests
Input:        test case set TCSet
Output:       test case set TCSet
Declare:      test case pair T1 and T2 as TestCase
Begin: TWCrossoverTests
   1. for each pair of test case in TCSet do
   2.   pick T1 and T2 from TCSet
   3.   if T1 and T2 have state_fitness as 0 then
   4.     perform two-way crossover and add all 4 offsprings in TCSet
   5.   else
   6.     perform conventional crossover and add both offsprings in TCSet
   7.   end if-else
   8. end for
   9. return TCSet
   end TWCrossoverTests
```

Fig. 5. Algorithm for two-way crossover

```
Algorithm:    TWCrossoverTests
Input:        test case set TCSet
Output:       test case set TCSet
Declare:      test case pair T1 and T2 as TestCase
Begin: TWCrossoverTests
   1. for each pair of test case in TCSet do
   2.   pick T1 and T2 from TCSet
   3.   if T1 and T2 have state_fitness as 0 then
   4.     perform two-way crossover and add all 4 offsprings in TCSet
   5.   else
   6.     perform conventional crossover and add both offsprings in TCSet
   7.   end if-else
   8. end for
   9. return TCSet
   end TWCrossoverTests
```

Fig. 6. Algorithm for adaptable biological mutation

6.3 Experiment Results

We have presented the results of experiments in Figs. 7, 8 and 9. Figure 7 shows average mutation score achieved by all the testing techniques and it can be noticed that the improved genetic algorithm [40] is able to gain highest score in a fixed number of attempts for every program.

In the second phase of experiments, we run the tool again on all twenty programs while trying to detect all the non-equivalent mutants to see how much effort is consumed by each technique in shape of total number of iterations and total number of test cases executed. The results show that again the improved genetic algorithm [40] has used less amount of effort (iterations) as compared to other three techniques. Figure 8 shows the results in form of bar chart; the horizontal axis shows techniques and vertical axis shows iterations used by them.

fridaya@uef.fi

Table 4. Source programs used in experiments

Programs	Classes	Methods	LoC	Mutants (M)
AutoDoor	1	9	112	111
BankAccount	2	14	116	180
BinarySearchTree	1	6	119	189
Calculator	1	7	60	97
CGPACalc	1	4	105	111
CLI	2	6	160	24
Collections	3	16	386	131
Compress	3	11	291	87
Crypto	1	2	81	120
CSV	2	4	129	145
ElectricHeater	1	9	120	146
HashTable	1	8	98	103
JCS	1	5	135	74
Lang	2	12	405	51
Logging	2	11	133	95
Math	1	4	109	103
Stack	2	13	144	107
TempConverter	1	8	60	106
Text	1	3	143	97
Triangle	1	5	99	147
Total	30	157	3005	2224

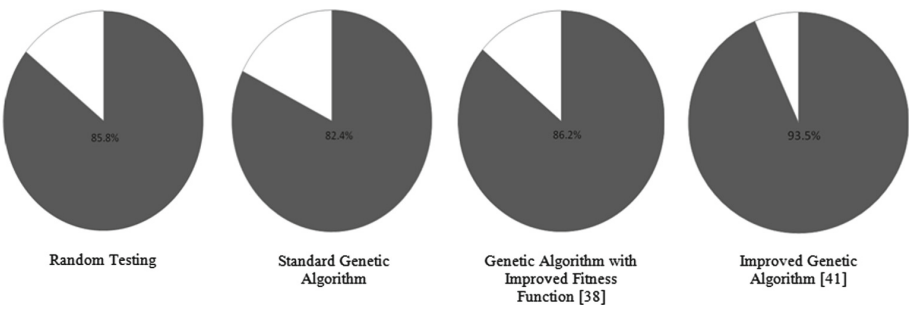


Fig. 7. Mutation score comparison among test case generation techniques

Figure 9 shows the number of iterations used by each testing technique to gain 100% mutation score. We have again used bar chart for this purpose; the horizontal axis shows the techniques and vertical axis shows the number of iterations used by them. Again the improved genetic algorithm [40] is found the best among all.

fridaya@uef.fi

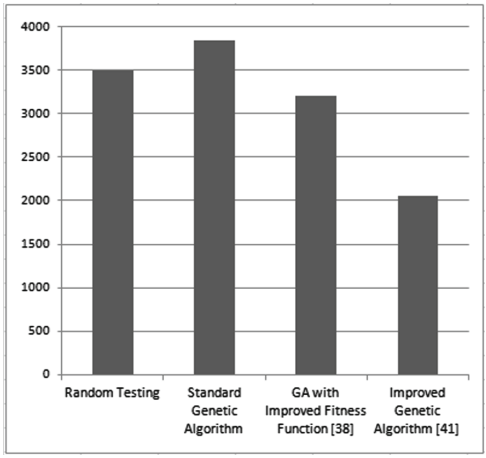


Fig. 8. Iterations used for test case generation

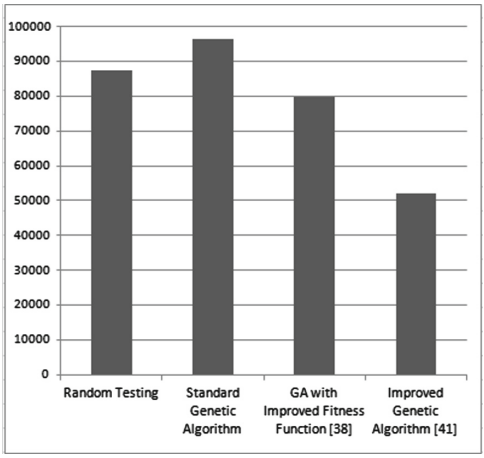


Fig. 9. Test cases executed for test case generation

6.4 Statistical Analysis

We have performed statistical analysis on experiment results to further strengthen our claims. The analysis shows that improved genetic algorithm [40] has produced significant results. We have used U-test and A measure to conduct statistical analysis. These statistical analysis is performed using R [44] tool, which is a famous statistical tool.

Normality Test. Most of the times, the data produced by random (for example GA) algorithms is not normalized. Before selecting the statistical test, we need

fridaya@uef.fi

to be sure if data is normalized or not. For this we have used Shapiro Wilk test. We have taken "0.050" for p-value to declare if a data set is normalized or not. The data sets having low p-value than 0.050 is non-normal and data sets with high p-value is normal. Next we present the details of our experiments.

1. Random Testing: The effect size is computed as 0.67 for improved genetic algorithm [40] in comparison to random testing.
2. Standard Genetic Algorithm: The effect size is computed as 0.76 for improved genetic algorithm [40] in comparison to standard genetic algorithm.
3. Genetic Algorithm with Improved Fitness Function [37]**:** The effect size is computed as 0.72 for improved genetic algorithm [40] in comparison to genetic algorithm with improved fitness function.

6.5 Discussion

eMuJava is a complete automation to test Java-based programs. It performs mutation testing in combination with four automated test case generation techniques. The ability to generate test cases automatically for mutation testing makes eMuJava attractive and useful tool. Besides detecting normal mutants during testing, it can also detect suspicious mutants. The concept of suspicious mutant was coined in [36], which refers to a special mutant, which although exercises different execution path on a test case as compared to original program, yet it produces same output. No other testing tool provides this feature.

The statistical analysis shows that random testing has performed poorly as compared to other approaches. After random testing, one of the variants of genetic algorithm with improved fitness function [37] has shown brief improvement. Among all the techniques, the improved genetic algorithm [40] has shown best performance that has also been proven by *U-test*. Finally the *A measure* also shows that improved mutation scores are attained by improved genetic algorithm.

These experiments have produced interesting results. We notice that more than one faults in a program can cause it to behave differently. For example a predicate `a>b` can be changed like `a>C` where `C` is a constant. This simple change can cause the program to become hard to kill or may be it becomes equivalent mutant. Similarly the program may become easy to kill mutant. The experiment results show that improved genetic algorithm [40] has caught maximum faults including some hard to kill mutants as well.

7 Conclusion

Evolutionary mutation testing combines mutation testing with evolutionary testing to handle shortcomings of mutation testing (high computational cost). With the help of evolutionary algorithms the effort to produce test cases can be saved. In fact the complete process of performing evolutionary mutation testing can be performed automatically. We have designed a tool that offers performing evolutionary mutation testing on Java-based programs. It has been developed in

fridaya@uef.fi

Java that makes it platform independent. Secondly, it is a complete automated solution that needs no other component to perform the testing, which makes it easy to install and configure. There are many other salient features including; it is a freeware available online for download, it implements both structured and object-oriented mutation operators that no other tool supports, eMuJava supports four types of test case generation techniques and its a GUI based tool that makes it easy to learn and use. The tool is open source so researchers can extend and modify it as per need. We have performed large number of experiments to verify and validate its operational capability and results will give confidence to its users to use it in their research studies.

8 Future Work

Currently, the tool offers random testing and three different implementations of genetic algorithm. We will be extending the tool to add support for more evolutionary techniques like artificial immune system, particle swarm optimization and so on.

References

1. Holland, J.H.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor (1975)
2. DeMillo, R.A., Offutt, A.J.: Constraint-based automatic test data generation. IEEE Trans. Softw. Eng. **17**(9), 900–910 (1991)
3. Offutt, J., Lee, A., Rothermel, G., Untch, R.H., Zapf, C.: An experimental determination of sufficient mutant operators. ACM Trans. Software Eng. Methodol. **5**, 99118 (1996)
4. Kim, S., Clark, J., McDermid, J.: The rigorous generation of java mutation operators using HAZOP. In: 12th International Conference Software & Systems Engineering and their Applications, December 1999
5. Kim, S., Clark, J., McDermid, J.: Class mutation: mutation testing for object-oriented programs. In: Proceedings of the Net.ObjectDays Conference on Object-Oriented Software Systems, Erfurt, Germany, Oct 2000
6. Bottaci, L.: A genetic algorithm fitness function for mutation testing. In: Proceedings of International Workshop on Software Engineering using Metaheuristic Inovative Algorithms, a workshop at 23rd International Conference on Software Engineering, p. 37 (2001)
7. Wegener, J., Baresel, A., Sthamer, H.: Evolutionary test environment for automatic structural testing. Inf. Softw. Technol., pp. 841–854 (2001)
8. Chevalley, P., Thvenod-Fosse, P.: A mutation analysis tool for Java programs, LAAS Report No 01356, Toulouse, France (2001)
9. Chevalley, P.: Applying mutation analysis for object oriented programs using a reflective approach. In: Proceedings of the 8th Asia-Pacific Software Engineering Conference, Macau SAR, China (2001)
10. Ma, Y.-S., Kwon, Y.-R., Offutt, J.: Inter-class mutation operators for Java. In: Proceedings of the 13th IEEE International Symposium on Software Reliability Engineering, pp. 352–363, Annapolis MD, November 2002

fridaya@uef.fi

11. Alexander, R.T., Bieman, J.M., Ghosh, S., Bixia, J.: Mutation of Java objects. In: Proceedings of IEEE 13th International Symposium on Software Reliability Engineering (2003)
12. McMinn, P., Holcombe, M.: The state problem for evolutionary testing. In: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2003), Lecture Notes in Computer Science, vol. 2724, pp. 2488–2497. Springer-Verlag, Chicago, USA (2003)
13. Harman, M., Hu, L., Hierons, R., Wegener, J., Sthamer, H., Baresel, A., Roper, M.: Testability transformation. IEEE Trans. Software Eng. **30**(1), 3–16 (2004)
14. Tonella, P.: Evolutionary testing of classes. In: Proceedings of the ACM SIGSOFT International Symposium of Software Testing and Analysis, pp. 119–128, Boston, MA, July 2004
15. Offutt, J., Ma, Y.-S., Kwon, Y.-R.: An experimental mutation system for Java. ACM SIGSOFT Software Engineering Notes **29**(5) (2004)
16. Andrews, J.H., Briand, L.C., Labiche, Y.: Is mutation an appropriate tool for testing experiments? In: Proceedings of the 27th International Conference on Software Engineering, St. Louis, MO, USA, 15–21 May 2005
17. Godefroid, P., Klarlund, N., Sen, K.: DART: directed automated random testing, In: Proceedings of the 2005 ACM SIGPLAN Conference on Programming Language Design and Implementation, pp. 213-223, New York, USA, 2005
18. Wappler, S., Lammermann, F.: Algorithms, using evolutionary, for the unit testing of Object-Oriented Software, GECCO05, Washington DC, USA, 25–29 June 2005
19. Ma, Y.-S., Offutt, J., Kwon, Y.-R.: MuJava: an automated class mutation system: research articles. Software Testing, Verification & Reliability **15**(2), 97–133 (2005)
20. Cheon, Y., Kim, M.Y., Perumandla, A.: A complete automation of unit testing for Java programs. In: The 2005 International Conference on Software Engineering Research and Practice (SERP), Las Vegas, Nevada, USA, 27–30 June 2005
21. Cheon, Y., Kim, M.: A specification-based fitness function for evolutionary testing of object-oriented programs. In: Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, Seattle, Washington, USA, 08–12 July 2006
22. Bradbury, J.S., Cordy, J.R., Dingel, J.: ExMAn: a generic and customizable framework for experimental mutation analysis. In: Workshop on Mutation Analysis (2006)
23. Seesing, A., Gross, H.: A genetic programming approach to automated test generation for Object-Oriented Software. ITSSA, vol. 1, no. 2, pp. 127–134 (2006)
24. Offutt, J., Ma, Y.-S., Kwon, Y.R.: The class-level mutants of mujava. In: AST 06 Proceedings of the 2006 International Workshop on Automation of Software Test, p. 7884. ACM, New York, NY, USA (2006)
25. Dharsana, C.S.S., Askarunisha, A.: Java based test case generation and optimization using evolutionary testing. In: Proceedings of International Conference on Computational Intelligence and Multimedia Applications, pp. 44–49, Sivakasi, Tamil Nadu (2007)
26. Liaskos, K., Roper, M., Wood, M.: Investigating data-flow coverage of classes using evolutionary algorithms. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation, London, England, 07–11 July 2007
27. Grun, B., Schuler, D., Zeller, A.: The impact of equivalent mutants. In: Proceedings of the 4th International Workshop on Mutation Testing, Denver, CO, USA (2009)

fridaya@uef.fi

28. Bashir, M.B., Nadeem, A.: A state based fitness function for evolutionary testing of object-oriented programs. In: Proceedings of 7th ACIS International Conference on Software Engineering Research, Management and Applications SERA, Haikou, Hainan Island, China, 2–4 December 2009
29. Schuler, D., Zeller, A.: (Un-)covering equivalent mutants. In: Proceedings of 3rd International Conference on Software Testing Verification and Validation, pp. 45–54 (2010)
30. Fraser, G., Zeller, A.: Mutation-driven generation of unit tests and oracles. In: Proceedings of the 19th International Symposium on Software Testing and Analysis, pp. 147–158 (2010)
31. Mishra, K.K., Tiwari, S., Kumar, A., Misra, A.K.: An approach for mutation testing using Elitist Genetic algorithm. In: Proceedings of 3rd IEEE International Conference on Computer Science and Information Technology, pp. 426–429, Chengdu, China (2010)
32. Madeyski, L., Radyk, N.: JudyA mutation testing tool for Java. IET Soft. **4**(1), 32–42 (2010). Institute of Engineering and Technology
33. Harman, M., McMinn, P.: A theoretical and empirical study of search based testing: local, global and hybrid search. IEEE Tran. Soft. Eng. **36**(2), 226–247 (2010)
34. Fraser, G., Arcuri, A.: EvoSuite: automatic test suite generation for object-oriented software. In: Proceedings of the 19th ACM SIGSOFT Symposium and the 13th European Conference on Foundations of Software Engineering, pp. 416-419, New York, NY, USA (2011)
35. Papadakis, M., Malevris, N.: Automatic mutation based test data generation. In: Proceedings of the 13th Annual Conference Companion on Genetic and Evolutionary Computation (GECCO 2011), pp. 247–248 (2011)
36. Bashir, M.B., Nadeem, A.: Control oriented mutation testing for detection of potential software bugs. In: Proceedings of 10th International Conference on Frontiers of Information Technology (2012)
37. Bashir, M.B., Nadeem, A.: A fitness function for the evolutionary mutation testing of object-oriented programs. In: Proceedings of the 9th International Conference on Emerging Technolgoies (ICET 2013), December 2013
38. Bashir, M.B., Nadeem, A.: A state-based fitness function for the integration testing of object-oriented programs. In: Proceedings of the 10th International Conference on Emerging Technologies (ICET 2014), December 2014
39. Fraser, G., Arcuri, A., Achieving scalable mutation-based generation of whole test suites, empirical software engineering. **20**(3) (2015)
40. Bashir, M.B., Nadeem, A.: Improved Genetic Algorithm to Reduce Mutation Testing Cost. Accepted for Publication in IEEE Access (2017)
41. Jumble. http://jumble.sourceforge.net/. Accessed on March 2018
42. Moore, I., Jester. http://jester.sourceforge.net/. Accessed on March 2018
43. Sourceforge, JUnit. http://www.junit.org/. Accessed March 2018
44. Team, R.D.C., R: A language and environment for statistical computing. In: https://www.r-project.org/. Accessed on 30 November 2017, R Foundation 1288 for Statistical Computing (2008)

fridaya@uef.fi

"STEPS" to STEM

Esther Pearson(✉)

School of Health Sciences, Lasell College, Newton, MA, USA
epearson@lasell.edu

Abstract. The "Science, Technology, Engineering, Precollege Studies" (STEPS) program was developed in 1988 by Dr. Esther Pearson. Programs like this and others provide consistent pipelines to STEM careers by taking into consideration cognitive, affective and psychomotor behaviors, as well as, addressing gender-bias in STEM career preparation. The STEPS program has served thousands of youth over the past two decades to provide academic support and mentoring to minorities and women students. The STEPS program focuses on demonstrating a connected learning approach to STEM academics. Students participate in mentoring through the STEM pipeline of course choices, extra-curricular activities, and exposure to STEM practitioners. Students learn to overcome the challenges that prevent successful matriculation into STEM fields. Minority and women students in elementary through college in the Boston and greater Boston areas learn how to navigate from a desire for a STEM career to achieving one.

Keywords: STEM · Science · Technology · Engineering · Precollege studies · Mentoring

1 Introduction

"STEPS to STEM" is a strategy and program to guide students and prepare them for STEM future. The acronym "STEPS" indicates the phrase, "Science, Technology, Engineering, Precollege Studies". The title infers that involvement in STEM fields must begin in a timeframe prior to college. It must begin in the precollege stages, specifically elementary and/or junior high or middle school.

Elementary grades provide the first formal exposures to STEM. In elementary school in many instances, an uninhibited acceptance to STEM occurs because students have not yet learned the stereotypes and phobias associated with STEM courses. Research indicates that phobias about STEM courses, specifically mathematics begins in elementary school (www.rethink-anxiety-disorders.com). Mathematics filters for STEM majors and careers; so, it is warranted that a view of Mathematic phobias occurs.

Mathematical phobias are a type of fear. It is a fear triggered, and develops based on events that occur. These events lead to anxiety. In other words, mathematics phobias are learned behavior. The phobia results in anxiety brought on by the event of performing mathematics problems. Phobias are harmful because mathematics is believed

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 866–872, 2019.
https://doi.org/10.1007/978-3-030-22868-2_59

fridaya@uef.fi

to assist in the development of critical thinking skills. Thus, mathematics is a key subject for those pursuing STEM majors and careers.

Mathematics not only assists in the development of critical thinking; it is also a filter for STEM. Piaget terms the preoperational stage, ages 2–7, and the concrete stage, ages 7–11 are helpful in progressing students to the formal stage or critical thinking stage. This progression of exposures with a STEM perspective in mind is helpful. This exposure can begin with the taxonomy of educational objectives. These objectives were credited to Benjamin Bloom in 1956. The objectives were revised over the years. We find that Bloom's taxonomy was organized into the Cognitive, Affective, and Psychomotor domains.

The Cognitive domain is a hierarchy structure for thinking and learning. The hierarchy records a student's movement from non-critical to critical thinking. The hierarchy includes from the bottom to the top: Knowledge, Comprehension, Application, and Analysis, Synthesis, and Evaluation. The first three structures for thinking and learning represent lower levels of cognition and the top three represents higher-order or critical thinking levels. Students must be guided to move through these levels of cognition. This can be performed using connected learning approaches to STEM.

The Affective domain is a hierarchy structure for dealing with things emotional, to include feelings, values, appreciation, motivations, attitudes, and enthusiasms. Affect is how beliefs, feelings, and philosophies are demonstrated in observable behavior. Thus, it is apparent that the affective domain gives a view of phobias and stereotypes that form and affect students' predispositions to STEM experiences, majors, and careers.

The Psychomotor domain is a hierarchy of motor-sensory exposures and experiences. These can be provided as guided-reception exercises in which a teacher or STEM practitioner demonstrates a STEM hands-on activity to students. It can also be provided as an inquiry-discovery exercise in which a student takes the lead and uses his or her knowledge to creatively develop and design a STEM hands-on project. This provides students with a precursor to scientific research, and technology, engineering, and mathematics experiences.

Each of Bloom's domains taps into how students learn. Each has an opportunity to enrich a visual, auditory, kinesthetic/tactile, or combination learner. Whether students learn through seeing, listening or doing, each approach to learning can take advantage of STEM exposures. The more the domain interaction and learning styles connect to the real-world or everyday life experiences, the more relevance the STEM experience, major or career will assume in the life of the student. This opens up a discussion of "relevancy" of mathematics as it applies to STEM.

Students shy away from or reject mathematics because they are unable to make the connection between mathematics, especially higher-level mathematics, to real life. In many instances mathematics teachers are asked "Why are we learning this?" and "When will I ever use it outside of this class?" The teacher in some instances states, "You will need it to get into college; you will need it to move forward to the next required math course; or the material we are covering will be on the test". Each of these responses by teachers makes no connection to math's importance in real life situations.

In his essay "A Mathematician's Lament" Paul Lockhart says, "I don't see how it's doing society any good to have its members walking around with vague memories of algebraic formulas and geometric diagrams, and clear memories of hating them" (www.

fridaya@uef.fi

edutopia.org). Paul Lockhart's comment highlights the lack of cognition, affect and kinesthetic behavioral engagement. This lack helps to perpetuate a poor view of the importance of mathematics because students can't or have difficulty seeing the relationship between mathematics and real life.

Flegg et al. (2012) indicate even when mathematics is directly required as part of a major for any area of STEM, students question its use, especially when the teacher or professor does not make a connection to the use of mathematics in a specific industry context. Therefore, math's connection to real life must be pointed out in practical ways or by practitioners such that direct application of mathematics to everyday life and current events is obvious. Research indicates that this is true globally and not just in the United States.

Pronin (2014) study "Perceptions of the Relevance of Mathematics and Science: An Australian Study" highlights perceptions and attitudes of girls regarding the relevance of mathematics and science based on their gender. The Jones and Young (1995) study indicates that girls shut themselves out of STEM studies because of stereotypes about girls' ability to perform mathematics and their perception of mathematics relevancy unless it was connected with people, living things, or social issues.

More recently, statements and methods of motivating girls to debunk the stereotypes and myths of mathematics performance rather than motivate simply set the standard of performance to a gender proposition. Research by Markman and Chestnut (2018) indicates, "Although 'Girls are as good as boys at math' explicitly expresses equality, we predict it could nevertheless suggest that boys have rawer talent. In statements with this subject complement structure, the item in the complement position serves as the reference point and is thus considered more typical and prominent."

Kogelman and Warren (1978) Mathematics stereotypes include the following:

- Some people have a math mind and some do not
- There is a magic key to doing math
- Men are better in math than women are
- Math requires logic not intuition
- You must always know how you got the answer
- Math is not creative
- There is a best way to do a math problem
- It is always important to get the answer exactly right
- It is bad to count on your fingers
- Mathematicians do problems quickly in their heads
- Math requires a good memory
- Math is done by working intensely until the problem is solved.

It is evident from some stereotypes that cognition, affect, and psychomotor skills are involved in mathematics stereotypes. Therefore, when mathematics' relevancy is combined with gender stereotypes, it takes on an even greater challenge.

fridaya@uef.fi

2 Strategies for STEM

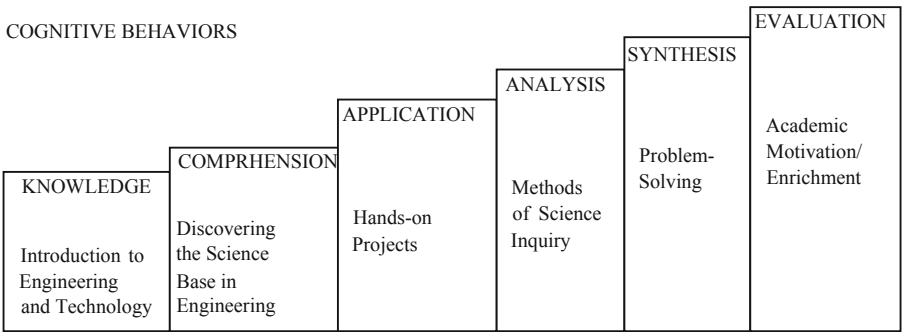
Foley et al. (2017), state, "Demand for science, technology, engineering, and mathematics (STEM) professionals is on the rise worldwide. To effectively meet this demand, many governments and private organizations have revamped STEM education and promoted training to enhance math and science skills among students and workers. Education and training programs typically focus on increasing individuals' math and science knowledge. However, data from laboratory studies and large-scale international assessments suggest that fear or apprehension about math, math anxiety, should also be considered when trying to increase math achievement and, in turn, STEM career success."

The Pearson (1988) STEM strategies I have used since the 1980's and continue to use today is Bloom's "Taxonomy of Educational Objectives" in combination with learning styles. The following graphics provide visual representations of the STEPS approach to lead to a STEM future for students.

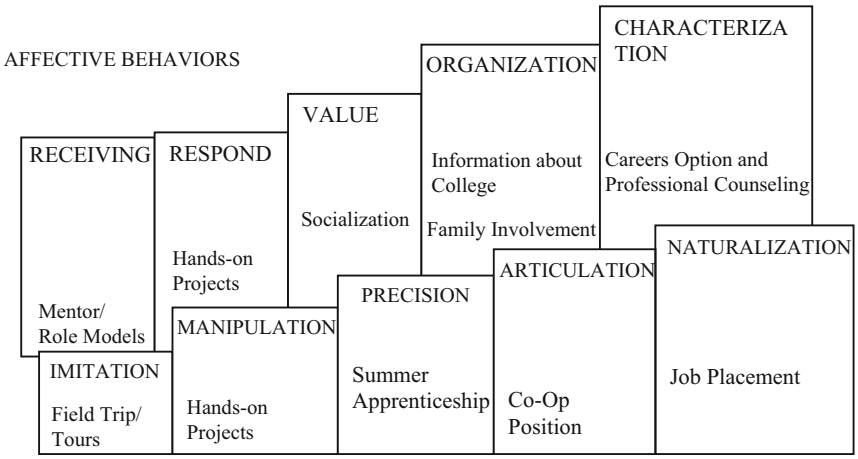
The STEPS approach is used as a framework for an action approach to increase the participation of underrepresented groups (Latino, Blacks, Native Americans, and Women) in science, technology, engineering, and mathematics related fields. Students became active participants in real-life science, mathematics, and engineering hands-on project experiences while interfacing with positive minority and women role models.

The students gained self-confidence, interest, and an awareness of STEM studies, college majors and careers. A visual of the STEPS strategy is shown below as it uses Bloom's Taxonomy of Educational Objectives in combination with STEPS program and project activities.

2.1 Pearson (1990) STEPS Program Model (with Bloom's Taxonomy of Educational Objectives)



fridaya@uef.fi



2.2 STEPS Project/Program for STEM Majors and Career Preparation

The STEPS projects and programs address:

- Students
- Courses
- Format of volunteer recruitment, student recruitment, funding
- Program evaluation

Students are recruited with an emphasis on early intervention. Evidence supports that an essential ingredient in the development of future STEM practitioners is exposure to hands-on experience and opportunities for exploration in science and mathematics. Recent research by Clements and Julie (2011) indicates exposures should occur as early as ages 3–5. Interventions designed to facilitate mathematical learning during ages 3 to 5 years have a strong positive effect on these children's lives for many years thereafter. Early childhood teachers often believe they are "doing mathematics" when they provide puzzles, blocks, and songs. Even when they teach mathematics, that content is usually not the focus, but is embedded in a fine-motor or reading activity. Cross et al. (2009) and Clements and Sarama (2009) indicate that evidence suggests that such an approach is ineffective, owing to a lack of explicit attention to mathematical concepts and procedures along with a lack of intentionality to engage in mathematical practices. The STEPS approach engages students ages 5 to 18 in cognitive, affective and psychomotor skills. Thus, it goes beyond doing math to actually experiencing it by engaging mind, emotions, and body.

Courses include:

- Computer Design
- Science Fair Projects
- Chemistry Measurement

fridaya@uef.fi

- Discovering Energy
- Physics in Electronic Design
- Engineering Principles
- Exploring Engineering Design
- Inventions, Patents, and Copyrights
- Engineering Entrepreneurs
- Radio Communications I, II, III
- Space Mission
- Fun with Graphics
- Making it in Math
- Robotics
- Technical Writing
- And Others

The format of volunteer recruitment, student recruitment, and funding are achieved by various methods. Volunteer recruitment for the STEPS program is obtained by contacting colleges, professional organizations and/or collaborating with companies. Colleges provide students seeking academic credit for community building projects. Professionals seeking to give back to the community and companies seeking to increase diversity in STEM by working with minority and/or women STEM organizations get involved.

Students are recruited by partnerships with school systems, individual schools, and youth organizations. The STEPS program is viewed as academic enrichment and supplemental for student off-hours STEM programing. Parental permission and engagement is obtained and parents are required to participate in a non-academic capacity (field trip monitors, meal preparation/serving, etc.).

Funding is obtained by grant writing to philanthropic organizations and corporations. Grant writing is critical for resource and equipment purchases, student field trips, and meals. A grant writer is obtained as a volunteer.

STEPS program evaluation is completed by administration of a pretest and posttest to students. These tests are designed to measure affective characteristics of the students. Secondly, the program conception and implementation is assessed to determine if improvements should be made to increase effectiveness and efficiency of the STEPS program. Students, parents, and volunteers are administered surveys to determine areas where improvements are needed.

Lastly, a formal "Certification of STEM Programs", specifically those with engineering focus was developed. The certification established criteria upon which program status, quality, and validity could be measured. The status criteria were intended to collect statistical data; quality and validity was meant to determine program effectiveness and fundability.

Future of STEM Careers

Historically, the United States circa 1960's–1970's, during the Sputnik era, was reasonably secure in the fields of STEM. However, the 1980's found the nation at risk of other countries, specifically the Pacific Rim, emerging as leaders. Today, STEM innovation is no longer the providence of the United States. The nation is at jeopardy of losing its' preeminence in commerce, industry, science and technology.

fridaya@uef.fi

Today, The U.S. Bureau of Labor Statistics (2015) indicate, "Seven out of the ten largest STEM occupations were computer related. Most of the largest STEM occupations were related to computers and information systems. There were nearly 8.6 million STEM jobs in May 2015, representing 6.2% of U.S. employment. Computer occupations made up nearly 45% of STEM employment, and engineers made up an additional 19%. Mathematical science occupations and architects, surveyors, and cartographers combined made up less than 4% of STEM employment.

The reason for the gradual eroding of STEM strength has many possible causes. However, it also has many potential solutions. The STEPS approach is one of the solutions because it provides early intervention, informal education opportunities, hands-on experiences, increased opportunities for women and minorities. Using cognitive, affective and psychomotor strategies to engage students in STEM connected learning provides a path to success as students learn how to navigate from a desire for a STEM career to achieving one.

References

Clements, D.H., Sarama, J.: Learning and Teaching Early Math: The Learning Trajectories Approach. Routledge, New York (2009)

Clements, D., Julie, S.: Early childhood mathematics intervention. AAAS/Sci. **333**(6045), 968–970 (2011). https://doi.org/10.1126/science.1204537

Cross, C.T., Woods, T.A., Schweingruber, H.: Mathematics in Early Childhood: Learning Paths Toward Excellence and Equity. National Academy Press, Washington, DC (2009)

edutopia.org. http://www.edutopia.org/blog/mathematics-real-world-curriculum-david-wees

Flegg, J., Mallet, D., Lupton, M.: Students' perceptions of the relevance of mathematics in engineering. Int. J. Math. Educ. Sci. Technol. **43**(6), 717–732 (2012)

Foley, A.E., Herts, J.B., Borgonovi, F., Guerriero, S., Levine, S.C., Beilock, S.L.: The math anxiety-performance link: a global phenomenon. Curr. Dir. Psychol. Sci. **26**, 52–58 (2017). http://journals.sagepub.com/doi/abs/10.1177/0963721416672463

Jones, J., Young, D.J.: Perceptions of the relevance of mathematics and science: an Australian study. Res. Sci. Educ. **25**(1), 3–18 (1995)

Kogelman, S., Warren, J.: Mind Over Math. McGraw-Hill, New York (1978)

Markman, E.M., Chestnut, E.K.: Cognitive science: "girls are as good as boys at math" implies that boys are probably better: a study of expressions of gender equality. Cogn. Sci. **42**(7), 2229–2249 (2018). https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12637

Pearson, E.: Pre-engineering program model: science technology engineering precollege studies (STEPS) (1988). Copyright1988

Pearson, E.: Science, technology, and engineering precollege studies (STEPS): a conceptual and theoretical framework (1990). Copyright1990

Pronin, E.: Women and Mathematics: Stereotypes, Identity, and Achievement. Princeton University, Princeton (2014). http://apcentral.collegeboard.com/apc/members/homepage/31986.html

rethink-anxiety-disorders.com. http://www.rethink-anxiety-disorders.com/math-phobia.html

U.S. Bureau of Labor Statistics (2015). https://www.bls.gov/spotlight/2017/science-technology-engineering-and-mathematics-stem-occupations-past-present-and-future/pdf/science-technology-engineering-and-mathematics-stem-occupations-past-present-and-future.pdf

fridaya@uef.fi

Named Entity Enrichment Based on Subject-Object Anaphora Resolution

Mary Ting[1(✉)], Rabiah Abdul Kadir[2(✉)], Azreen Azman[1], Tengku Mohd Tengku Sembok[3], and Fatimah Ahmad[3]

[1] Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 Serdang, SGR, Malaysia
gigimitty.mt@gmail.com
[2] Institute of Visual Informatic, Universiti Kebangsaan (UKM), 43600 Bangi, SGR, Malaysia
rabiahivi@ukm.edu.my
[3] Faculty of Defence Science and Technology, Department of Computer Science, National Defence University of Malaysia, Kem Sungai Besi, 57000 Kuala Lumpur, Malaysia

Abstract. Named Entity Recognition (NER) is an early stage processing of an Information Extraction, which identifies and classify entities in text. The outcomes of the task have become the foundation of building complex Information Extraction applications. With the enormous amount of information available everywhere, the area has gained lots of attention from the research community. Currently, there are two main approaches to perform the Named Entity Recognition; rule-based approach and machine learning approach. In order to improve the accuracy of the classification and performance of the recognizer, some researchers have implemented a hybrid approach, which is the combination of both approaches. Even though many research and works have been done in the Named Entity Recognition, there is still room available for improvement. This paper proposed to increase the accuracy of entities detected by implementing anaphora resolution during the preprocessing phase and a hybrid approach to classify the detected tokens during the classification phase. The hybrid approach is combined the Conditional Random Field (CRF) classifier with a gazetteer and pattern rules to perform classification. The result has shown that the application of anaphora and gazetteer has increased 46% accuracy of the detected entities for the person class.

Keywords: CRF classifier · Named Entity Recognition · Anaphora resolution · Rule-based extraction · Gazetteer

1 Introduction

In the current information-driven world, the amount of information available on the Internet promotes the needs to extract relevant information from resources rather than reading every single document. Even though information required by different systems varies according to its needs, normally these information are the entities reside in a context. The term "entity" refers to an identifiable object in natural language text,

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 873–884, 2019.
https://doi.org/10.1007/978-3-030-22868-2_60

fridaya@uef.fi

which is known as a subject or object of a clause. In language, a clause is a group of words that can express a complete proposition. A typical clause contains of a subject and a predicate. There are several Named Entity Recognition (NER) techniques have been invented to identify entities in a context. Invented techniques can be categories into rule-based, machine learning and a hybrid of rule-based and machine learning technique. These techniques focus mainly on the proper noun of a person, organization, and location.

Named Entity Recognition (NER) is an early stage processing of Information Extraction, which is used to identify and categorise entities in text into predefined categories. The recognition process plays an important role in information extraction, where the outcomes of the process become the foundation to build a complex Information Extraction System (IES). As of now, all the available named entity recognizer tools are able to identify general entities such as a person, organization, location, medical codes, time expressions, quantities, monetary values, percentages, etc. Some of the well-known NER tools have the abilities to identify more specific entities such as date, time expression, percentage and monetary. Other than the well-known tools, there are some custom NER tools that have been designed and implemented for a specific domain. These NER will be able to identify more specific entities such as medical, biomedical and chemistry terms.

The most common challenges faced in NER are the entity categorization, varieties and ambiguity for each entity name, writing style and format of the context. Although many researchers have invented many Named Entity Recognition techniques, solve these issues, there still a room for improvement. The list of recognised entities from a passage might not be accurate due to ambiguity and the writing style of the word, and the same word falls into various classes. An example can be seen in Fig. 1, where the recognizer categories the *two* as language, and *weeks* as date which is inappropriate. The pronouns such as *his, him* and *he* in the passage are referring to the Thomas Edison, which is an object/entity. However, the recognizer does not have the ability to recognise those words as a person entity.

[PERSON Thomas Alva Edison] lit up the world with his invention of the electric light. Without him, the world might still be a dark place. However, the electric light was not his only invention. He also invented the phonograph, the motion picture camera, and over [CARDINAL 1,200] other things. About every [LANGUAGE two] [DATE weeks] [PERSON Edison] created something new.

Fig. 1. Result of an entity recognizer tool

This paper focuses on solving the entity categorization and increase the accuracy of entities detected. Other than focusing on the existing issue, this research also includes a pronoun coreference to increase the number of entities identified in the context by implementing Subject-Object (SO) Anaphora Resolution.

fridaya@uef.fi

2 Named Entity Recognition

"Named entity" task was developed by Grishma and Sundhei in the sixth series of Message Understanding Conferences (MUC) to identifying the names of all the people, organisation and geographic location in a text [1]. This task has widely implemented in the NLP applications, especially in Information Extraction and Question Answering System. The task of identifying and extracting the names from the text at the early stage of processing is important to ensure the accuracy of an Information Extraction System. Named Entity Recognition (NER) is an important process in the Information Extraction field. It is a process used to identify entities in the text and associate them with its' respective class by placing a label beside them.

NER process will be executed at the earlier stage of Information Extraction System to identify entities, which are available in the text and group it in predefined class. Predefined class is grouping a list of entities, which has similar behaviours into a single category. Basic predefined classes are the person, company/organisation, location and MISC [2]. In order to extract more entities from the text, the number of predefined classes has increased over the years. Currently, NERs has the ability to recognise number like date, time, percent and money [3]. Other than a number, recognisable entities have also been extended to law, product, work of art, etc. [4]. For an example, a passage in Thomas Edison has been fed into Illinois extended Named Entity Recognizer (Cognitive Computation Group), an NER tool that has the ability to recognise 18 different entity classes. The outcome of the Illinois NER can be seen in Fig. 1, where each recognised entity is annotated with a category name. A colour is designated to each identified category, thus other than attaching a label with identified entity; each entity is highlighted with different colours in accordance with the class. For example, a red colour to represent the person class, violet colour for number class, peach colour for the language class, cyan colour for date class, etc.

The NER process generally starts with extracting content from the document. These content needs to be preprocessed before it can be used to identify the entity for the content and categorized it. Figure 2 shows the general flow of the Named Entity Recognition process.

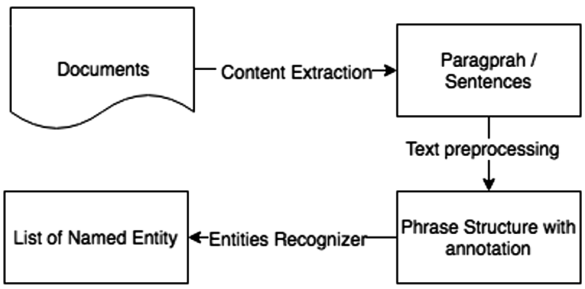


Fig. 2. NER process flow

fridaya@uef.fi

During the content extraction process, all the irrelevant information such as pictures, diagrams, tables, etc. will not be extracted from the document. Relevant content from documents will be further processed by breaking down the sentences/paragraph into individual tokens before it can be further processed. These tokens will be grouped together into different classes with a part of speech tag in accordance with their syntactic behaviour. Annotated tokens will be used to form a syntactic structure before it can be processed with a sequential labelling algorithm to detect an entity. The process will segments and labels, multi-token sequences based on grammar rule. Each segment will be further evaluated to identify the embedded entity.

This research has a process flows as all the NER tools except the techniques implemented to classify the entities during the entity recognition process. Normally, techniques used to identify and classify entities can be grouped into 3 different categories; rule-based, machine learning and hybrid of the rule-based and machine learning [5]. Rule-based is a traditional technique used to detect and identify entities in the text. To further enhance the performance of NER, machine-learning technique has been implemented to detect the entities according to the word's pattern. However, in this research, the hybrid approach is used and combined the Conditional Random Field (CRF) classifier with a gazetteer and pattern rules to detect entities during the recognition process.

3 Related Works

Various techniques have been invented to identify and classify the entities embedded in the text. The process of entity detection and recognition can be implemented using different techniques. These techniques can be grouped into three main categories, which are linguistic or rule-based, machine learning or hybrid of rule-based and machine learning techniques.

The rule-based technique is a classic approach used to identify and classify entities through a set of rules. These rules are a set of patterns, which is based on the linguistic grammar through some natural language processing techniques such as part-of-speech tagging, syntactic analysis and orthographic feature detection. Figure 3 illustrates an example of a set of rules defined to match the features of each class. Each rule will have a set of features corresponding to a specific class, which is based on a linguistic grammar rule. For example, the token must start with capital or entire token is capital, symbol in the token. Examples of the word token start with a capital letter are Malaysia, Putrajaya, Singapore, etc. As for all-capital letter token are IBM, UN, UPM, UK, etc. The symbol in the token is abc@gmail.com, ©Microsoft, $100.00, etc. If an annotated token matches any of the stated rules, then the token will be classified in accordance with a correspondence class for that rule.

Rule-based technique implemented to recognize entities can be found in Greek Information Extraction System [6], a NER tool for Malay language by [7], ANNIE Information Extraction System [8] and earlier NER tools such as NLTK, RegexNER (Stanford NER), LingPipe, OpenNLP.

In order to improve the accuracy of the detected entities using a rule-based technique, the machine learning detection technique has been implemented to detect entities

fridaya@uef.fi

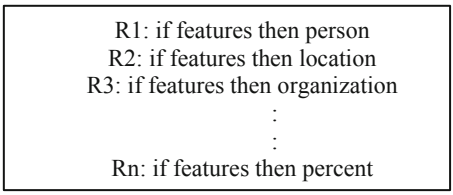


Fig. 3. A set of rules for rule based detection approach

with a statistical model. The technique uses a machine-learning algorithm with a statistical model to generate a model through the identification of the pattern and its correlation in the context. The applied technique can either be supervised, semi-supervised or unsupervised learning. Application of machine learning technique in NER can automatically induce rule-based or any sequence labelling algorithm from the training data.

Clustering features, a widely used unsupervised learning technique in NER to group similar features entities together. The technique relies on a set of patterns that is based on the word lexicon or word features. Authors in [9] applied the existing Brown clusters and word embedding technique to improve the accuracy of NLP systems. They induce 1000 Brown clusters supersets as word features and embeds with 100 dimensions over 5-g windows. Another word clustering work can be seen in [10], where Jensen-Shannon divergence is used to measure the similarity between probability distribution.

Instead of using a clustering method to detect entities, many named entity tools have implemented Conditional Random Field (CRF), a semi-supervised learning algorithm for labelling and segmenting structured data. Stanford NER is a Java based CRF classifier, which implements a linear chain Conditional Random Field sequence model [11]. The Stanford classifier labels the sequence of words in the text according to a set of predefined class like a person, location, time, location, organisation, money, date and percent. In [12], the author has implemented a bootstrapping approach to NER to identify and classify the entities based on word and context features given a set of training data.

Supervised learning technique requires a large set of the annotated corpus to construct a statistical model. The annotated corpus consists of a list of entities, where each entity contains several features, which is corresponding to a particular class. According to [5], most of the current NERs implements the supervised machine learning as a way to automatically induce rule-based systems or sequence labelling algorithms. Maximum Entropy Models, Support Vector Machines and Hidden-Markov Models are supervised learning techniques, which have been widely implemented to perform classification.

Hidden Markov model (HMM) is a generative model that assigns the joint probability of paired observation and label sequence. In [13], the author implemented HMM to directly generate the original NE tags from the output words of the noisy channel. Their model assumes mutual information independence, which is much looser instead of conditional probability independence.

fridaya@uef.fi

Maximum Entropy (ME) is a statistical modelling, which estimates probability distributions based on the principle of making as few assumptions as possible. The distribution should be as uniform as possible when nothing is known. Whereas [14] implemented ME model using generalised Iterative Scaling Algorithm. Another application of ME can be seen at [15] work, where they implemented two systems; one is fully based on training data and another with additional features derived from name lists. A well known Named Finder provided by OpenNLP package used pre-trained name finder model, which is based on the principle of Maximum Entropy to detect entities embedded in the text. The maximum entropy classifier belongs to the class of exponential model for feature classification [16].

Support Vector Machine (SVM) is based on the Structural Risk Minimization principle, which is very useful for classification [17, 18]. Mickelin [18] applied SVMlin to recognised Swedish named entities through a set of pattern in the training corpora. While Takeuchi et al. [17] implemented SVM using Tiny SVM package provided by NAIST, where each training pattern based on lexical feature will be defined as a vector.

Among the above techniques, machine learning approach has provided more accurate result as it's able to handle a large amount of data being processed and ability to predict an entity class through a set a predefined pattern. A further improvement has been made on the existing approach by integrating rules based and machine learning techniques together. A hybrid between rules based and Maximum Entropy Models or Hidden-Markov Models have been implemented previously [19–21]. Other hybrid approaches were integrated decision list learning to chunking technique [22] and integrated fuzzy algorithm into Support Vector Machine [23]. Further work was LingPipe Named Entity Recognition tool by Alias-i and Illinois by Cognitive Computation Group. LingPipe NER integrated n-gram character language models, Hidden Markov Models and CRF methods to perform expression matching [24]. Illinois extended NER integrated Hidden Markov Models, multilayered Neural Network, text chunking, gazetteer and Learning Based Java (LBJ) to train using OntoNotes corpus, which provides larger set of entity type.

4 Subject-Object Anaphora Named Entity Recognition

This research proposed a method to extend the existing NER process by adding an anaphora resolution module as a pre-processing of recognition process. The proposed method integrates version 3.6.0 CRF classifier API provided by The Stanford Natural Language Processing Group [25]. Hence, the proposed model only has the capabilities to recognise seven different classes based on the existing CRF model. These predefined classes are Person, Location, Organization, Monetary, Percentage, Date and Time expression. Instead of just depending on the CRF model provided by Stanford, the WordNet dictionary [26] and a gazetteer will be used to enhance the existing knowledge available.

fridaya@uef.fi

4.1 Anaphora Module

Anaphora is a technique of using a word, which refers to or replaces another word in a sentence. For example, the *his* in the sentence "Thomas Alva Edison lit up the world with *his* invention of the electric light" is referring to Thomas Alva Edison. Figure 4 illustrates the algorithm use to determine the antecedent of a pronoun for anaphora resolution. The module will search for the pronoun's antecedent through the existing entity's collection given as the token lexical information. The anaphora module will replace the identified pronoun such as *he, it, they, her,* etc. with its antecedents. While, if the pronoun appears in the first sentence of a paragraph, the proposed algorithm will look for the antecedent from the sentence instead of empty entities collection. Each pronoun word identified will determine it will act as the subject or object of a clause.

```
Find antecendent
 En_type ←SubObjVerification(pronoun)
 IF entities collection is nil then
     return SearchCurrentSen(En_type)
 ELSE
     index = EntityCollection size - 1
     repeat
         IF Entity_type in EntityCol(index) then
             return EntityCol(index).get(En_type)
         ENDIF
         index reduced by 1
     while index >= 0
 ENDIF
 return pronoun
END
```

Fig. 4. Anaphora algorithm

5 Subject-Object Anaphora NER Framework

The NER is divided into two parts, where the first part deal with pre-processing of the raw documents feed into the system and the second part deal with the classification each token. When a document is fed into the system. The system will break the document down into sentences and pass it into the Stanford parser to perform tokenization, stemming and part-of-speech (POS) process. The system will search for pronoun token after POS process to determine the need to executing the anaphora resolution module.

The Proper noun has been used in the NER system to determine the named entity. Proper noun tokens can be used to classify a person, location and company, but not suitable for a time, date, percent and many classes. Thus, the system will examine all the noun tokens during the classification process in this system. In order to reduce the cost of analysing every single word, only proper noun and noun tokens will be further analysed with WordNet dictionary. WordNet dictionary has the capability to determine

fridaya@uef.fi

the token based on its lexical group. The dictionary has the capability to determine a word as a person, location, time, animal, abstract, body, etc. Since the system only required to identify certain classes, thus only token that fall under the lexical group person, location and time be labelled with its lexical group name. The outcome of the WordNet passes into the second part of the system.

The second part of the system dealing with the classification process, where the tokens received from pre-processing will be divided into two groups; labelled and unlabelled. The labelled token will be directly stored in an entity collection, whereas unlabelled tokens will be passed into the classifier to detect its class. The classification process will receive a list entities together with its class from a gazetteer. An entity list, pattern rules and Stanford CRF classifier are integrated to perform classification for the unlabelled token. Only token identified from the predefined classes will be updated into the entities collection. When the entire document has been processed, the system will generate a list of entities from the tokens stored in entities collection. Figure 5 illustrates the framework for the Subject-Object Anaphora NER process.

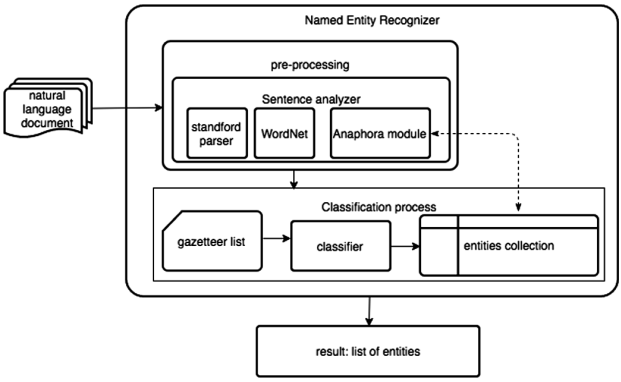


Fig. 5. Subject-Object Anaphora Named Entity Recognition framework

6 Experimental Result and Discussion

The proposed NER has been tested with Level 2 of the Remedia Publication dataset. Level 2 dataset consists of 28 children's stories in different documents with an average of 200 words per document. Entities presented from these 28 documents can be categories into six different named entity classes; person, organisation, location, time, date and money. Among these NER tools, such as ANNIE only has the ability to detect 3 classes that are a person, organisation and location. Whereas, Illinois has the ability to detect the highest number of classes such as number, artefact, law, nationality, quantity, work-of-art, event, language, facility, product and also the seven other classes identifiable by Stanford and proposed NERs.

Figure 6 shows the total number of accurate entities detected in 28 documents. Among the NERs, the proposed method yields the highest number of accurate entities due to the additional implementation of anaphora resolution in pre-processing phase,

fridaya@uef.fi

thus gazetteer and pattern rules in the classifier. In overall, ANNIE NER only produces 133 accurate entities, which is comparatively lower among other NERs. ANNIE NER actually yields a better result than the others for an individual class, especially for a person and location classes. This is due to the massive number of gazetteer lists reside in the NERs.

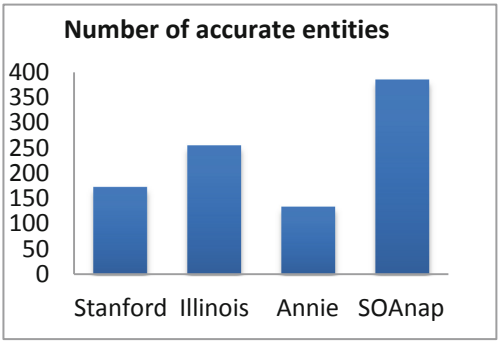


Fig. 6. Total number of accurate entities in Remedia level 2 dataset

Table 1 shows the result of accurate entities retrieved from Stanford, Illinois, ANNIE and the proposed NERs. From the result, it has clearly shown the significant improvement in identifying a person when anaphora resolution has been integrated into the pre-processing process. Thus, the use of the gazetteer and pattern rules has increased the accuracy of the organisation and location entities. Among the existing NERs, Illinois generally yield the highest number of accurate entities detected due to the number of predefined classes. Other than recognising a variable number of classes, Illinois also yields the highest amount of errors due the number resides in the context. The numerical digit in date expression and monetary categories might be considered at a cardinal number.

Table 1. NER accuracy for dataset of Remedia level 2

	Stanford	Illinois	Annie	SO Anaphora
Person	0.12	0.19	0.17	0.65
Organisation	0.23	0.59	0.36	0.68
Location	0.25	0.35	0.39	0.35
Time	0.37	0.20	0.00	0.32
Date	0.69	1.00	0.00	0.73
Money	0.18	0.00	0.00	0.18

Based on the result, none of the tested NERs has the ability to detect monetary entity category due to the writing style such as RM2.10. For example “millions of people lost some or all of their money in the stock market”, “some money” or “all

fridaya@uef.fi

money" should be recognised as monetary. Stanford and proposed NER recognise the "*two million dollars*" words as a monetary. However, Illinois NER recognised "*more than three million dollars to move*" as Date expression.

One of the weaknesses found among those NERs are proper noun token recognition, which usually will be recognised either as a person, location and organisation. In some situation, the proper noun is used to represent an object such animal, car and house instead of a person, location and organisation. "Shamu", "Kandu" and "Millie" found in document RM2.15 and RM2.29 is used to represent an animal. Yet, all the NERs recognised it as a person entity.

7 Conclusion and Future Work

The proposed method applied anaphora resolution technique during the pre-processing to transform the pronoun into a proper noun to increase the number of names entities category. Integration of the existing CRF model and WordNet dictionary with the gazetteer and pattern rules during the classification process has been improved the percentage of the accuracy entities detected. While, the existing NER technique does not have the ability to find all the named entities reside in a document due to by ignoring the pronouns in a sentence or paragraph. Without anaphora resolution, the pragmatics between the sentences that using pronouns would be missing in the process of entity detection.

In the future work, the research would apply the proposed approach to enrich the word sense to solve the ambiguities and semantic issue in Question Answering System and Information Extraction System. The broaden scope of answer and information extraction would be increased with higher F-measure in information retrieval.

Acknowledgment. This research has been supported by the Fundamental Research Grant Scheme funded by the Ministry of Higher Education (MoHE), Malaysia under Project Code FRGS/1/2017/ICT04/UKM/02/8 and Industrial Fund registered under Universiti Kebangsaan Malaysia Grant with Project Code ZG-2018-001.

References

1. Grishma, R., Sundhei, B.: Design of the MUC-6 evaluation. In: MUC6 '95 Proceedings of the 6th Conference on Message Understanding, pp. 1–11 (1995)
2. McCallum, A., Li, W.: Early results for named entity recognition with conditional random fields; feature induction and web-enhanced lexicons. In: Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL, vol. 4, pp. 188–191. Association for Computational Linguistics (2003)
3. Manning, C.D., Bauer, J., Finkel, J., Bethard, S.J.: The Stanford CoreNLP Natural Language Processing Toolkit. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55–60 (2014)
4. Redman, T., Sammons, M., Roth, D.: Illinois Named Entity Recognizer: Addendum to Ratinov and Roth'09 reporting improved results (2016)

fridaya@uef.fi

5. Nadeau, D., Sekine, S.: A survey of named entity recognition and classification. Lingvisticae Investigationes **30**(1), 3–26 (2007). https://doi.org/10.1075/li.30.1.03
6. Farmakiotou, D., Karkaletsis, V., Koutsias, J., Sigletos, G., Spyropoulos, C.D., Stamatopoulos, P.: Rule-based named entity recognition for Greek financial texts. In: Proceedings of the Workshop on Computational Lexicography and Multimedia Dictionaries (COMLEX 2000), pp. 75–78 (2000)
7. Alfred, R., Leong, L.C., On, C.K., Anthony, P.: Malay named entity recognition based on rule-based approach. Int. J. Mach. Learn. Comput. **4**(3) (2014). http://doi.org/10.7763/IJMLC.2014.V4.428
8. Named Entity Recognition with ANNIE. http://services.gate.ac.uk/annie/. Accessed 17 Sept 2015
9. Turian, J., Ratinov, L., Bengio, Y., Turian, J.: Word representations: a simple and general method for semi-supervised learning. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, pp. 384–394 (2010). http://doi.org/10.1.1.301.5840
10. Chrupała, G.: Hierarchical clustering of word class distributions. In: Proceedings of the NAACL-HLT Workshop on the Induction of Linguistic Structure, pp. 100–104. Association for Computational Linguistics (2012)
11. Stanford NLP Group. Stanford Named Entity Recognizer (NER). http://nlp.stanford.edu/software/CRF-NER.shtml. Accessed 12 Oct 2015
12. Thenmalar, S., Balaji, J., Geetha, T.V.: Semi-supervised Bootstrapping approach for Named Entity Recognition. ArXiv E-Prints (2015)
13. Zhou, G., Su, J.: Named entity recognition using an HMM-based chunk tagger. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics ACL 02, 6744, pp. 473–480 (2002). http://doi.org/10.3115/1073083.1073163
14. Borthwick, A.: A Maximum Entropy Approach to Named Entity Recognition (1999)
15. Chieu, H.L., Ng, H.T. Named entity recognition with a maximum entropy approach. In: Proceedings of the 19th International Conference on Computational Linguistics, vol. 4, pp. 160–163. http://doi.org/10.3115/1119176.1119199
16. OpenNLP. http://www.opennlp.sourceforge.net. Accessed 8 Aug 2015
17. Takeuchi, K., Collier, N.: Use of support vector machines in extended named entity recognition. In: Proceedings of the 6th Conference on Natural Language Learning-Volume 20. Association for Computational Linguistics, pp .1–7 (2002). http://dl.acm.org/citation.cfm?id=1118882
18. Mickelin, J.: Named Entity Recognition with Support Vector Machines Master of Science Thesis Named Entity Recognition with Support Vector Machines. Royal Institute of Technology, Stockholm (2013)
19. Ferreira, E.: Combining Rule-based and Statistical Methods for Named Entity Recognition in Portuguese. Language and Speech, pp. 1615–1624 (2007)
20. Srihari, R., Niu, C., Li, W.: A Hybrid Approach for Named Entity and Sub-Type Tagging. In: Applied Natural Language Processing Conference, pp. 247–254 (2000). http://doi.org/10.3115/974147.974181
21. Jayan, J.P., Sherly, E.: A Hybrid Statistical Approach for Named Entity Recognition for Malayalam Language, pp. 58–63 (2013)
22. Sassano, M.: Named Entity Chunking Techniques in Supervised Learning for Japanese Named Entity Recognition. Training, 3676, pp. 17–19
23. Mansouri, A., Affendey, L.S., Mamat, A.: Named Entity Recognition Approaches. J. Comput. Sci. **8**(2), 339–344 (2008). https://doi.org/10.1075/li.30.1.03nad
24. Named Entity Tutorial. http://alias-i.com/lingpipe/demos/tutorial/ne/read-me.html. Accessed 8 Dec 2015

fridaya@uef.fi

25. Finkel, J. R., Grenager, T., & Manning, C.: Incorporating non-local information into information extraction systems by gibbs sampling. In: ACL, pp. 363–370 (2005). http://doi.org/10.3115/1219840.1219885
26. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press, Cambridge (1998)

fridaya@uef.fi

Tour de Force: A Software Process Model for Academics

Zeeshan Haider Malik(✉), Habiba Farzand, Muhammad Ahmad, and Awais Ashraf

Forman Christian College (A Chartered University), Lahore, Pakistan
zeeshanmalik@fccollege.edu.pk

Abstract. Software engineering functions on divergent models. Models play the most significant role in the triumph of a project. Models can be considered as skeleton of the software. The development models are the innumerable processes and methodologies that are being tabbed for development of the project reliant upon the aims and goals of the project. A software process model is a standardized format for planning, running and organizing a development project. Out there exists many software process models for the industry but they fail to apply at academics. While students are in the process of becoming software engineers, it is essential to give them a taste of the professional world and this could be accomplished by facilitating them with a process model applicable for Final Year Project and by ensuring the constant support of Supervisor throughout the project. For fulfilling this need, a customized process model is designed and evaluated. The results of evaluation portray the successful achievement of the goal.

Keywords: Software engineering · Software process model · Tour de Force · Software process model in academics · Final year project

1 Introduction

Towards the final year of computer science baccalaureate program, students are asked to opt for a final year project. The Final Year Project is regarded as the crowning learning experience of the undergrad program. The measure of excellence of student productivity is used as a gauge for the quality of the program [1]. The final year project acts as a mechanism to introduce the students to many skills required for engineering graduates, especially, professional skills. This has become vital since the revised evaluation criteria for accreditation of engineering programs highlight the development of professional skills [2]. A majority of Final Year Projects are about developing software. At this point of time, students look for a software process model to be used in their projects.

Software development projects pass through a series of phases that account for their inception, inaugural development, productive operation, upkeep, and retirement from one age bracket to another. A software process model comprises of activities and

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 885–901, 2019.
https://doi.org/10.1007/978-3-030-22868-2_61

fridaya@uef.fi

related information that is required to develop a software system. The goal of a software process model is:

- A list of activities that need to be performed
- The input to and output from each task
- The preconditions and post conditions of each task
- The order of flow of tasks [3].

The failure of plenty of projects in 1960s led to rise of innovation and creativity in the field of software process models. In the following years, many process models like waterfall, scrum, RUP were introduced to the industry. Despite all these software process models, there remains a need for a software process model that meets the needs of the student in his final year project. Practical implementation of software process model is necessary in order to train the students about the professional world which they are about to enter. A student has 8 months of time altogether (2 semesters) and is expected to produce software that meets all the requirements and is bug-free with complete system and user documentation and that too, on time. This proves to be a big challenge for students. The available process models are applicable on large scale projects and demand more time and a larger team; which is why they fail to apply at academic level. Students face difficulty in managing and applying software process models due to which at certain times, they fail to accomplish all software related tasks and sometimes they are able to manage by adopting unfair means.

Along with all these factors, the lack of interest of Supervisor further heightens the difficulties for the students. The role of supervisor is so important that a number of researchers consider it more important than the role of a personal tutor [4].

A proper process model that is customized according to the student needs and requirements keeping in mind the scope of project and time availability in combination with the support of supervisor is missing. Hence, there arises a need for a software process model that not only meets the goals of a software process model but is designed keeping in mind the level and time available for a student to complete the software development project and with confirmation that the student receives maximum support from the Supervisor.

2 Literature Review

A research by Anand Kr. Shukla and Archana Saxena discusses various process models. Table 1 summarizes the findings of their research [5].

Out of the many process models, waterfall process model is the oldest and most widely used. This model constitutes of concept phase, requirements phase, design phase, implementation phase, test phase and installation/check out phase. These activities are carried out in order with little or no iteration [6]. This process model brings along some assumptions which includes that requirements are known upfront and are unambiguous and rarely change. In addition to this, users of the system know what they want and the system is not so complex. No working software is produced until the developer has reached the testing phase [7].

fridaya@uef.fi

Table 1. Comparative analysis of software process models

Features	Spiral model	Prototype model	Waterfall model	Iterative model
Requirement specification	Well understood	Not understood	Not understood	Well understood
Cost estimation	High cost	High cost	Difficult to estimate cost	Less costly
Risk analysis	Very high	Minimum risk	Risk not cover	Less risk
Software compatibility	For big projects	Complex	For big project	For big project
Complexity	Complex	Flexible	Complex	Less complex
Feasibility	Flexible	For both large and small project	Poor flexible	Flexible
User involvement	User involvement exists	High user involvement	No user involvement exists	User involvement exists
Success rate	High success rate	High success rate	No success rate	High success rate

This process model cannot be applied to the Final Year Projects as the project is not a small project and there are high chances of changing requirements. Students are in the learning and practicing phase, there is a high probability of iterating between difference phases. On the contrary, waterfall model does not allow iteration between phases. All skills learnt in the 4 year bachelors program are put into practice through this Final Year Project. When a student steps into the world of professionalism, Final Year Project is something which is most talked and enquired about. Keeping in consideration these factors, Final Year Project cannot be small scale and hence waterfall model cannot be applied.

Code and fix is another process model available. Coding and error fixing together with ad-hoc tests are the only activities performed consciously [8]. The constraint with this model is that system does not conform to the real requirements and cost for corrections are disproportionately high. Important concepts and decisions are not documented. When a student is in the learning stage, it is highly important and required that he documents every detail of the project.

Barry Boehm's reaction to the critics of waterfall model resulted in spiral process model. The main focus of this model is on risk analysis. This model has four phases; planning, risk analysis, engineering and evaluation. The project repeatedly passes through these phases [9]. This model eliminates the risk to a greater extend but still this model cannot be implemented on Final Year Projects as risk analysis require highly specific expertise which the students of the final year might not have. The success of the project is highly dependent upon risk analysis phase. Elimination of risks can result into high costs of the project which can be difficult for the students or university to manage.

fridaya@uef.fi

Incremental model presents another model for software development. Each module of the software passes through requirements, design, implementation and testing phase [10]. A working version of software is produced at the completion of each module. Progress is measured by executable systems. The problem with this model is that it requires good planning and design. This model can be used when requirements are well known and understood of the complete system. Students might not be able to draw requirements that well and complete as they are experimenting and testing their skills of software engineering. Furthermore, they might experience problems in breaking down components of the system and building incrementally and this can result into exceeding the time limit of completion of the project.

Another approach for software development is iterative. This development involves continuous revision of the software. A new and improved version of the software is produced in every iteration step [11]. Iterative development may produce costly system architecture or design issues may arise because not all requirements are collected upfront for the entire progress of the software. This will result into more time, cost and effort in terms of people which might be not possible for the students of Final Year Project.

Prototyping is hardware or software development tool in which preliminary version of part or all of the hardware/software is developed for user response and determining user feasibility [6]. Prototyping is best to use when the desired system has a lot of interaction with the user. Examples include online systems, websites, mobile apps and much more [12]. Prototyping can be used for some projects of Final Year which involve a lot of interaction between the system and the user but the other projects which do not involve much of interaction, cannot take much benefit from this model. Thus, there arises a need for a standardized model which can be applied on every sort of project.

Scrum Process Model offers allocation of roles i.e. product owner, scrum team and scrum master. Product owner represent all stakeholders. Scrum team is self-organized interdisciplinary team. The scrum master ensures that the scrum process is performed as planned and removes blockers and problems. The scrum process involves pre-game phase, development phase and post-game phase. Scrum proposes the idea that a team agrees on developing a group of requirements from the requirements backlog in a short period of time and then the cycle continues for the rest of the requirements [13]. The problem with scrum model is that it cannot be applied on every sort of project as experienced scrum master is mandatory and the students of Final Year have just started preparation to step into the practical world. Hence, the experience part is missing. With little or no experience, it is hard to take decisions which result into positive outcomes in terms of project success. On the other hand, scrum of scrum will be required to handle larger projects which can be a complex task for the newly educated software engineers to handle.

Rational Unified Process (RUP) model makes use of iterative and incremental approach. It is based on use cases and prototyping. The phases included are inception, elaboration, construction and transition [14]. RUP is large and time consuming. Having time of 8 months to completion and presenting a bug-free project with user and system documentation, it can be a big challenge for students to manage and cope with.

fridaya@uef.fi

Moreover, it could lead to over-documentation. It is difficult to customize according to the project as the process elements are inter-linked.

3 User Study

Survey was chosen as the technique to conduct user study. The purpose behind this was to find out the current situation regarding Final Year Project and process model selection. 109 respondents were chosen; 56.88% males and 43.12% females.

Computer Science is a practical field which asks for more and more practice. The more practice one gets during the academic life, the more experienced and well managed he is in the professional life. Figure 1 shows that 45.87% respondents have worked on more than 5 projects during their graduate program. This stat leads to the point that there has to be a process model which can be implemented and practiced well at academic level.

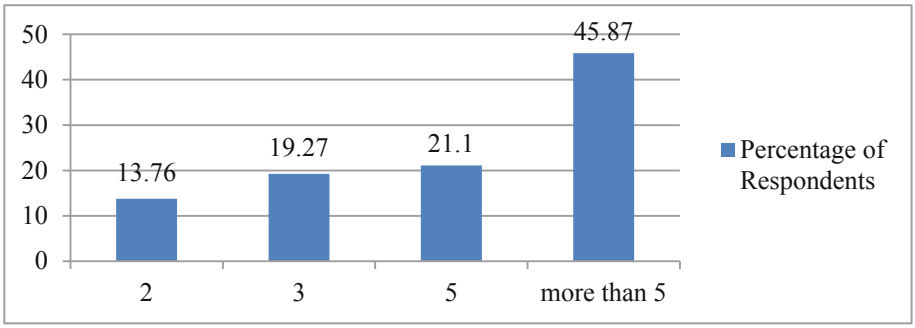


Fig. 1. Number of projects worked on during graduate program

Figure 2 shows the number of respondents who have used process models for their projects. 65.14% respondents have used process models which shows that students are eager to learn and implement the learnings; no matter if they are able to fully apply it or not. On the other side, 34.86% respondents did not use process models; mainly due to their drawbacks and constraints.

50–70% features have been implemented successfully by 55.96% respondents as shown in Fig. 3. Whereas 70–100% features were accomplished by only 36.70% respondents. 100% features implementation was achieved by 7.34% respondents only. This depicts that the students are unable to meet all the requirements of their proposed project. Thus, there arises a need for a way out in terms of process model which helps them in accomplishing all features successfully.

When asked about the effect of unaccomplished features on project success, 56.88% respondents (as shown in Fig. 4) said that the success of their project was affected badly. Back in 1960s, when projects failed tremendously, different process models for the industry were introduced. Now, when it comes to project failures at academic level, there has to be a process model which saves the projects from failure.

fridaya@uef.fi

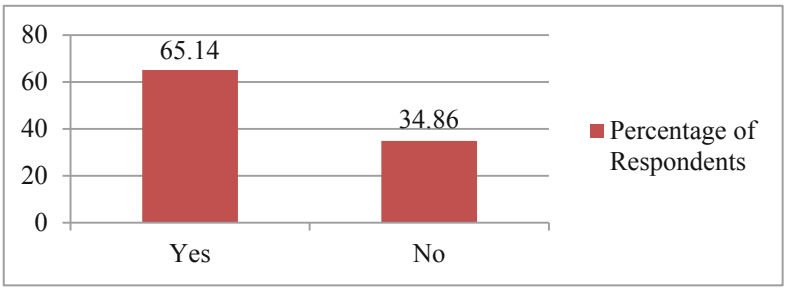


Fig. 2. Usage of process model for projects

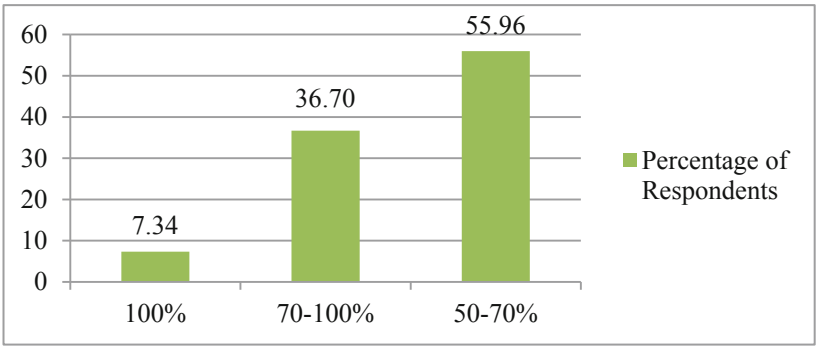


Fig. 3. Percentage of accomplishment of features

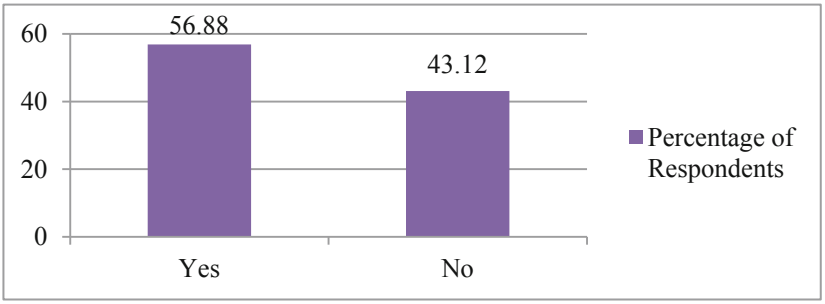


Fig. 4. Effect of unaccomplished features on project success

Students were enquired if they feel a need for a customized process model according to their level, scope and other constraints. Figure 5 shows that 88.99% respondents voted in the favor of a customized process model.

Figure 6 shows 66.06% respondents are using process model in their Final Year Projects. This indicates that there is an immense need to design a process model that suits their needs and is in accordance with their level. Students need a process model to

fridaya@uef.fi

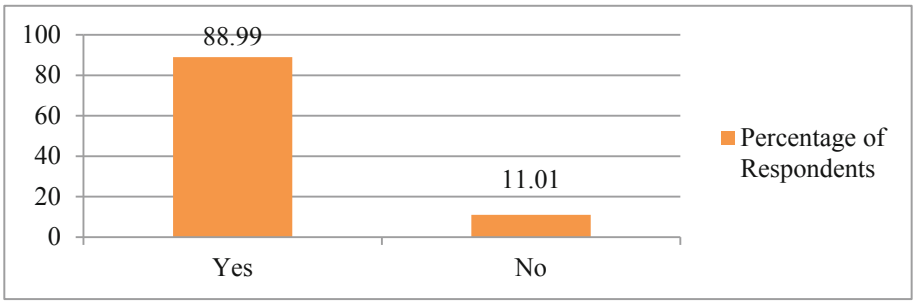


Fig. 5. Need of Software Process Model for Developing Software

shadow and complete their Final Year Project. The rest 33.94% participants are not implementing any process model mainly due to the reason that they are unable to implement it fully and achieve the desired results.

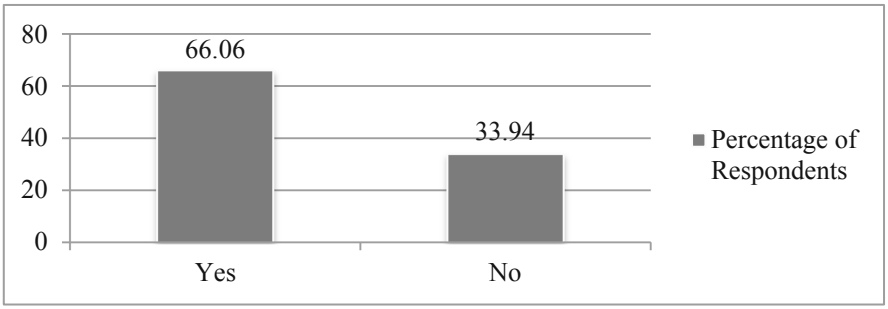


Fig. 6. Use of software process model in Final Year Projects

Figure 7 illustrates the choice of process model for Final Year Project. 38.53% respondents had no idea about what process model would suit their Final Year Project. On the other side 26.6% respondents chose waterfall model. Waterfall model does not allow iteration between phases which leads to incomplete features and a lot of time consumption. The existing process models are causing hindrance for students in accomplishing their goal of the Final Year Project.

When asked about if there exists a need for a customized process model, 72.48% respondents voted in favor of a new customized process model. Figure 8 illustrates the results.

85.32% respondents wish to use process model in Final Year Project as shown in Fig. 9. This shows students eagerly want to learn the application of process model. Thus, there has to be a model that suits their needs and prepare them for the professional life.

fridaya@uef.fi

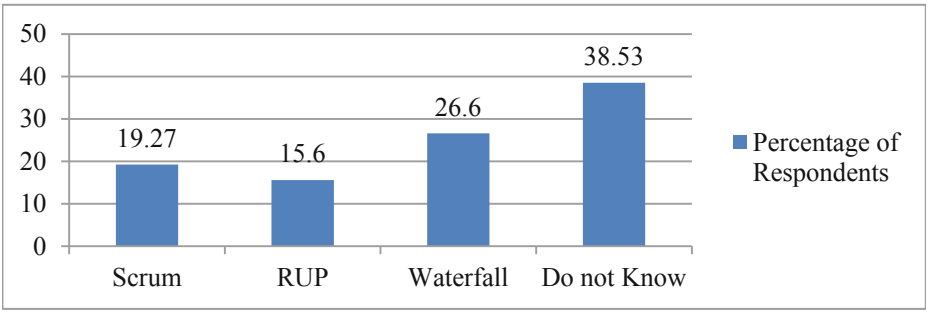


Fig. 7. Choice of process model for Final Year Project

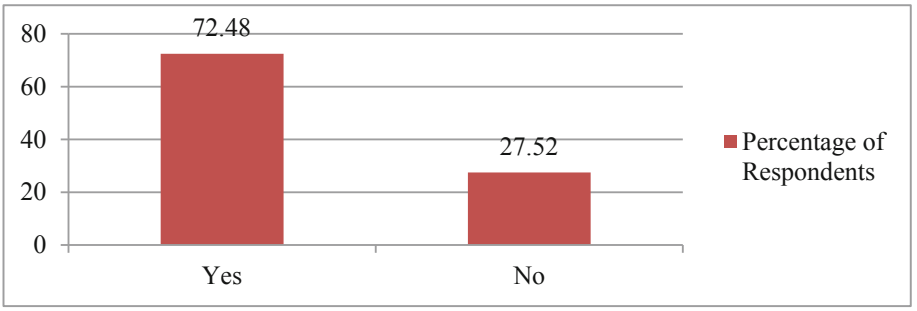


Fig. 8. Need for a customized process model

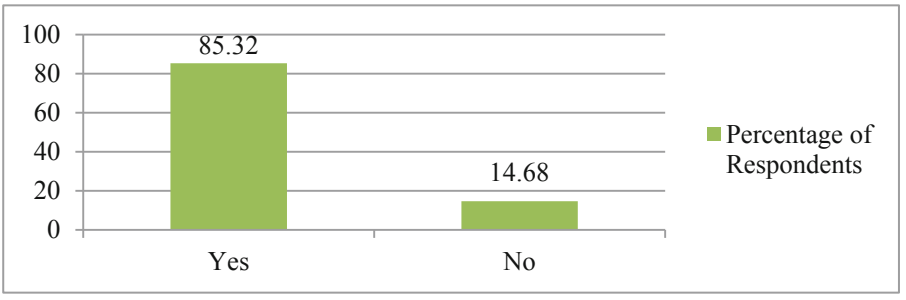


Fig. 9. Wish to use process model in Final Year Projects

Figure 10 shows the involvement of the advisor in Final Year Project. 86.24% respondents answered that there was no involvement of the advisor whereas only 13.76% said that there was an involvement of the advisor.

fridaya@uef.fi

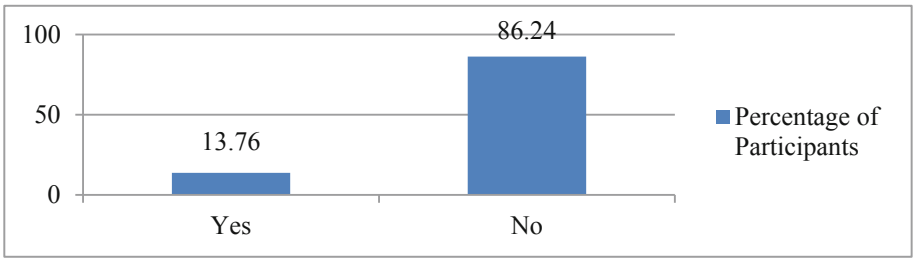


Fig. 10. Involvement of the advisor

4 Focus Group Session

How do people consider an experience or an experiment? What can be done easily by gathering people in groups and easing interaction between them? Focus Group is one of the qualitative research methods that provide answers to such queries. Focus Group was selected to discuss and explore the problems in depth faced by the students related to Final Year Project. Focus Group was chosen as it is easier to conduct and generates opportunities to collect data from group discussions. It has high face validity and low cost in relation to other methods [15]. The focus group session was held at the university campus and total 10 students were selected and out of which 7 were from Final Year and 3 of them were recent graduates. All participants were presented consent form prior to the study. The focus group session was digitally recorded and lasted for about 50–100 min. Field notes were also taken.

Three sessions of Focus Group Study were conducted. The results of Focus Group session clearly illustrated the problems faced which were as follows:

1. 3 out of 8 months allocated for Software Requirement Specification document. Less time for development of software.
2. Lack of interest of Advisor.
3. No follow ups from Advisor.
4. No systematic way of handling the project.
5. Coordination issues among group fellows.
6. Incomplete features/requirements due to lack of time.
7. Outsource help for the completion of project.
8. No practical experience of a software process model.

Keeping in consideration the results of Focus Group session, a customized process model was designed.

fridaya@uef.fi

5 Proposed Model

Tour de Force software development process model is designed for Final Year Projects. This model will assist the students in accomplishing their goals towards the final year project. This model will also shower light upon how process models are used in the practical world. This model comprises of the following attributes:

- Ace: This title will be allocated to the Advisor. Ace will play the role of Product Owner. Ace will guide the students about how to achieve desired results and will assist them in times of difficulty. He/she will also be responsible to keep a check on the project progress.
- Crew: This title will be granted to the students undertaking Final Year Project. The members of one crew can vary from two to four members but not more than four. Workload will be equally divided among students.
- Time Chart: This will be designed in accordance with workload and deadlines and in coordination with Advisor. Ace and crew will be responsible to keep a check for upcoming deadlines and goals to be accomplished within the time limit.
- Daily Meetup: In order to keep a regular check on the progress of project, crew will be organizing a small meeting daily and discuss the progress or shortcomings. The meeting minutes will be recorded and later presented to the Ace. This way Ace will also be in touch with the progress of the project and contribution by each crew member.
- Requirement Inventory: This will be the list of all significant goals to be accomplished along with time required and deadline to achieve it. In addition to this, it will be the replacement of Software Requirement Specification Document.
- Activity List: This will be the list of all features, functionalities, requirements and changes to be required to be completed by next meetup. A time chart indicating every task that has been completed along with estimated and actual time taken will also be included.
- Leap: Duration of 2 weeks will be called one leap. In one leap, a set of requirements from requirement inventory will be selected and placed in the activity list. The selected set of requirements will have to be completed until next leap meetup.
- Leap Meetup: For the maintenance and scheduling of Requirement Inventory and Activity List, bi-weekly meetings will be conducted by Ace with the crew. This will also serve the purpose of updating the project progress and future plan. This meetup will be known as “Leap Meetup”.

6 Evaluation

For the evaluation of the proposed model, a website was developed. This web portal was used as a test case for the proposed Tour de Force model. This was necessary in order to evaluate the proposed model before handing it over to students.

The web portal was related to imports, exports and vendor-ship of customs department. This portal had different NTN holder companies registered on it with their detailed profiles. Furthermore, it had a list of licensed clearing and forwarding agents

fridaya@uef.fi

with their complete profiles. The portal displayed and maintained updated lists of auctions. The facility of online registration was also available for the companies and agents to ease upthings. Furthermore, the users could also subscribe and get weekly updates about the latest activities that will take place on the portal.

The requirement inventory of this website included:

- A Home page
- Company description
- List of registered companies for import
- List of registered companies for export
- List of Custom Agents
- List of categories
- Contact Form
- Database for the form

Requirements were set in accordance with estimation (days) and priority. The estimation was set by considering the complexity of the user story. In one day team worked three hours a day per person. And all the requirements were developed with this map. Table 2 shows the estimated days required and the priority sequence.

Table 2. Estimated days and priority

User story	Estimation (days)	Priority
Home page	3	1
Company description	1	2
List of register custom agents	5	4
Category list	1	3
List of registered companies for imports	5	5
List of registered companies for exports	5	6
Contact us form	3	7
Database for the form	5	8

Week 1

Table 3 shows the work done in first week of the development. In the first week, prototype of the software was designed and was tested with users.

Week 2

Table 4, i.e. activity list shows the work done in week 2 of the development. User story 1 was taken and developed after the prototype evaluation. The home page with slider was developed. There are three different sliders and those slider shows the slogan and the company motive. These sliders can be moved right and left. The start now button will help to move on the main web and show other information.

Week 3

User story 2 and 3 were developed in this week. The page showed the description of the company when clicking of the start now button. Further the page for categories was

fridaya@uef.fi

Table 3. Work done in first week

Task	Work plan (meeting)	Estimated time	Actual time
T1: Paper Prototype	Day 1 Discussion	2 h	2 h, 18 min
T2: 2nd Iteration Prototype	Day 2 Design Discussion	2 h	1 h, 35 min
T3: 3rd Iteration	Day 3 Design Discussion	2 h	1 h, 20 min
T4: Final Prototype	Day 4 Design Discussion	2 h	2 h, 12 min
Evaluation of the Prototype	Day 5 Design Evaluation Discussion	2 h	1 h, 30 min
Ace Meeting	Next Leap Planning	1 h	1 h, 15 min

Table 4. Activity list

Task	Work plan (meeting)	Estimated time	Actual time
T1: Main Page	Day 1 What and How	2 h	2 h, 07 min
T2: Start Now Tab and Sliders	Day 2	3 h	2 h, 35 min
T3: Click any tab and next page (HTML/CSS) No data	Day 3	3 h	2 h, 20 min
T4: HTML/CSS Pages	Day 4	3 h	3 h, 12 min
Evaluation of the work	Day 5	2 h	1 h, 30 min
Ace Meeting	Next Leap Planning	1 h	1 h, 15 min

also developed with the list of categories that were being dealt on this website. All the sections of the web were defined on the navigation tab. Table 5 illustrates the activity table for week.

Week 4

This was the week in which the major user stories are developed. The data was collected about the companies and then it was entered to the pages of importers and exporters. The next page of agents was developed. The logo of the website was also designed in this week. Table 6 illustrates the activities of Week 4.

fridaya@uef.fi

Table 5. Activity table

Task	Work plan (meeting)	Estimated time	Actual time
T1: About Us Page with background image	Day 1 Company Description	2 h	1 h, 30 min
T2: HTML/CSS Categories Page with proper list	Day 2 get the list of categories	3 h	2 h, 45 min
T3: New tab, click any tab and next Page (HTML/CSS) No data	Day 3 Data Collection	3 h	3 h, 20 min
T4: HTML Page for Importers	Day 4 Data Collection	3 h	2 h, 12 min
T5: HTML Page for Exporters	Day 5 Data Collection	2 h	1 h, 50 min
Ace meeting	Next Leap Planning	1 h	45 min

Table 6. Activities of week 4

Task	Work plan (meeting)	Estimated time	Actual time
T1: Data Entry in Importers Page	Day 1	2 h	2 h, 15 min
T2: Data Entry in Exporters Page	Day 2	3 h	2 h, 45 min
T3: Agents Page with raw data	Day 3 Collection of Data	3 h	2 h, 30 min
T4: Insertion of agents data	Day 4	3 h	2 h, 50 min
Logo Designing	Day 5	2 h	1 h, 35 min
Ace Meeting	Next Leap Planning	1 h	45 min

Week 5

The data of the agents was added to the website. Logo, that was previously designed, was added to the header and the footer of the website. The "Contact Us" form was also added to the website and integration with the database was accomplished in this week. Table 7 lists down the activities performed.

Table 7. Activities performed

Task	Work plan(meeting)	Estimated time	Actual time
T1: Agents List and Contact Information	Day 1	2 h	2 h, 15 min
T2: Logo on web header and footer	Day 2	1 h	1 h, 18 min
T3: Contact Us Page	Day 3	3 h	2 h, 30 min
T4: Creation of table at MySQL	Day 4	1 h	1 h, 25 min
T5: Integration of the contact form with Database Text File	Day 5	4 h	3 h, 35 min
Ace Meeting	Next Leap Planning	1 h	45 min

fridaya@uef.fi

7 Refinement of the Process After Evaluation

The process model was implemented to the project and some refinements were required. The biweekly meetings were changed to weekly meetings. The reason for this change was that if in case the student's work of two weeks is not up to the mark or requires some major changes, then that should be incorporated as soon as possible. For this reason, the time of two weeks is little more and should be changed to one week in order to avoid time wastage. In this case recourses for only a week will be wasted and they have enough time to develop the refined work in the next week. Prototype iterations are important when it comes to good design. They should be iterated to improve the design until appropriate design for the software is achieved. There should be at least three prototype iterations.

- Name: Tour de Force
- Team: Crew
- Advisor: ACE
- Daily meetings of students: Daily Meetup
- Weekly meetings: Leap Meetup (Changed from biweekly to weekly)
- Requirements of project: Requirement Inventory
- Portion of work to be done: Activity List
- Prototype Iterations.

8 Revised Process Model

Changes were made in leap meetup and prototype iteration was introduced to the model. The changes made are reflected in the following points.

- Leap Meetup: For the maintenance and scheduling of Requirement Inventory and Activity List, weekly meetings will be conducted by Ace with the crew. This will also serve the purpose of updating the project progress and future plan. This meetup will be known as "Leap Meetup".
- Prototype Iterations: Prototypes will be mandatory as they portray the first image of the system. There should be at least three iterations.
- Name: Tour de Force
- Team: Crew
- Advisor: ACE
- Daily meetings of students: Daily Meetup
- Weekly meetings: Leap Meetup
- Requirements of project: Requirement Inventory
- Portion of work to be done: Activity List
- Prototype iterations (At least three iterations).

fridaya@uef.fi

9 User Study of Proposed Model

After revising the design of the proposed model, 42 students of Final Year were asked to implement this model in their Final Year Projects and later were presented a questionnaire based on ratings about proposed model. They were asked to rate out of ten on satisfaction, fulfillment of requirements, accomplishment of features, time management, and chances of interaction, involvement of advisor, understanding of process, success rate and complexity. The average ratings are illustrated in Fig. 11.

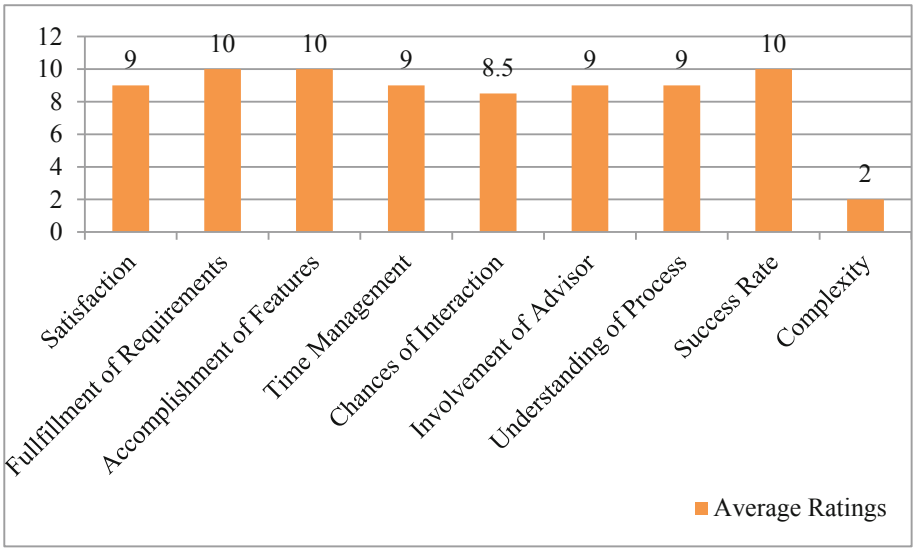


Fig. 11. Validating factors of process model

It was noted that Tour de Force was able to achieve 9/10 satisfaction points and so was the understanding of the process. The complexity of process model achieved only 2/10 points. This shows that the process model is not complex and is easy to understand and comprehend. Fulfillment of requirements and accomplishment of features scored 10/10 which is the most important and significant success factor of a process model. Time management scored 9/10 points which indicates that the way in which time was managed was helpful in accomplishing the tasks on time. Chances of interaction were seen as most noticeable point i.e. 8.5/10. Interaction helps in making better decisions and accomplishing tasks in lesser time. The secret behind every success is interaction. Involvement of Advisor achieved 9/10 points which conveys the idea that the Advisor is much more involved and is taking more interest than earlier. The last but not the least, success rate was rated. Students rated success rate as 10/10. This shows that the purpose of designing a customized process model was achieved successfully.

fridaya@uef.fi

10 Discussion and Analysis

Students find it hard to implement the existing software process models in their Final Year Projects as students have minimal budget and have small team and fixed amount of time. As they are still under training and learning phase, they are not much experienced to handle large software process models. Hence, there arises a need for a software process model which is designed keeping in consideration the budget, time and skills available. Waterfall and prototype models are not well understood whereas Tour de Force was able to achieve 9/10 points in understanding. Spiral and iterative model are understood but their cost estimation is very high. Cost estimation in waterfall is very difficult task to perform. RUP model offers too much documentation whereas in Tour de Force model, Software Requirement Specification document is replaced with Requirement Inventory. Not only documentation work is reduced but it also saves time.

Scrum requires experienced Scrum Master which is not possible for the students of the Final Year to have as they are still in the phase of training and learning. Looking at the experience aspect, if we allocate the title of "Scrum Master" to the advisor, it will still not solve the problem as the advisor is not much involved and does not show interest towards the Final Year Project. Tour de Force eliminates this constraint by allocating the role of Ace to the Advisor and provides a way out in which the Advisor is more involved in the project. In this way, the expertise of Advisor is also made use of and the advisor is more aware about the happenings and progress of the project. The advisor also plays the role of the Product owner which means he provides his feedback at every step and at the end the desired product is achieved. The chances of interaction are also high in Tour de Force Model which is a significant factor in the success of any project.

Tour de Force is a process model similar to existing process models but with a flavor of academic level. It makes it easier for the students to apply and get the taste of how work is done in professional life. It provides the most suitable way of time management which not only contributes in completion of project on time but also limits and reduces help from outsources and encourages students to produce their own work.

11 Conclusion

In this paper, different software process models are examined and concluded that due to certain restrictions and limitation they cannot be applied at academic level. Students face difficult in choosing and implementing process models in their Final Year Project which results into incomplete features, time wastage, less chances of interaction, lack of interest of advisor and help from outsources. Tour de Force is designed keeping in consideration the level and sources available for the students of Final Year. 56.88% respondents said that their project progress was affected due to incomplete accomplishment of features and requirements. While 88.99% respondents said that there is a need for a customized software process model for the Final Year Project. 38.53% respondents had no idea about which existing process model would suit their need for the Final Year Project. 85.32% respondents showed their willingness to use software process model in their Final Year Project and 72.48% respondents acknowledged the

fridaya@uef.fi

importance of process model. All above listed factors were incorporated in Tour de Force process model and the average ratings depict the success of the model. Tour de Force was able to achieve 9/10 satisfaction points, 8.5/10 chances of interaction points, 10/10 success rate, fulfillment of requirements and accomplishment of features points. Involvement of advisor and understanding of process achieved 9/10 points.

In light of these results, it can be concluded that Tour de Force is the best model to implement in Final Year Projects.

References

1. Jawitz, J., Moore, R., Shay, S.: Management and assessment of final year projects in engineering. Int. J. Eng. Educ. **18**(4), 472–478 (2002)
2. Shuman, L.J., Besterfield-Sacre, M., McGourty, J.: The ABET professional skills - can they be taught? Can they be assessed? J. Eng. Educ. **94**(1), 41–55 (2005)
3. Tsui, F.: Essentials of Software Engineering, 2nd edn. Jones & Bartlett Learning, Sudbury (2010)
4. James, H.A., Hawick, K.A., James, C.J.: Teaching students how to be Computer Scientists through student projects. ACM Int. Conf. Proc. Ser. **106**, 259–267 (2005)
5. Shukla, A., Saxena, A.: Which model is best for the software project? a comparative analysis of software engineering models. Int. J. Comput. Appl. **76**(11), 18–22 (2013)
6. ISO/IEC/IEEE 29119–2:2013(E), pp. 1–68 (2013)
7. What is Waterfall model- advantages, disadvantages and when to use it? ISTQB Exam Certification. http://istqbexamcertification.com/what-is-waterfall-model-advantages-disadvantages-and-when-to-use-it/. Accessed 13 Apr 2018
8. Ludewig, J., Lichter, H.: Software Engineering: Grundlagen, Menschen, Prozesse, Techniken. Heidelberg (2010). ISBN 978-3-89864-662-8
9. What is Spiral model- advantages, disadvantages and when to use it? ISTQB Exam Certification. http://istqbexamcertification.com/what-is-spiral-model-advantages-disadvantages-and-when-to-use-it/. Accessed 13 Apr 2018
10. What is Incremental model- advantages, disadvantages and when to use it? ISTQB Exam Certification. http://istqbexamcertification.com/what-is-incremental-model-advantages-disadvantages-and-when-to-use-it/. Accessed 13 Apr 2018
11. What is Iterative model- advantages, disadvantages and when to use it? ISTQB Exam Certification. http://istqbexamcertification.com/what-is-iterative-model-advantages-disadvantages-and-when-to-use-it/. Accessed 13 Apr 2018
12. What is Prototype model- advantages, disadvantages and when to use it? ISTQB Exam Certification. http://istqbexamcertification.com/what-is-prototype-model-advantages-disadvantages-and-when-to-use-itLast. Accessed 13 Apr 2018
13. Schwaber, K., Sutherland, J.: The Scrum Guide™, November 2017
14. Kruchten, P.: The rational unified process: an introduction. Addison-Wesley, Boston (2004)
15. Freitas, H., et al.: The Focus Group: A Qualitative Research Method, February 1998

fridaya@uef.fi

Virtual Testbeds with Meta-data Propagation

Thomas Kuhn(✉), Pablo Oliveira Antonino,
and Andreas Morgenstern

Fraunhofer Institute IESE, Kaiserslautern, Germany
{thomas.kuhn, pablo.antonino,
andreas.morgenstern}@iese.fraunhofer.de

Abstract. Testing of embedded software today is via Hardware in the Loop (HiL) testbeds. Virtual HiL testbeds provide simulated testing environments that resemble important components of systems under test and enable continuous testing of embedded software components. Creation of virtual testbeds requires coupling of specialized simulators. Existing simulator couplings focus on functional data, which are inputs and outputs of software components under test. In this paper, we discuss the additional propagation of meta data as enabler for additional evaluation scenarios that cover for example timing failures, error propagation, and unwanted feature interaction. With traditional simulator coupling, those effects are difficult to evaluate. We discuss the integration of meta data into testbeds, the development of virtual prototypes with simulator coupling, as well as challenges related to meta data propagation and simulator coupling. We highlight our solution for supporting meta data in virtual HiL testbeds, and evaluate its applicability in context of two realistic application examples.

Keywords: Virtual testbed · Simulation · Simulator coupling · Meta-data · Feral

1 Introduction

The next generation of embedded systems will be much more integrated with each other. The increasing level of integration between formerly independent systems yields unwanted feature interactions that result in unexpected system behavior. For example, driver assistance systems might apply force on the steering wheel to notify the driver that the car is about to change lane. The ESP system needs to know that this force is a notification only, and does not indicate a steering request of the driver. Today, Hardware-in-the-Loop (HiL) testing is used for the integration testing of automotive systems. It integrates physical control devices with simulated physics [8]. The aforementioned integration challenge brings traditional HiL testing to its limits.

Virtual Harware-in-the-Loop (vHiL) tests substitute real hardware components of HiL environments with simulation models for platform components and networks to create virtual prototypes. They are therefore much more flexible and can realize lot more system configurations and test beds. Furthermore, they can be replicated at very low cost and are available in early stages of development processes. The accuracy of

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 902–919, 2019.
https://doi.org/10.1007/978-3-030-22868-2_62

fridaya@uef.fi

virtual prototypes is determined by the accuracy of simulation models and by the accuracy of the used simulator coupling framework.

The required accuracy of a virtual prototype depends on the type of tests that it needs to support. Functional testing only tests the function of a component, and does not include platform effects. The functional mockup interface (FMI) specification [1] for example defines an interface for the exchange of functional simulation models. Virtual prototypes on this level of abstraction therefore only need to couple functional components. Platform testing that evaluates for example the effect of communication or scheduling delays to functional behavior require additional platform simulation components, which simulate task scheduling, CPU, or network behavior.

Despite of substituting physical HiL test components with simulation models, vHiL testing is similar to HiL testing. The vHiL environment executes test-cases that stimulate components under test and compare their outputs with expected results. Optionally, they consider timing as well. Virtual HiL testing furthermore has the potential to address additional integration challenges in embedded systems that traditional HiL testing cannot address. One example is sensor data fusion, which relies on sensor data that has been sampled at the same time. Pre-processing, filtering, and communication of sensor data, as well as scheduling effects add processing delays. This might may yield the fusioning of sensor data sampled at different times, and therefore erroneous results. Traditional testing only detects wrong behavior of algorithms that use this data, tracking this effect to sensor data fusioning would be difficult. In vHiL environments, the detection of this kind of error would be easy, if the basic error, the fusioning of data that has been sampled at different times, could be detected.

We therefore have extended our vHiL framework FERAL [2] with the ability to use meta data during testing. When adding sampling time to raw sensor data, FERAL can define an assertion that detects incorrect sensor data fusioning. In this paper we describe our approach of attaching meta data to vHiL testing. We describe the propagation of meta data in virtual HiL environments with coupled simulators, and we discuss application scenarios of simulator meta data. We evaluate the potential of meta-data in vHiL simulations using two realistic examples.

The remainder of this paper is structured as following: Sect. 2 surveys related work. Section 3 describes functional coupling of simulators. Section 4 describes the propagation of meta data in virtual prototypes. Section 5 evaluates our approach with two application examples. Section 6 draws conclusions and lays out future work.

2 Related Work

A common domain for simulator coupling is the integration of network simulation and functional behavior. With the increased connectivity of embedded systems, network performance becomes a key factor for system development. The authors of [7] have coupled the ns-2 network simulator with Simulink to evaluate networked control loops. The created simulation solution, PiccSIM, evaluates the performance of networked control loops, with respect to delay and jitter. The presented simulator coupling focuses on functional data. The ns-2 internally uses meta data to mark erroneous frames, but the coupling does not propagate this information to other simulators. The authors of [6]

fridaya@uef.fi

present another coupling in the same domain. They describe the coupling of Simulink and OPNeT simulators to evaluate the performance of networked control systems over wireless networks. Again, only functional data is considered. The authors of [10] couple OMNeT++ network simulator [5] and the road traffic simulator SUMO [9] to study impacts of communication networks to Car-to-X systems.

Since the development of simulator couplings from scratch is highly time consuming, several frameworks support the development of coupled simulations: Ptolemy [3] defines an approach for integrating heterogeneous behavior models using actor oriented semantics. It explicitly implements models of computation and communication (MOCC) as directors, which control the execution and communication between actors. Both actors and directors define the same inter-face, which enables their nesting. The resulting approach hierarchically couples semantic models and enables their integrated execution.

The Functional Mockup Interface (FMI) [1] of the MODELISAR project is another approach for the coupling of simulation models. It focuses on functional coupling that enables the simulation of system behavior without deployment effects. The standard defines a common interface and semantics for simulation models that extends the basic actor interface defined by [3]. Functional Mockup Units (FMU) encapsulated simulation models, which may expose two interfaces: the co-simulation interface and model exchange interface. Co-Simulation interfaces realize a black box interface, model exchange interface enable white box access that supports communication of the internal state of FMUs and the processing of numeric events like discontinuities. A host algorithm uses both interfaces to control the execution of FMUs; developers are free to implement any host algorithm that conforms to the basic semantic model. Compared to the basic actor interface described in [3], the functional mockup interface enables the definition of maximum times that a FMU may advance before it needs to return control to the host algorithm.

The Distributed Interactive Simulation (DIS) [11] standard defines exchange formats for the coupling of simulation models. Simulation models execute in real-time, coupling is via UDP or TCP protocols. Coupled simulators need to send periodic heartbeats to indicate their presence. A management of the coupled simulation models is not part of the standard. The High Level Architecture (HLA) [4] is a more recent standard for simulator coupling. It extends the concepts of DIS by a central manager and the ability to reduce network load by providing simulation models only with the data they need. Simulation models are integrated as federates; their execution is controlled by a central manager. HLA defines the overall architecture of simulator coupling, but leaves the semantics of the central manager to implementers. It provides no guidelines for the integration of semantic models.

The aforementioned simulator couplings are only a fraction of published simulator couplings. However, to our knowledge, all simulator couplings only propagate functional data between simulators. Meta data is never considered. We therefore analyzed the challenges related to inter-simulator meta data propagation and devised our solution approach from it.

fridaya@uef.fi

3 Simulator Coupling

Virtual HiL testing requires the virtual integration of the hardware and software components that form the embedded system under test. We differentiate system and simulation components. System components are components of the system under test, e.g. software functions that will become part of the final product. Simulation components encapsulate simulation models for the vHiL.

3.1 Virtual Prototypes of Embedded Systems

Figure 1 illustrates a simple vHiL simulation setup that consists of both system and simulation components. Two sensor components periodically sample data, timestamp this data, and forward it to two filter components that remove invalid values. A simulated Network connects filters with the receiver that implements sensor data fusioning. When the data is fusioned at receiver, the evaluator functions checks the resulting fusioned data set for errors, e.g. due to network and scheduling delays.

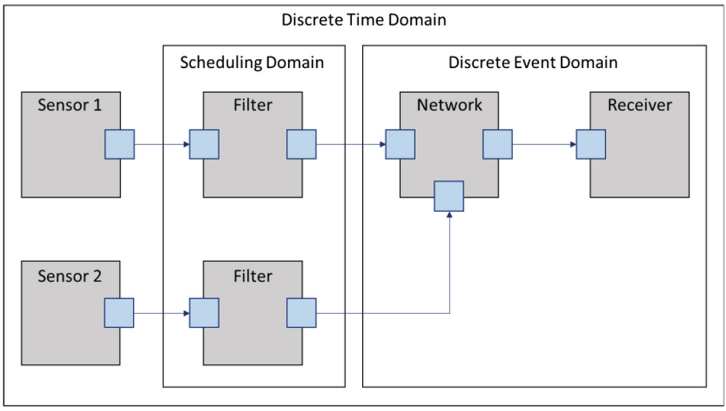


Fig. 1. Simulator coupling example

Our simulation model realizes message-based inter component communication. Message exchange is via links that connect output to input ports, as shown in Fig. 1. Virtual prototypes need to integrate different components of the system under test and platform simulation models. Each of these components conforms to a model of computation and communication (MOCC) that describes the basic semantics of a component in a real system. It defines when a component executes and how it communicates. This might be for example a time triggered task that conforms to a discrete MOCC, or a hardware component that reacts immediately to signal change, which is a discrete event MOCC. An accurate simulator coupling needs to integrate the semantics of all simulation components into one holistic simulation.

Complex virtual prototypes integrate simulation models that conform to different MOCCs. The example above illustrates the hierarchical coupling of three MOCCs.

fridaya@uef.fi

The MOCC of the *Discrete Time Domain* is the top-level MOCC. It executes all contained MOCCs in discrete time steps and ensures their synchronized execution. The MOCC of the nested *Scheduling Domain* simulates the task scheduling of an operating system. The *Discrete Event Domain* processes incoming events and executes receiver components. We realize MOCC coupling by coupling directors and simulation components into a tree structure. A MOCC controls the execution and communication of all directly contained components, which are either system components, simulation components, or MOCCs. Publications [2] and [3] describe principles and semantics for hierarchical coupling of MOCCs.

3.2 Simulation Meta Data

Every component of a vHiL simulation represents a part of the system under test. A component $c_i = (D_i, D_o, f_{beh})$ defines input ports D_i, output ports D_o, and a behavior function f_{beh}. Behavior functions define the functional behavior of a component, e.g. how it transforms input signals to output signals. Our virtual prototypes use an event-based mechanism for inter-component communication. When a component executes, it therefore reads events from input ports and transmits events to output ports. An event $e = (e_{ID}, e_{type}, e_{val}, e_{time}, e_{meta})$ is a five tuple of a unique event ID e_{ID}, the event type e_{type}, the value e_{val} that is transported by the event, the expiration time of the event e_{time}, and optionally additional meta data e_{meta}.

We attach simulation meta data to inter-component communication events. It carries key/value pairs that describe additional information to functional data, which is for example the sampling time of a value. Compared to functional data, simulation meta data is not actively processed by simulation components, and therefore will not directly affect the functional behavior and the outcome of a test case. However, it supports testers with identifying the causes for test case failures, and with detecting improper system behavior that was not revealed by failing test cases.

Since meta data is attached to inter-component communication events, meta data propagation depends on the MOCC that controls the execution of a simulation component. To understand meta data propagation, it is therefore necessary to understand common MOCCs for simulation.

3.3 MOCCs for Simulation

MOCCs control execution and communication semantics of components. Most components conform to either discrete time or discrete event semantics. The discrete time MOCC periodically calls all contained components. Receivers for component input ports D_i implement the behavior of buffers. Every input port buffers one value only, and this value remains in the port until it receives a new value. A component can therefore read an input port value multiple times while it executes, without ever receiving a changed value. Discrete time MOCCs commonly execute discrete control and simulation algorithms. For example, Simulink models and simulation models that conform to the Functional Mockup Interface (FMI) [1] require a discrete time MOCC. The director behavior for the discrete time MOCC is as following:

fridaya@uef.fi

```
1: fire (t_start, d_active):
2:   if (t_now > t_start+d_active) return;
3:   if !(t_now) t_now = t_start;
4:   forwardMessages(S_inputPortProxies);
5:   while ((t_start + d_active) < (t_now + d_step)) {
6:     ∀ c∈C (c.fire(t_now, d_active-(t_now-t_start)))
7:     t_now = t_now + d_step;
8:   }
```

The *fire* operation activates the director that implements a MOCC. When the discrete time MOCC activates, it first checks the active period. This period defines the maximum time that the local clock may advance during execution to prevent simulation errors. The director executes all controlled components, and then advances the clock by the duration of the time step. This repeats until another execution would exceed the duration of the active period. In every execution, the contained components create a new set of output values. The resulting response behavior of a system component that conforms to the discrete time MOCC is shown below: a system component, when executed, samples values from any number of input ports p_i, executes the behavior function f_{beh}, and generates output values on output ports p_o. Components that process only one input value (from one port) at a time are single event processing components; other components are multi-event processing components.

```
1: fire(t_start, d_active):
2:   (p_o1,…,p_on) = f_beh(p_i1,…,p_in);
```

The director behavior for the discrete event MOCC is shown below. It manages an ordered event queue that sorts events based on their expiration time. Input port behavior places received events into the central event queue of the director. When the director executes, it processes waiting events in the queue until the queue is either empty, or until the expiration time of an event would exceed the director active period. Directors implement event processing by calling the *fire* operation of contained components that receive the event. A component therefore receives every event only once.

```
1: fire(t_start, d_active):
2:   forwardMessages(S_inputPortProxies);
3:   t_now = t_start;
4:   while (e = queue.peekNext()) {
5:     if (e.t_Event > t_start + d_active) return;
6:     if (e.t_Event > t_now) t_now = e.t_Event;
7:     e.comp.fire(t_now, d_active-(t_now-t_start))
8:     queue.remove(e);
9:   }
```

fridaya@uef.fi

System components that conform to the discrete event MOCC create output events as response to received events, during their first invocation, or when a previously set timer expires. Timer expiration is based on timer events that are placed into the central queue to ensure their execution at a defined time. When a worker is invoked, it processes a single event to ensure the proper ordering of processed events. We refer to all system components that conform to discrete event semantics as single input processing components.

```
1: fire(t_start, d_active):
2:   t_now = t_start;
3:   e = queue.getNext();
4:   handleEvent(e);
```

A MOCC therefore may execute the same component at different times, enabling it to processes input events to output events. Meta data attached to events is directly affected by that. Discrete time MOCCs create several output values out of one input value, even if that input value is unchanged between time steps. Likewise, different meta data might be created by invocations in different time steps out of the same input meta data. Discrete event semantics processes an input event only once. Discrete event components might create several output events from the same input event.

4 Meta Data Propagation

Meta data e_{meta} consists of meta data items $m_i = (m_{type}, m_{data}) \in m$, which are tuples of meta data type m_{type} and meta data information m_{data}. Meta data needs to propagate together with simulation events through the virtual prototype.

4.1 Event and Meta Data Propagation

The main challenge for propagating meta data through coupled simulations is that simulators process functional data contained in simulation events without knowing about meta data. Input events e_{in} are consumed and output events e_{out} are created based on their functional payload by simulation workers or by existing (native) simulators. Meta data is quickly lost in this process. Figure 2 illustrates this problem with an example of four system components that propagate events. The sensor component *Sensor* samples input values and attaches a time stamp as meta data of type m_{sens} to generated events. The network simulation component simulates a communication network. It forwards events that arrive on an input port to the network simulator. It transmits the payload that is contained in these events through a simulated network that simulates message loss, message re-ordering, message duplication, and transmission delays. The application processes received events from the network and generates output events that are received by the actuator component. By evaluating the time stamp that every event carries as meta data, the developer can evaluate end-to-end communication delay and jitter with the virtual prototype.

fridaya@uef.fi

As mentioned before, this kind of evaluation requires the correct propagation of meta data through all components. We categorize virtual prototype components into black box and white box components from the viewpoint of meta data propagation. White box components are aware of meta data. Their generated output events e_{out} are tuples $(e_{ID}, e_{type}, e_{val}, e_{time}, e_{meta})$ that include event meta data. Black box components are not aware of any meta data. Their generated output events e_{out} do not include event meta data. Instead, the simulator coupling needs to process meta data of input signals into meta data of output signals. This needs an explicit response behavior specification that defines how output meta data is created based on input event meta data. Simulator coupling populates output signals with meta data based on response behavior after a component finishes execution.

4.2 Response Behavior Specification

A response behavior specification defines possible reactions of simulation components to input event sequences. The simulator coupling uses this information to identify the behavior of simulation components, match input events with response behavior specifications, and to populate output events with meta data. One response behavior specification defines one possible reaction to a sequence of input events, and the meta data propagation for this alternative. The complete set of response behavior specifications defines all possible reactions of a component. We define response behavior r_{comp} for simulation component *comp* as following:

$$\mathrm{r_{comp}}(\mathrm{p_i} : \mathrm{e_i}, \ldots, \mathrm{p_n} : \mathrm{e_n}) = \{\mathrm{P_{o1}}, \ldots, \mathrm{P_{on}}\} \quad (1)$$

Each response behavior specification r_{comp} defines an input event type e_i for every component input port p_i. For single event processing components, a response behavior specification defines event types for one input port only. The same component may define multiple response behavior specifications r_i. The simulator coupling selects valid response behavior specifications by comparing types of specified input events for input ports with waiting event types at component input ports. Valid response behaviors match the type of all specified waiting events at all input ports. Instead of formulating equivalent response behavior specifications for a collection of input ports p_i, formulation of port ranges $[p_a, \ldots, p_n]$ is valid. A range matches if a message is waiting at exactly one port out of $[p_a, \ldots, p_n]$.

In some cases, it is not sufficient to match only input events for selecting response behavior for a simulation component, because e.g. due to event values or internal state, the same component might react differently on the same input event sequences. Response behavior specifications therefore also support the definition of output

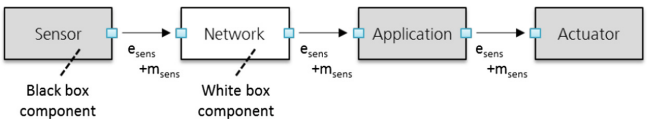


Fig. 2. Event propagation

fridaya@uef.fi

sequences P_{oi} for output ports p of simulation components. An output sequence P defines a sequence of generated outputs for every output port p:

$$P_{oi} = p : \{E_1, \ldots, E_n\}_1 \ldots \{E_1, \ldots, E_m\}_k \quad (2)$$

Output port p is either a named port of the simulation component, or the generic placeholder p_{gen} that matches any output port. The sequence of generated outputs E for every port consists of any number of output events e_{out} or meta-data operations (cf. Table 1). Generation of outputs is bound to the semantics of the MOCC that controls the execution of a component. Components that conform to the discrete time MOCC output one values per output port. They can also combine input events from all ports into outputs. Components conforming to discrete event MOCC process individual events and output sequences of events on output ports. For the response behavior to match, generated output sequences need to match the output sequences E defined in the response behavior specification.

Every generated output event e_i is specified by the response behavior specification as 4-tuple $e = (e_{type}, e_{val}, e_{time}, e_{meta})$ of event type e_{type}, event value e_{value}, delay e_{delay}, and optionally the meta data propagation specification e_{meta}. Delay e_{delay} is relative to the previously generated event on the same output port, or to the time of the component invocation in case of the first event in a sequence. It defines the delay until event e appears on that port. Response specifications may refer to types and values of events received events at input ports and previously generated events at output ports. Input and output ports are referred to by port names. This includes the event that is currently being generated. E.g., variable $e_{ip.type}$ represents the type of the event at input port i_p. Variable $e_{ip.value}$ then represents the event value. This enables defining a response behavior with output event values depending on the value of an input event or of previously generated output events.

Components provide output values based on the result of calculations. Response behavior should not need to define exact algorithm implementations. Especially in early stages of development, algorithm behavior is not fully known. Therefore, architects or developers may specify ranges of values without defining the exact response of a component to input events:

- Intervals *[a, b]* define a value interval that includes values a, b, and all values in between.
- Set of values *{a, b, c}* specifies any one of the given values in the set.

It is also possible to combine intervals and sets in response behavior. This way it is possible to define events that potentially carry values from different interval ranges. Our approach also permits the specification of processing delays e_{delay} with intervals and sets.

After a black box component was executed according to its MOCC, the simulator coupling needs to decide about the meta data propagation based on both the sequence of input events and the resulting output event sequence, considering type, value, and delay for every output event. If the coupling framework identifies no valid response behavior specification, it generates an error. If more than one response behavior

fridaya@uef.fi

specification is valid, the coupling framework randomly selects one specification and its meta data propagation.

If the response behavior specification for a single event-processing component does not define explicit meta data propagation for an output event, all meta data from the triggering input event is copied to this output event. Multi event processing components always require an explicit meta data propagation specification for output events. The meta data propagation specification $e_{meta} = \{m_1, ..., m_n\}$ consists of a sequence of meta data propagation operations (cf. Table 1). These operations enable developers to define how meta data needs to be passed from input events to output events. If, for example, a received erroneous value affects the current output of a component and all following outputs, an error marker may be stored in a variable v. All subsequently generated output events may have the value of variable v added to indicate an erroneous value.

5 Evaluation

We evaluate the added benefits of meta data propagation in virtual HiL testing in context of two examples. Both examples resemble realistic, but anonymized problems from automotive industry. They highlight problems that are difficult to identify using traditional HiL or vHiL testing without meta data use. We highlight the necessary overhead for defining response behavior specification for black box components and the benefit in return from the added meta data to inter component communication.

Table 1. Meta data propagation operations

Operator	Description
p_i	Copy meta data of input event at port p_i
$(p_i, m_{tin}) \rightarrow m_{tout}$	Rename meta data m_{tin} of input event at p_i to m_{tout}
$+p_i$	Add meta data from input event at port p_i
+v	Add meta data from variable v
+v, m_{type}	Add meta data type m_{type} from variable v
$-m_{type}$	Remove meta data of type m_{type}
−v	Remove meta data types contained in variable v
/	Delete all meta data
store(p_i, v)	Store meta data from input event at port p_i in v
clear(v)	Clear variable v
add(p_i, v)	Add meta data from input event at port p_i to v
add(p_i, m_{type}, v)	Add meta data m_{type} from input event at port p_i to v
remove(v, m_{type})	Remove meta data type m_{type} from variable v
put(t, k, p_i)	Put meta data from event at p_i into map t with key k
get(t, k)	Get meta data with key k from map t

fridaya@uef.fi

5.1 Feature Interaction Example

Our first application example covers unwanted feature interaction when integrating seemingly independent functions into one system. We integrate two functions: an air-conditioning control function *AcCtrl*, and a hill holder function *HillHoldCtrl* (cf. Figure 3). The function *AcCtrl* controls the air conditioning of a car. The driver sets the desired temperature of the car using a user panel. A temperature sensor measures the temperature of the driver cabin. Depending on the difference between the requested and the actual temperature, the *AcCtrl* function requests more or less power from the compressor. If more power is necessary, the compressor issues a torque request (*torqueReq*) signal to the engine controller. The hill holder function *HillHoldCtrl* is a comfort function. When the driver invokes the function by pressing a button, the brakes hold the car to prevent it from rolling backwards while the driver operates clutch and gas pedals. As soon as the driver wants to start driving, the hill holder detects a torque request and releases the brakes. The hill holder controller detects the intent to start driving by monitoring the powertrain CAN bus for the *torqueReq* signal. When the driver presses the electronic gas pedal, it transmits the *torqueReq* signal.

The virtual prototype for evaluating possible feature interactions defines component shells for black box components *AcCtrl*, *Compressor*, and *HillHoldCtrl*. Components *UserPanel*, *TempSensor*, *HillHolderBtn*, and *GasPedal* are white box components and attach meta data to inter-simulator events that identify the sender component. If no feature interaction occurs, output port p_{BA} of component *BrakeActuator* reacts to events from sender *HillHolderBtn* and *GasPedal* only. The response behavior specification for components *AcCtrl*, *Compressor*, and *HillHoldCtrl* is as following:

$$\mathrm{r_{AcCtrl}}(\mathrm{p_r} : \mathrm{r_{temp}},\ \mathrm{p_i} : \mathrm{r_{temp}}) = \{\mathrm{p_p} : \{\mathrm{e}(\mathrm{r_{pwr}},\ [0,\ 100],\ 0\mathrm{s},\ +\mathrm{p_r} + \mathrm{p_i})\}, \{\}\} \tag{3}$$

$$\mathrm{r_{Compressor}}(\mathrm{p_r} : \mathrm{r_{pwr}}) = \{\mathrm{p_t} : \{\mathrm{e}(\mathrm{r_{trq}},\ [0,\ 100],\ 5\mathrm{ms},\ \mathrm{p_r})\}, \{\}\} \tag{4}$$

$$\mathrm{r_{HillHoldCtrl}}(\mathrm{p_h} : \mathrm{r_{hold}},\ \mathrm{p_r} : \mathrm{r_{trq}}) = \{\mathrm{p_o} : \{\mathrm{e}(\{\mathrm{r_{hold}}, \mathrm{r_{release}}\}, , , + \mathrm{p_h} + \mathrm{p_r})\},\ \{\}\} \tag{5}$$

Function *AcCtrl* is a time triggered function that reacts to values on input ports p_r and p_i by outputting a new power request value r_{pwr} on output port p_p, or by providing no new output. The compressor function outputs a new torque request value as response to a power request r_{pwr}. The response behavior for the hill holder control function indicates the possible output values r_{hold}, and $r_{release}$ as reaction to input values at ports p_h, and p_r.

We used our simulator coupling framework FERAL [2] to implement a simulator coupling that did execute the test suites T_{AcCtrl} and $T_{HillHolder}$ for both systems *AcCtrl* and *HillHolder*. The virtual prototype executes both test suites independently on the integrated topology shown in Fig. 4. Functions connect to networks of a simulated car. A gateway connects both networks. The response behavior for CAN bus simulation components *CANSim* and for the Gateway is defined as following:

fridaya@uef.fi

$$\begin{aligned} r_{\mathrm{CANSim}}([p_1, \ldots, p_n]) &= \{\mathrm{put}(t_{\mathrm{id}}, e_{\mathrm{id}}, p_i)\} \\ r_{\mathrm{CANSim}}() &= \{p_{\mathrm{gen}} : \{e(,,,\mathrm{get}(t_{\mathrm{id}}, e_{\mathrm{id}})\} \end{aligned} \tag{6}$$

When the CAN simulator receives an event on any input port p_i, the meta data of the input event is put into map t_{id} with key e_{id} of the received event. Output events are matched by the response behavior defined by Eq. 6. The specified response behavior matches any output event on any port, as it defines no condition for event type e_{type}, event value e_{value}, and delay e_{delay}, and accepts event on generic port p_{gen}. Meta data for output events is retrieved from map t_i with the unique event id as key. This is possible, because the event propagation semantics of simulation component *CANSim* is known; the component forwards events and retains their unique ID e_{ID}.

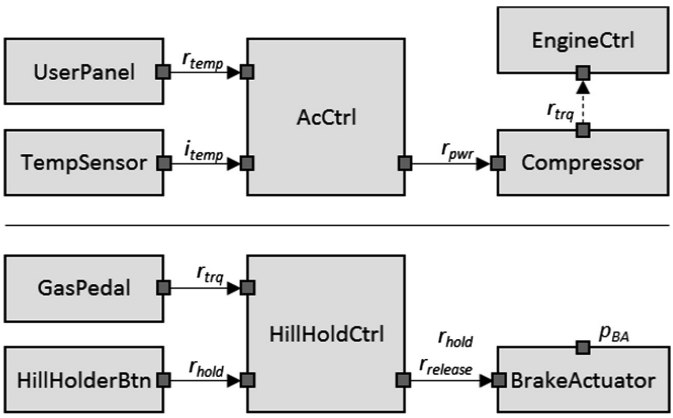


Fig. 3. Functions *AcCtrl* and *HillHoldCtrl*

Gateway developers in our example context did realize that issuing of a torque request command from the comfort CAN could yield a safety relevant behavior, therefore, the gateway only forwards this kind of signal from the comfort CAN to the power train CAN (*PowerTrainCAN*) if the car is standing still and the clutch is neutral. For the response behavior this is not relevant. The gateway either forwards received messages or not. This is covered already by the specification of r_{CANSim}. The response behavior specification for the Gateway is therefore similar.

$$\begin{aligned} r_{\mathrm{Gateway}}([p_1, \ldots, p_n]) &= \{\mathrm{put}(t_{\mathrm{id}}, e_{\mathrm{id}}, p_i)\} \\ r_{\mathrm{Gateway}}() &= \{p_o : \{e(,,,\mathrm{get}(t_{\mathrm{id}}, e_{\mathrm{id}})\} \end{aligned} \tag{7}$$

To identify feature interaction, we attach meta data to event sequences that are send as triggers to system components as part of a running test case. This meta data identifies the current test case. If test cases from two test suites are fully independent of each other, then they must not affect the outcome of each other. For example, the developers of the hill-holder system considered it independent of climate control in all cases. After executing all test cases individually, the virtual prototype executes test cases again.

fridaya@uef.fi

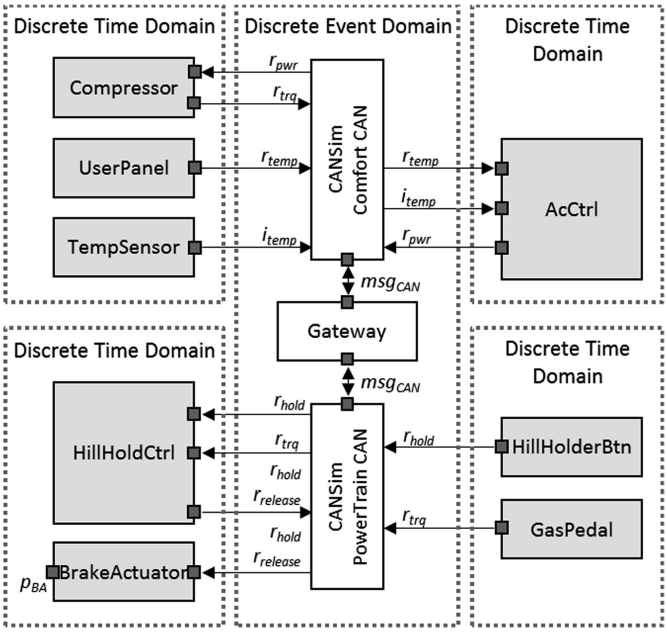


Fig. 4. Integrated functions *AcCtrl* and *HillHoldCtrl*

This second execution combines and permutates the execution of test cases from both systems. If the same actuator reacts to events from different test-cases, both are not considered independent. Response specifications propagate meta data specifications of input data to generated output events, testers may easily identify the test suite that did originate a signal.

In the example for our case-study, the developers of the *AcCtrl* function are unaware of the detailed implementation of the hill holder function that is part of some configurations; the hill holder developers on the other hand are unaware of the details of the *AcCtrl* function behavior. The resulting system yields an unwanted and safety relevant feature interaction. When the car was standing on a hill with hill holder activated, clutch being set to neutral, and the compressor of the air conditioning system did issue a torque request, this request did trigger as well the releasing of the brakes. This is not intended.

5.2 Response Times and Sensor Data Fusioning

Our second evaluation analyses end-to-end communication timing and whether components did combine input values that were sampled at different times. The latter is a considerable problem for algorithm developers in system with virtual sensors. Virtual sensors combine and process measured values from physical into virtual sensor values that cannot be measured directly.

In the example shown in Fig. 5, it is not possible to directly measure the wheel slip value. Therefore, a virtual sensor calculates this value based on measured vehicle speed

fridaya@uef.fi

and wheel speed. When the system is deployed, and processing of some sensor values take longer than the processing of other sensor values, it is possible that virtual sensors combine values that physical sensors did measure at different times. Depending on the control algorithm, this could yield significantly different behavior from pure functional models and in worst cases unstable control algorithms.

We created a virtual prototype with our FERAL simulation framework to simulate the deployment of the function net from Fig. 5. Figure 6 illustrates the prototype with three ECUs and one CAN bus network. Sensors are connected to ECUs, functions are deployed to the ECUs. A CAN bus network connects all ECUs. When the sensors sample an input value, they stamp the outgoing event with meta data m_{spl} that carries a time stamp. We did perform the evaluation by time stamping every output signal from physical sensors. Time stamps are propagated as meta data. When sensor values are combined, meta data time stamps are combined as well. For every event, it is therefore possible to identify whether it was generated based on sensor values that were sampled at the same time. Developers can furthermore evaluate the reaction time of the system, which is the time between measuring a value at a sensor and monitoring a reaction on the brake actuator.

The response behavior of wheel slip sensor function Fun_{WSS} is as following: The wheel slip sensor calculates wheel slip values based on measured wheel speeds (*ws*) and vehicle speeds (*vs*). The wheel slip value on output port of Fun_{WSS} carries a value between 0 and 100. The meta data propagation specification M_{wss} defines that meta data of both input events on input ports p_{ws} and p_{vs} should be renamed and forwarded to the output signal.

$$\begin{aligned} \mathrm{r_{wss}}(\mathrm{p_{ws}}, \mathrm{p_{vs}}) &= \{\mathrm{p_{wso}} : \{\mathrm{e}(\mathrm{ws}, [0, 100], 0\mathrm{s}, \mathrm{M_{wss}})\}\} \\ \mathrm{M_{wss}} &= \{\mathrm{m_{spl}} : ((\mathrm{p_{ws}}, \mathrm{m_{spl}}) \rightarrow \mathrm{m_{spl.ws}} + (\mathrm{p_{vs}}, \mathrm{m_{spl}}) \rightarrow \mathrm{m_{spl.vs}})\} \end{aligned} \tag{8}$$

The response behavior specification r_{bs} for the brake sensor defines inputs from type *brakeSens* on input port of Fun_{BS} to be forwarded as event type *brakeInd*. The value and meta data of the input event at port p_{bs} is renamed. Processing takes no time.

$$\begin{aligned} \mathrm{r_{bs}}(\mathrm{p_{bs}} : \mathrm{brakeSens}) &= \{\mathrm{p_{bso}} : \{\mathrm{e}(\mathrm{brakeInd}, \mathrm{e_{bs.val}}, 0\mathrm{s}, \mathrm{M_{bs}})\}\} \\ \mathrm{M_{bs}} &= \{\mathrm{m_{spl}} : ((\mathrm{p_{bs}}, \mathrm{m_{spl}}) \rightarrow \mathrm{m_{spl.bs}})\} \end{aligned} \tag{9}$$

Response behavior r_{abs} for abs controller function Fun_{ABS} defines a bt_{Ind} event as response to values on input ports p_{ms} and p_{be}. The output event carries an integer value between 0 and 100, processing takes between 0 and 5 ms. Meta data of type m_{spl} from both input ports is forwarded as meta data $m_{spl.ms}$ and $m_{spl.be}$.

$$\begin{aligned} \mathrm{r_{abs}}(\mathrm{p_{ms}}, \mathrm{p_{be}}) &= \{\mathrm{p_{bt}} : \{\mathrm{e}(\mathrm{btInd}, [0, 100], [0, 5_{\mathrm{ms}}], \mathrm{M_{abs}})\}\} \\ \mathrm{M_{abs}} &= \{\mathrm{m_{spl}} : ((\mathrm{p_{ms}}, \mathrm{m_{spl}}) \rightarrow \mathrm{m_{spl.ms}}, (\mathrm{p_{be}}, \mathrm{m_{spl}}) \rightarrow \mathrm{m_{spl.be}})\} \end{aligned} \tag{10}$$

The response behavior specification for the brake actuator defines an output event of type w_t as response to inputs on input ports p_{be} and p_{bt}. Value and processing time is not specified. Therefore, the response behavior is valid for any value.

fridaya@uef.fi

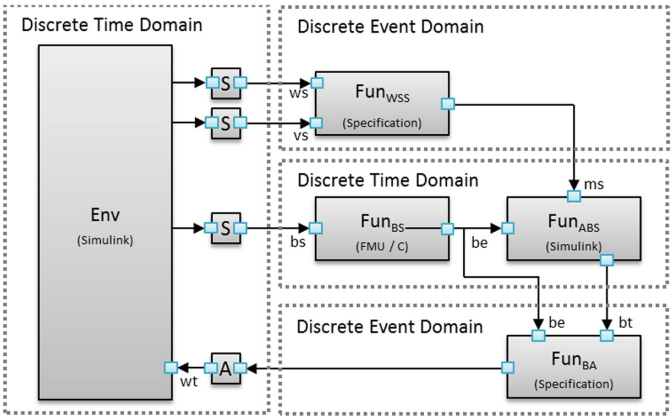


Fig. 5. ABS controller function set

$$\begin{aligned} r_{ba}(p_{be}, p_{bt}) &= \{p_{wt} : \{e(wt, , , M_{ba})\}\} \\ M_{ba} &= \{m_{spl} : ((p_{be}, m_{spl}) \rightarrow m_{spl.be}, (p_{bt}, m_{spl}) \rightarrow m_{spl.bt})\} \end{aligned} \tag{11}$$

During the execution of every test case, we monitor the sampling times of all processed events at the brake actuator function Fun_{BA}. As shown in Fig. 7, the time stamps of sampled values from wheel speed w_s and vehicle speed v_s sensors are consistent (both values were sampled at time 2 ms) when they are fusioned in the wheel slip sensor. When the wheel slip sensor value is processed together with brake sensor value by the ABS controller function, both input values carry different time stamps. The virtual wheel slip sensor requires time to process input data. During this time, the brake sensor samples a new value at time 3 ms. The system architecture does not account for that and processes both values as inputs for the ABS controller function Fun_{ABS}. The same holds for input values b_e and b_t of function brake actuator (Fun_{BA}). These inputs carry different time stamps well.

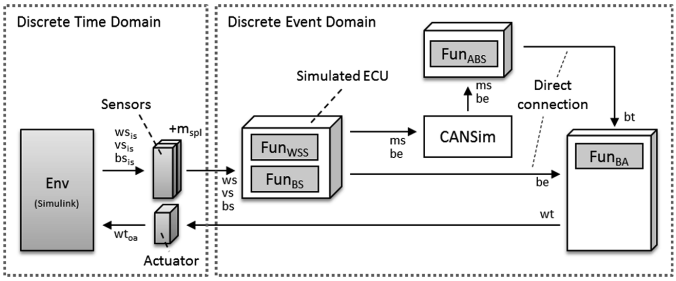


Fig. 6. ABS controller deployment

fridaya@uef.fi

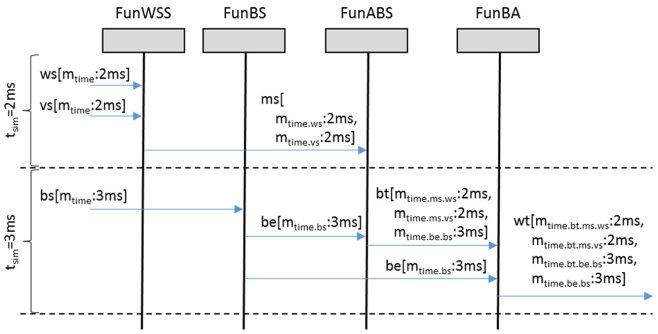


Fig. 7. ABS controller event sequence

5.3 Discussion

Both evaluations did cover testing challenges that are difficult to test in real-systems, or in traditional testing environments without meta data propagation. The reason for that is that both feature interaction and timing problems have the potential to significantly affect system behavior, but both effects to not predictably change the outcome of test cases. Therefore, they go undetected in many traditional test cases. Meta data enables the testers to add new types of assertions to test environments, which quickly identify feature interaction and timing problems, even if those do not have a direct negative impact on test case results.

Our application examples reflect realistic but not real testing challenges from the automotive industry. As part of the application examples, we evaluate the complexity of necessary response behavior specifications for black box simulation components. We did cover response behavior specifications for sensors, control functions, network components, and actuators. Our examples did show that the response behavior specifications were relatively simple and did sufficiently cover the behavior of black box simulation components to enable meta data propagation. In all cases it was possible to describe the behavior of system components, as well as the behavior of simulation components with response behavior specifications.

The simulation meta data did prove itself useful in both examples in detecting otherwise hard to identify errors. The examples did highlight that meta data complements functional testing very well – while functional test cases are executed, meta data reveals additional defects that are not detected by comparing functional output values from test cases, such as sensor data fusion errors and unwanted feature interaction. Both of these system aspects are difficult to test in a regular test setting.

6 Conclusion

In this paper, we have described the propagation of meta data in coupled simulations. We have highlighted the problems that arise due to unknown event propagation semantics and we did propose response behavior specifications as a feasible solution for defining high-level specifications of components.

fridaya@uef.fi

We have discussed two common MOCCs that control the execution of simulation components, and that are responsible for the propagation of simulation events. Considering the propagation of meta data attached to simulation events, we differentiate between black box and white box simulation components. White box simulation components have known semantics; based on observed output events, the simulator coupling framework automatically propagates meta data from matching input events. Network simulators are good candidates for white box components, because they only forward events from senders to receivers. For black box components, additional information is necessary to implement meta data propagation. Response behavior specifications provide this information. They describe the high-level behavior of simulation components based on sequences of input and output events. Every response behavior specification defines one meta data propagation strategy. Based on observed input and output event sequences, the simulator coupling can decide for a matching response behavior specification and can use the meta data specification to propagate meta data from input to output events.

Our evaluation discusses realistic application examples that highlight benefits of simulator coupling with meta data propagation. All examples realized analyses that would have been much more difficult without meta data propagation. The necessary response behavior specifications for black box components were relatively simple in all examples that we did consider.

Future work includes the identification and definition of re-useable response behavior specifications for common types of meta data and for common simulation components.

References

1. Blochwitz, T., Otter, M., Arnold, M., Bausch, C., Clauß, C., Elmqvist, H., Junghanns, A., Mauss, J., Monteiro, M., Neidhold, T., Neumerkel, D., Olsson, H., Peetz, J., Wolf, S.: The functional mockup interface for tool independent exchange of simulation models. In: Proceedings of the 8th International Modelica Conference, 20.-22 März, Dresden (2011)
2. Kuhn, T., Forster, T., Braun, T., Gotzhein, R.: FERAL – framework for simulator coupling on requirements and architecture level. In: Proceedings of the 11th ACM-IEEE International Conference on Formal Methods and Models for System Design (MEMOCODE ‘13), Portland (Oregon), USA (2013)
3. Eker, J., Janneck, J., Lee, E., Liu, J., Liu, X., Ludvig, J., Sachs, S., Xiong, Y.: Taming heterogeneity - the Ptolemy approach. Proc. of the IEEE **91**(1), 127–144 (2003)
4. Dahmann, J.S., Fujimoto, R. M., Weatherly, R.M.: The department of defense high level architecture. In: Andradóttir, S., Healy, K.J., Withers, D.H., Nelson, B.L. (eds.) Proceedings of the 1997 Winter Simulation Conference (WSC ‘97). IEEE Computer Society, Washington, DC, USA, pp. 142–149 (1997)
5. Varga, A., Hornig, R.: An overview of the OMNeT++ simulation environment. In: Simutools ‘08: Proceedings of the 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems and workshops, Marseille, France (2008)

fridaya@uef.fi

6. Hasan, M.S., Yu, H., Carrington, A., Yang, T.C.: Co-simulation of wireless networked control systems over mobile ad hoc network using SIMULINK and OPNET. IET Commun. **3**(8), 1297–1310 (2009)
7. Björkbom, M., Nethi, S., Eriksson, L., Jäntti, R.: Wireless control system design and co-simulation. Control Eng. Pract. **19**(9), 1075–1086 (2011)
8. Keranen, J.S., Raty, T.D.: Model-based testing of embedded systems in hardware in the loop environment. IET Softw. **6**(4), 364–376 (2012)
9. Krajzewicz, D., Erdmann, J., Behrisch, M., Bieker, L.: Recent development and applications of SUMO - simulation of urban mobility. Int. J. Adv. Syst. Meas. **5**(3&4), 128–138 (2012)
10. Schumacher, H., Schack, M., Kurner, T.: Coupling of simulators for the investigation of Car-to-X communication aspects. In: IEEE Asia-Pacific Services Computing Conference, (APSCC), Kuala Lumpur, Malaysia (2009)
11. 1278.1-2012 – IEEE Standard for Distributed Interactive Simulation – Application Protocols (2012)

fridaya@uef.fi

Challenges and Mitigation Strategies in Reusing Requirements in Large-Scale Distributed Agile Software Development: A Survey Result

Syeda Sumbul Hossain(✉)

Department of Software Engineering,
Daffodil International University, Dhaka, Bangladesh
syeda.swe@diu.edu.bd

Abstract. Requirements re-usability in a distributed software development project is applied to increase system productivity, reliability, quality, decreasing system development sprint and maintaining consistency between two identical systems, which later help to reduce both project time and cost. Nowadays, most of the projects are driven by market so that the intention of this research is to identify the challenges faced by practitioners in requirements re-usability in distributed large-scale agile projects and to find out how practitioners apply the concept of re-usability to mitigating those challenges in distributed large-scale agile software development from requirement engineering or re-usability perspective. In this study, a survey is used to identify requirement re-usability challenges and mitigation approach from practitioners. From a series of semi-structured interview, we have identified 14 challenges and 10 mitigation approaches into three categories, such as communication, coordination, and control from the global software engineering perspective. The findings from this research will help industry people to make a decision in their industry oriented activities.

Keywords: Requirement re-usability · Agile · Distributed · Large-scale · Software reuse · Global software engineering · Survey

1 Introduction

Agile methods are mainly designed to satisfy customers needs with an early and continuous delivery of software [1]. Reusing the requirements can help the practitioners to reduce both times and cost in software development. When requirement reuse is done in a large-scale project, it becomes a challenging job to manage and trace requirements.

Nowadays, most of the organizations are adopting agile software development methodology to overcome certain constraints. Many SDLC processes might meet the challenges of time, budget, and specification, but sometimes fail to satisfy

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 920–935, 2019.
https://doi.org/10.1007/978-3-030-22868-2_63

fridaya@uef.fi

users' contentment. Agile comes with the behavior of fulfilling users' satisfaction [2] by releasing and delivering product features (according to the priority) with rapid new versions simultaneously. After each version, the Agile team meets with the stakeholders and refine continuously as per customers' satisfaction. By doing so, an Agile team gets the customer' feedback frequently and enable to increase the quality of the product.

Agile methods focus on incremental and iterative development where requirements, design, implementation, and testing are continued throughout the project life cycle [3]. In traditional SDLC processes, all requirements are gathered and got a sign off from the customers, which puts a limit on changing and controlling requirements. Whereas in agile requirements can be changed even later in many iterations. After completion of each iteration to deliver rapid and repeatable components of projects, agile teams meet the customers by means of face to face communication. It is a belief among agile process proponents that people can respond more quickly and transfer ideas more rapidly when talking face to face then they can when reading or writing documentation [4]. Along with maintaining the quality of a product, some other benefits of agile SD are increased productivity, expand text coverage, improved quality with fewer defects, reduced time and costs, understandable, maintainable and extensible code, improved morale, better collaboration, and higher customer satisfaction. Although agile helps companies produce the 'right product' [3], there are many limitations of agile SD discussed in [5,6]. As Agile stands on continuous development, releasing software early, agile practitioners prefer to work with globally distributed teams. An agile company building a large product and facing time to market pressure needs to quickly double or quadruple productivity with a constrained budget [7], the distribution will help to add motion in deployment. For a distributed agile project, there are limited supports in communication, coordination, and control. In distributed SD, we cannot always do physical face-to-face communication, which is highly recommended in Agile SD. Distributed teams often work with fixed requirements. On the other hand, agile teams rely on ongoing negotiation among developers and the customers for determining the quality of requirements at different stages of SDLC. That is not quite possible in globally distributed projects. In short, agile methods are formal while distributed SD depends on formal process [8,9] focus on gaps between agile SD and distributed SD. As software organizations need to develop software at internet time [5], we can benefit by blending both of the advantages of agile and distributed software development [8]. Moreover, customization and modification are required when agile and distribution work together on a large project. The intention of this research is to identify the practitioners' strategies in reusing requirements in distributed large scale agile developments. Therefore, the research question (RQ)was set as:

RQ: *How practitioners apply the concept of re-usability in distributed large-scale agile software development?*

The rest of this paper is structured as follows: Sect. 2 presents background study. Section 3 presents the Research Methodology. Section 4 presents the

fridaya@uef.fi

results and analysis of the survey. Sections 5 and 6 present the discussions of the results and conclusion of the paper respectively.

2 Background Study

Though reuse started in 1969, Doug McIlroy first introduced the idea of systematic reuse as the planned development and widespread use of software components in 1968 [32], but Software companies do not want to reuse resources due to risk and technical support. In most cases, they do not feel comfortable to reuse artifacts for lots of constraints.

Different literature claimed that distributed software organizations face lots of problem with reuse requirements such as a change in customer requirements and technology, increased cost, delayed schedules, unsatisfied requirements and lack of software professional [12,34–36]. Another challenging issue is the satisfaction level of the customers to reuse requirement from previous identical projects and they feel insecure when reusing requirements from the similar identical project [37]. On the other hand, it shows that integrity with a distributed team is a difficult job to maintain or trace out the project resource [34,38,39].

The remedies of common problems such as communication, collaboration, and control of globally distributed agile projects are described in [7,8]. In the globally distributed agile project, team members share project-specific knowledge through frequent face-to-face interaction, effective communication and customer collaboration [14]. The use of the computer system in every discipline of working network forces to grow the complexity of software. Safety-critical systems in aircraft control, hospital patient monitoring, and power station management are done by software. This large-scale software needs more focus on quality, which can only be fulfilled through agile development. Large-scale distributed agile projects need more focus on requirements prioritization and specifications as well as emphasize rich documentation of requirements, whereas agile team is not put emphasis on documentation. The number of reported experiences is still limited and more expertise on how to apply the agile practices to different kind of distributed setting and teams is still needed [10]. Though there are insufficient literature's on a distributed large-scale agile project, a single case study described how large project worked in DAD [11]. In this case, efficient communication schema has to be established so that teams at different sites can interact virtually in real time.

The prospect of increased product reliability and stability with shorter development time and reduced cost continues to fuel the on-going interest in development. This approach advocates the acquisition, adaptation, and integration of reusable software components, including commercial-off-the-shelf (COTS) products, to rapidly develop and deploy complex software systems with the minimum engineering effort and resource cost [13].

fridaya@uef.fi

3 Research Methodology

3.1 Interview Design

Software engineering research seeks pragmatic results and empirical research in software engineering which is motivated us to do exploratory research [21–24]. According to Robson's [15], exploratory research is done to find out what is happening, seeking new insights and generating ideas and hypotheses for future research work. Exploratory research can be done both for qualitative and quantitative research. In this research, exploratory qualitative research has been done which will generate deeper knowledge rather than the same results. Interviewing is flexible to collect qualitative data which provides richer and deeper understanding or knowledge. Interviewing people provides insight into their world, their opinions, thoughts, and feelings [24]. There are three types of interviews - structured, unstructured and semi-structured [25,26,28]. In this research, a semi-structured interview was conducted in which researchers asked predefined questions but overall discussion covered with open-ended questions [27].

Sampling: To avoid prejudice in research results, the interview were conducted with multiple interviewees of different organizations. The criteria of selecting the interviewees are based on software companies (practices at least CMMI level 3) in which they work in, their experiences in distributed agile development in large-scale projects (at least 3 years), their role in distributed teams. The interviewees for this research work in different companies as IT manager, Project Manager, Software Architect and Developer, and Senior Java Developer in distributed agile projects located in different countries i.e. Bangladesh, China, India, Japan, Spain, Sweden, and USA.

Interview Guide: Interview questions were designed (see at Appendix Interview Questions) in two parts Tables 5 and 6, one part containing demographic or introductory questions (A.1) and the other one domain specific questions (A.2). As each interview output would be the input for the next interview, frequent updates are added in these interview questions according to the research needs.

Data Collection: The researcher mailed to the interviewees with the attached invitation letter. According to their given scheduled time, the interview were conducted. Two interviewers were conducted the interview; one led the interview and other took notes, asked follow up questions, and aided the primary interviewer when necessary are mentioned in [24]: (i) interviewee talked more with two interviewers than with one, and (ii) more follow up questions were asked in the two interviewer's case. A telephone interview provides the best source of information [28]. Skype conversation were conducted for collecting data. Interview were started with introductory questions, and then slowly turned into domain specific questions. While interviewing, the conversation had been recorded for further analysis.

fridaya@uef.fi

3.2 Data Analysis

As different data sources were used, this data and methodological triangulation will be turned towards the research goal. Thematic analysis helps to conduct qualitative research. Thematic analysis is considered as a foundation method for qualitative analysis [16]. It is used for identifying, analyzing and reporting patterns (themes) within data [17,18]. However, thematic analysis becomes one of the most used analysis methods in Software Engineering stated in [17,18]. Through this analysis, researcher will map themes based on the research questions. The recommended steps for thematic synthesis in SE discussed by Daniela and Dyba for systematic review methodology in [18] were followed.

Step 1: Extract Data
The analysis process was started with the verbal data, which was recorded throughout the interview. The data need to be transcribed first into written form for further analysis. Throughout the extracting process, transcribers get accustomed to the data by repeated listening verbal data. Data have extracted carefully from immersed findings from interviewees experience. Interviewees company details, project's context descriptions, team structures, and interviewees responsibilities have extracted into special segments. The overall extraction of data has rechecked by another researcher.

Step 2: Code Data
Code means the most basic segment, or element of raw data [16], which is more meaningful to the analysts. Researcher coded the transcribed raw data into meaningful groups as per interview guide was prepared. Researcher performed coding through Xmind software. Coding helps for evolving deep meaning and extracting specific information from an individuals perspective. Relevant findings from surrounding data are also coded.

Step 3: Translate Codes into Themes
When all data has been initially coded, this phase is begun. In this phase, researchers refocus on the coded data to find out patterns or theme among data sets. A theme is an outcome of coding, categorization, and analytic reflection, not something that is, in itself coded [19]. As mentioned before, researcher will map patterns or themes according to the research questions. On the translation progression to themes, coded data need to be rearranged and reclassified. For visual representation theme software tool can be used e.g. thematic networks, treemaps, mind maps, or xmind; and xmind is used for deriving theme (see Fig. 1).

Step 4: Create a Model of Higher-Order Themes
The theme which is emerged in previous steps can be explored and interpreted into a higher-order theme. The final product of this step can be a description of higher-order themes, a taxonomy, a model, or a theory [18], which will be the return of research questions, a distinction of theoretical data (from background studies) and practitioners' experiences, and indecision's of conflicts. The aim of this higher-order theme is to return to original research questions. From studying literature, researcher got the primary findings. Then connections between

fridaya@uef.fi

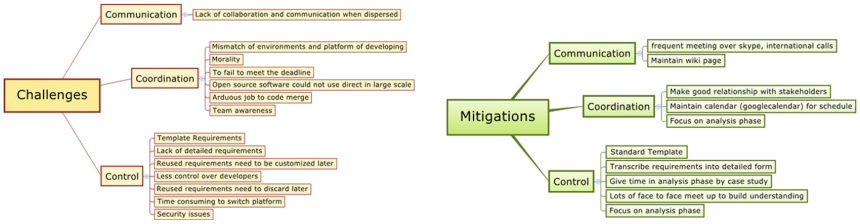


Fig. 1. Translate code into theme using Xmind

the primary data and the secondary one (secondary data which we get from semi-structured interview) were established. The relationships between primary findings and the previously extracted theme will be the key aspects of the higher-order theme. By comparing and contrasting these relationships, the final higher-order theme (see Fig. 2) was identified and defined.

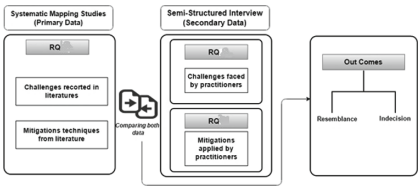


Fig. 2. Model of SE motivation from Beecham et al. [33]

Step 5: Assess the Trustworthiness of the Synthesis
Research findings and outcomes should be trustworthy as per the research methodology adopted. The trustworthiness of a synthesis depends on both the quality and quantity of the studied data. Poor methodological quality of primary studies and included literature can also affect the trustworthiness of the data synthesis. Trustworthiness will also depend on the methods used, e.g. measures taken to minimize bias, and weighting of studies according to quality [20].

4 Validity Threats

To handle validity threats, it is important to identify all possible factors that might affect the accuracy of results and analysis. Internal validity threats for qualitative research in general work with research biases and data transcription. External validity threats encounter to a reduced possible effect of research outcomes and in general, this approach applies in qualitative research [29]. Different types of validity threats ware encountered for both literature and interview such as construct validity threat, internal validity threat, external validity threats and reliability.

fridaya@uef.fi

4.1 Interview Validity Threats

For selecting interviewee at first their Linkedin profiles are visited to select same level and experienced industry people. After in general data collection of the interviewee, a formal mail was sent with an invitation letter to join interview and share their experience. To reduce the biases of the interview the majority of interview question were open-ended. Researchers also took help from experienced adviser to design interview questions. Every interview maintained the same sequence and started with same introductory and clarification questions. After completing interview, each interview was transcribed immediately to reduce might miss important information. Furthermore, researchers sent transcribe report of the interview to the interviewee in order to check whether interview data correctly transcribe and to confirm that transcribe content are the replica of participants thoughts, viewpoints, feelings, and experience. To maintain reliability during data analysis and understanding interview inner words, thematic qualitative data analysis was applied that helps to identify, analyze and theme within data. After completing extraction from transcribed data, researchers checked twice to identify any discrepancy. To validate outcomes of research, any in general comments have not been included from five interviewees in this finding.

4.2 Qualitative Validity Threats

This research encountered lots of major threats to turnout information from industry. To collect information from industry completely depends on willingness and availability of motivation of industry people. There are three types of validity threats with relevant to qualitative research [30], these are descriptive validity, interpretive validity, and theoretical validity [31]. Validity techniques are applied to reduce qualitative validity threats.

5 Results and Analysis

5.1 Interviewee Overview

Requirement reusability related challenges and mitigation strategies facing by practitioners are collected through interview of the distributed large-scale agile projects team from the different countries. Finally, 5 interviews with practitioners who job in different software company on the large-scale distributed project have been completed. In the Table 1, total overview of interview has been shown.

5.2 Findings from Interview

In global software development, the development team encounters communication, coordination and control challenges due to distributed nature of the project. In this interview, we have divided the findings into three categories such as communication, coordination, and control. In general, we have identified

fridaya@uef.fi

Table 1. Interview overview

Interviewee	Designation	Methodology	Team structure	Distribution type
I1	IT manager	Scrum	Sweden (4): Project Manager, Developers, Analysts, Testers China: 3 Tester	Distributed
I2	Project manager	Scrum, Waterfall	Spain (3): Requirements Analyst, Designers Bangladesh (9): QA, Developers	Dispersed
I3	Project Manager	Scrum	Bangladesh (20):2 testers, 4 R & D, 10 developers USA (35): Requirements Analysts India (7): Developers	Hybrid
I4	SW Architect & Developer	Not any specific agile but has fewwith scrum	Sweden:1 Developer China: 1 Project Manager, 1 Developer, 1 Software Architect, 1 Designer & Developer	Dispersed
I5	Senior Java Developer	Scrum	Japan: 1 Team Leader, 3 Developers, 1 Tester India: 10 Developers, 3 Testers China: 2 Testers	Hybrid

14 requirement reusability challenges and 10 mitigations approaches applied by the practitioners in the large-scale distributed agile project (see in Fig. 3 maximum challenges found in both coordination and control section, see in Fig. 4 maximum mitigations approach found in control section).

From Table 2 we observe that practitioners are seeking for solution of four challenges which have not solved yet.

5.3 Practitioners' Experience in Requirements Re-Usability

In this study, it is divided and graphically represented practitioners' experience in requirements re-usability in distributed large-scale agile project and classified

fridaya@uef.fi

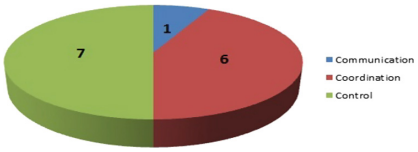


Fig. 3. Challenges faced by practitioners

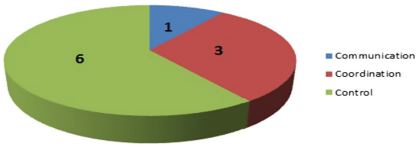


Fig. 4. Mitigation approaches applied by practitioners

those data according to communication, coordination and control in aspect of global software engineering (See in Figs. 3 and 4).

6 Discussion

The main objective of this research is to identify challenges and mitigation techniques in requirements reusability in distributed large-scale agile projects in regards to 3 C's factors in global software development. Those challenges and mitigation techniques are identified both from studies and survey results. Finally, both findings are related by applying thematic data synthesis technique.

6.1 An Overview of Different Challenges

Communication: Due to distributed nature of the project, team members have less knowledge about other team members which causes reusability challenging of requirements. Communication gap also occurs due to both temporal distance and team awareness. Inconsistency, low understanding of others role and knowledge transfer are the main challenges in literature. On the other hand, from interviewee's perspective communication gap does not affect more as daily scrum meeting, available of communication media minimize the communication challenges. But, some of our interviewees mentioned that communication gap impacts more in dispersed team than distributed one.

Coordination: Internal and external forces often cause rapid changes of the requirement in different phases of the development process. Along with that, it is also that found pair programming and share resources among distributed team members, low architectural repository, and unstructured documentation make requirement reusability challenging. Unclear requirements also arise

fridaya@uef.fi

critical situation at the time of reuse and match with existing resources. Platform dependency, open source software issues, code merging are the critical issues for industry development environment. Some of the ethical issues e.g. morality, trustworthiness and team awareness are the common issues identified in both findings. Moreover, coordination challenges are emphasized in both methodologies.

Control: Different literature focused that tracing requirements is a challenging job when the large-scale project is distributed. Along with, developing, deploying, and supporting systematically reusable software assets requires a deep understanding of application developer needs and business requirements. It is also found that practitioners feel low-level motivation at the time to contribute in the area of reuse and resources. Security violation problems are main controlling issues stated by the practitioners'; which have not yet logged in any literature.

Table 2. Challenges and mitigation strategies by practitioners

3 C's	Challenges	Mitigation strategies
Communication	Lack of collaboration and communication when dispersed	Frequent meeting over Skype, international calls, maintain wiki page
Coordination	Mismatch of environments and platform of developing	Still seeking for solution
	Morality	Make good relationship with stakeholders
	To fail to meet the deadline	Maintain calendar (Google calendar) for schedule
	Open source software could not use direct in large-scale	Still seeking for solution
	Arduous job to code merge	Focus on analysis phase
	Team awareness	Still seeking for solution
Control	Template requirements	Standard template
	Lack of detailed requirements	Transcribe requirements into detailed form
	Reuse requirements need to be customized later	Give time in analysis phase by case study
	Less control over developers	Lots of face to face meet up to build understanding
	Reused requirements need to discard later	Focus on analysis phase
	Need time to switch platform	Focus on analysis phase
	Security issues	Still seeking for solution

fridaya@uef.fi

Table 3. Common challenges from both literature and practitioner

Challenges	Descriptions
Lack of collaboration & communication when dispersed	When project and resource are distributed then control and maintain collaboration between them is so much challenging job and in a distributed development team communication gap is challenging issue for regional context
Arduous job to code merge	Tuff job to pair programming and share resources, need same skilled developer to reuse code
Team awareness	Low understanding the roles of others in distributed teams and they do not clear about all members role in a team and their activities
Morality	To reuse code developer are familiar with each others code, and trust each other but sometimes they do not trust each other
Template Requirements	Unstructured documentation sometimes are not clear but quality product depend on good and clear documentation quality
Mismatch of environments and platform for developing	Architectural changes and code changes dependency due to high-level architectural changes and low-level code changes
Lack of detailed requirements	Unclear requirements arise critical situation at the time of reuse and match with storage resources

Considering the common challenges faced by practitioners (i) communication and collaboration problem with the dispersed team, in this case, the project manager has less control over-dispersed team members. (ii) Reused code merging becomes more challenging in agile SD, as later changing in requirements and code discarding makes complexity in development. (iii) The most important issue is team awareness; team members should be clear about others responsibility which is foremost part of working in the distributed environment. (iv) The absence of morality among stakeholders might emerge apartness. (v) Different templates are often used in different projects; when to reuse requirements this unstructured requirements description might affect system quality and productivity. (vi) The inconsistency of environments, platforms, or architecture design makes SD a tedious job. (vii) Unclear and unstructured documents cause consuming more time in SDLC phases (see in Table 3, for more details about common challenges from both literature and practitioner).

6.2 An Overview of Different Mitigations Techniques

Communication: Communication problem should be the most affecting problem in Distributed Agile, as Agile supports frequent meeting and changes in requirements later on the development process. In a locally distributed team,

fridaya@uef.fi

communication is not detrimental for requirements reuse whereas in offshore development it impacts more. But, conveniences of communication media do not let that happen. Practitioners' use Skype, hangouts, international calls, Google drive, Google calendar (for scheduling meeting), mail, and wiki pages for frequent communication.

Coordination: To mitigate requirement reusability challenges practitioners apply formal modeling approaches. Integrated management of experience data and project related data to knowledge utilization and acquisition also help the team to reduce requirement reusability challenges. Different literature also claimed that practitioners apply test-driven development (TDD) approach for tightly coupled reusable requirements. The shared repository also used to manage reusable software component classification, project resources and ensure distributed access. Object-oriented programming concept also helps the team to reuse code, libraries, refactoring to design patterns and reusable feature sets. Industrial people stated that they emphasize on analysis phase to minimize the problem in reusing. They also make good relationship with stakeholders for maintaining coordination.

Control: Lesson learned from previously completed project repository helps to reuse requirements in a new project. Along with that, a graphical representation is used to ensure requirements interaction for individual requirements and extracting trigger events to trace reusable requirements. Software Product Line Engineering approach or feature driven approach also used to ensure faster delivery of similar or identical product and reuse knowledge and resources. To control the challenges practitioners' prefer standard template requirements, requirements should be transcribed in detailed form, long-term analysis phase.

Sometimes practitioners' applied approaches are different from existing literature. All of the practitioners' start their day and end with short-term meetings; some of them conduct a weekly meeting for long periods. These types of communication make themselves easier to resolve challenges like communication gap, team awareness, and morality issues. Schedule periodic checkups often are done for making good relationship with team members. Understanding wrapped requirements, requirements are transcribed into detailed form before providing requirements to the distributed teams. They use one standard template to maintain requirements, which also helps to trace requirements for further reusing. Most of the interviewee stated that they spent more times on analysis and design phases. Proper analyzing of requirements can be minimized challenges of reusing requirements. Practitioners' are still finding the solution for some challenges like the security issue, changes in later development (see in Table 4 for more details about common mitigations approach from both literature and practitioner).

fridaya@uef.fi

Table 4. Common mitigations from both literature and practitioner

Mitigations	Descriptions
Local workshop	Local workshop and frequent meeting over skype, international calls, maintain wiki page
Transcribe requirements into detailed form	Conceptual model, Rationale Model, Semantic wiki for requirements reasoning used to automatic distributed requirements analysis
Standard template	Graphical representation is used to ensure requirements interaction for individual requirements and extracting trigger events to trace reusable requirements
Lots of face to face meet up to build understanding	Face-to-face meeting conduct to find out team member motivation and reduce communication gap between teams
Focus on analysis phase	Prior in analysis phase due to low formality level in requirements diagrams
Schedule periodic check-ups	To ensure traceability between previous reuserequirements and current requirements and ensure project deadline maintained in calendar (google calendar)

7 Conclusions

In this section, it is summarized all activities performed in this thesis according to research questions and its outcomes. The main contribution of this research is to identify requirement reusability challenges and mitigation approach to improve software development productivity. Requirement reusability is efficient to reuse of requirements and resources which helps to minimize development cost, time and increase the reliability of product through applying knowledge from the lesson learned. Requirement reusability also helps to maintain sustainable software development by applying the concept of reuse, reduce and recycle.

In response to RQ, through a series of semi-structured interviews it is identified 14 challenges and 10 mitigations approach applied by practitioners. Interview results are divided into three categories according to global software engineering to make interrelation between both interview and current studies.

From the interview results, it is identified that practitioners and industry still seeking for four challenges solutions those are the mismatch of environments and platform of developing, open source software could not use directly in large-scale, team awareness and security issues. It is also mapped the common challenges and mitigations approaches applied by practitioners.

fridaya@uef.fi

A Interview Questions

A.1 Introductory Questions

Table 5. Introductory questions

IQ	Introductory questions
IQ1	What is your designation in the organization?
IQ2	What are your responsibilities in that project?
IQ3	How long have you been working with large-scale agile project?
IQ4	How long have you been working with globally distributed projects?
IQ5	Is in this project requirement reusable from previous project?
IQ6	Team Size and number of developing sites? Who does what? (Distributed? Dispersed feature teams? Or Hybrid team?
IQ7	Name, Age, Gender, Contact (email)

A.2 Domain Specific Questions

Table 6. Domain specific questions

DQ	Domain specific questions
DQ1	Which software development methodologies do you follow?
DQ2	Which agile practices are followed locally and globally? (Scrum, Kanban, XP, AUP etc.) (a) What is the team structure? (b) How many team members in each team? (c) How many number of sites? (d) With which team you frequently communicate?
DQ3	Through your experience, what are the benefits you have seen with reuse requirements in a large-scale project?
DQ4	Which type of requirements you generally reuse in agile SDLC different phases and Why?
DQ5	Which type of problems did you face when project was distributed and large-scale?
DQ6	What are the challenges to trace reusable requirements when it is distributed?

fridaya@uef.fi

References

1. Turk, D., France, R., Rumpe, B.: Agile Software Processes: Principles, Assumptions and Limitations. Technical Report, Colorado State University (2002)
2. Beck, K., et al.: Manifesto for agile software development (2001). 2006
3. Highsmith, J., Cockburn, A.: Agile software development: the business of innovation. Computer **34**(9), 120–127 (2001)
4. Keith, E.R.: Agile software development processes: a different approach to software design. A White paper (2002). http://cf.agilealliance org/articles/system/article/file/1099/file.pdf. Accessed 08 Oct 2012
5. Turk, D., France, R., Rumpe, B.: Limitations of agile software processes. arXiv preprint arXiv:1409.6600 (2014)
6. Turk, D., France, R., Rumpe, B.: Assumptions underlying agile software development processes. arXiv preprint arXiv:1409.6610 (2014)
7. Sutherland, J., et al.: Distributed scrum: agile project management with outsourced development teams. In: 40th Annual Hawaii International Conference on System Sciences, HICSS 2007, IEEE (2007)
8. Ramesh, B., et al.: Can distributed software development be agile? Commun. ACM **49**(10), 41–46 (2006)
9. Nisar, M.F., Hameed, T.: Agile methods handling offshore software development issues. In: Proceedings of INMIC 2004, 8th International Multitopic Conference, IEEE (2004)
10. Dingsyr, T., et al.: A decade of agile methodologies: towards explaining agile software development. J. Syst. Soft. **85**, 1213–1221 (2012)
11. Paasivaara, M., Durasiewicz, S., Lassenius, C.: Distributed agile development: using Scrum in a large project. In: IEEE International Conference on Global Software Engineering, ICGSE 2008, IEEE (2008)
12. Fortune, J., Valerdi, R.: Considerations for successful reuse in systems engineering. In: AIAA Space 2008 Conference & Exposition (2008)
13. Gill, N.S.: Reusability issues in component-based development. ACM SIGSOFT Softw. Eng. Notes **28**(4), 4–4 (2003)
14. Sahay, S., Nicholson, B., Krishna, S.: Global IT Outsourcing: Software Development Across Borders. Cambridge University Press, Cambridge (2003)
15. Robson, C.: Real World Research, vol. 2. Blackwell, Malden (2002)
16. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**(2), 77–101 (2006)
17. Cruzes, D.S., Dyb, T.: Research synthesis in software engineering: a tertiary study. Inf. Softw. Technol. **53**(5), 440–455 (2011)
18. Cruzes, D.S., Dyba, T.: Recommended steps for thematic synthesis in software engineering. In: 2011 International Symposium on Empirical Software Engineering and Measurement (ESEM), IEEE (2011)
19. Padilla, M.A.S., Saldaa, J.: The Coding Manual for Qualitative Researchers, pp. 121–126 (2011)
20. Dyb, T., Dingsyr, T.: Strength of evidence in systematic reviews in software engineering. In: Proceedings of the Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, ACM (2008)
21. Dingsyr, T., Dyb, T., Abrahamsson, P.: A preliminary roadmap for empirical research on agile software development. In: Conference Agile, 2008, AGILE 2008, IEEE (2008)

fridaya@uef.fi

22. Kitchenham, B.A., et al.: Preliminary guidelines for empirical research in software engineering. IEEE Trans. Softw. Eng. **28**(8), 721–734 (2002)
23. Wohlin, C., Hst, M., Henningsson, K.: Empirical research methods in Web and software Engineering. In: Web Engineering, pp. 409–430. Springer, Heidelberg (2006)
24. Hove, S.E., Anda, B.: Experiences from conducting semi-structured interviews in empirical software engineering research. In: 11th IEEE international symposium on Software metrics, IEEE (2005)
25. Seaman, C.B.: Qualitative methods in empirical studies of software engineering. IEEE Trans. Softw. Eng. **25**(4), 557–572 (1999)
26. Turner, I.I.I., Daniel, W.: Qualitative interview design: a practical guide for novice investigators. Qual. Rep. **15**(3), 754 (2010)
27. Flick, U.: An Introduction to Qualitative Research. Sage (2014)
28. Creswell, J.W., David Creswell, J.: Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications (2017)
29. Curtis, B., Krasner, H., Iscoe, N.: A field study of the software design process for large systems. Commun. ACM **31**(11), 1268–1287 (1988)
30. Bjrnson, F.O., Dingsyr, T.: A survey of perceptions on knowledge management schools in agile and traditional software development environments. In: International Conference on Agile Processes and Extreme Programming in Software Engineering. Springer, Heidelberg (2009)
31. Bjrnson, F.O., Dingsyr, T.: Knowledge management in software engineering: a systematic review of studied concepts, findings and research methods used. Inf. Softw. Technol. **50**(11), 1055–1068 (2008)
32. Mohagheghi, P.: Impacts of Software Reuse and Incremental Development on the Quality of Large Systems (2004)
33. Beecham, S., et al.: Motivation in software engineering: a systematic literature review. Inf. Softw. Technol. **50**(9–10), 860–878 (2008)
34. Mugridge, R.: Managing agile project requirements with storytest-driven development. IEEE Softw. **25**(1), 68–75 (2008)
35. Engstrm, E., Runeson, P.: Software product line testing a systematic mapping study. Inf. Softw. Technol. **53**(1), 2–13 (2011)
36. Thakur, S., Singh, H.: FDRD: feature driven reuse development process model. In: 2014 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), IEEE (2014)
37. Gunasekaran, A.: Agile manufacturing: enablers and an implementation framework. Int. J. Prod. Res. **36**(5), 1223–1247 (1998)
38. Biddle, R., Martin, A., Noble, J.: No name: just notes on software reuse. ACM Sigplan Not. **38**(12), 76–96 (2003)
39. Singh, S., Chana, I.: Enabling reusability in agile software development. arXiv preprint arXiv:1210.2506 (2012)

fridaya@uef.fi

Harvesting and Informal Representation of Software Process Domain Knowledge

R.O. Oveh[1(✉)] and F.A. Egbokhare[2]

[1] Department of Mathematics and Computer Science, Western Delta University, Oghara, Delta State, Nigeria
omo_rich@yahoo.com
[2] Department of Computer Science, University of Benin, Benin City, Nigeria
fegbokhare@uniben.edu

Abstract. Knowledge Management is the process of creating, storing and managing knowledge. Knowledge harvesting being an integral part of Knowledge Management aids in eliciting, organizing and deploying explicit/tacit knowledge in a form that is easily reusable. Reusability of knowledge is one of the key benefits of Knowledge Management, as it prevents reinventing the wheel. This paper harvested and informally represented software process domain knowledge using concepts maps (an informal tool for knowledge representation). Semi-structured interview, socialization and focus group method was used for the knowledge harvesting from four different software organizations. Processes and procedures from the different phases of software development was harvested and informally represented.

Keywords: Knowledge management · Knowledge · Knowledge harvesting

1 Introduction

Knowledge harvesting is a means to draw out, express, and package tacit knowledge to help others adapt, personalize, and apply it; build organizational capacity; and preserve institutional memory [13]. Knowledge harvesting is not a catchall solution but an integral part of KM. It hinges on trust and is engendered by shared context. It cannot succeed in conflict environments, where potential knowledge contributors think they will jeopardize their status or security if they share their knowhow. However, in learning organizations, it can be leveraged judiciously to codify some human expertise in such ways that others can make use of it, for instance during staff induction or through learning and development programs, good practices, and how-to guides. Numerous benefits can flow from enabling the sharing of knowledge stocks between entities: (i) the knowledge of individuals (but also groups) is made available to those who might need it independently of human memory – thereby, it bolstering institutional memory/capacity development; (ii) a wide range of solutions to organizational issues are produced; (iii) the ability to manage change is enhanced as knowledge is packaged for easy access and understanding; (iv) the likelihood of repeated mistakes is reduced; (v) the learning curve of new personnel is shortened; (vi) precious knowledge is not lost when personnel leave; (vii) the tangible knowledge assets of the organization can

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 936–947, 2019.
https://doi.org/10.1007/978-3-030-22868-2_64

fridaya@uef.fi

be increased to create organizational value; and (viii) knowledge is communicated as guidance and support information. Knowledge harvesting therefore, elicits, organizes, and deploys explicit, tacit/implicit knowledge as key knowledge assets [8].

Due to the criticality of continuity in an organization and the importance of employee knowledge to the operations, organizations should include knowledge harvesting as part of their overall continuity management plan. Knowledge harvesting has the ability to capture employee knowledge [3, 5–7]. Software process domain knowledge entails the knowledge obtained in the production of software. Domain knowledge is a form of enabling knowledge gained by software developers during software development projects. This knowledge is utilized throughout the software development process. Software process can be defined as the set of related activities that are used in developing software. There are several software process models. They describe the sequence of activities carried out in developing software. They are a standard way of planning and organizing a software process. The major phases are requirement gathering, design & coding, implementation and maintenance.

2 Related Literature

Advocates of Knowledge Management (KM) posit that its potential is limitless and vital. Proponents of KM argue that new world economy is one in which knowledge capital, intellectual capital, learning, intangible assets, and social capital form new types of socio-economic value [10]. In essence, KM strategies can be utilized to: (a) create new knowledge and knowledge assets; (b) curb duplication of efforts and reinvention of the wheel; (c) get the right knowledge to the right people at the right time (d) enhance and leverage indigenous knowledge assets; (e) promote efficiency, effectiveness, creativity and empowerment; and f) boost intellectual capital, innovation and competitiveness [8]. Harvesting employee knowledge is not enough; it needs to be managed to be effective [4]. Knowledge management makes information available. It can refer to the process of managing knowledge as well as the product that manages knowledge, although the latter is often referred to as a knowledge management system (KMS). Both the process and the product exhibit certain features. Knowledge management collects, stores and manages the use of its content [1].

The management of knowledge and experience are key means by which systematic software development and process improvement occur. Within the domain of Software Engineering (SE), quality continues to remain an issue of concern. Knowledge Management (KM) gives organisations the opportunity to appreciate the challenges and complexities inherent in software development [14]. Software Engineering (SE) is a discipline that is yet to reach maturity, despite the tremendous amount of research it has engendered. During the 1990's, increased consideration was given to the process used for software development and its potential to improve software quality. KM in a software organization is seen as an opportunity to create a common language of understanding among software developers, so that they can interact, negotiate and share knowledge and experiences.

There is currently a gap in literature concerning the KM process for Software Engineering (SE) [14]. SE knowledge is dynamic and evolves with technology,

fridaya@uef.fi

organizational culture and the changing needs of an organization's software development practices. Author in [9] argues that software processes are essentially knowledge processes, structured within a KM framework. In [2], the author points out that software development can be improved by recognizing related knowledge content and structure, as well as appropriate knowledge and engaging in planning activities. Organizations can view KM as a risk prevention and mitigation strategy, because it explicitly addresses risks that are too often ignored, such as:

1. Loss of knowledge due to attrition.
2. Lack of knowledge and an overly long time to acquire it due to steep learning curves.
3. People repeating mistakes and performing rework because they forgot what they learned from previous projects.
4. Individuals who own key knowledge becoming unavailable.

Software engineering involves several knowledge types—technical, managerial, domain, corporate, product, and project knowledge. Knowledge can be transferred through formal training or through learning by doing. Formal training is often time consuming and expensive, and if done externally does not cover local knowledge. Learning by doing can be risky because people continue to make mistakes until they get it right. KM does not replace organized training, but supports it. Documented knowledge can provide the basis for internal training courses based on knowledge packaged as training material [12]. The author in [11] carried out a study to assess the current level of knowledge management in Software development organization in Nigeria. The knowledge maturity level during the process of Software development was computed at Level 1. It was observed that though there exists some level of knowledge management practice during software development, the knowledge maturity is at Level 1. A major characteristic of organizations at knowledge maturity Level 1 organizations is that there is no defined knowledge management culture. There is therefore the need to harvest knowledge for software development organizations, formally representing the knowledge, and storing such knowledge in a repository for reuse. This paper seeks to harvest and represent knowledge for software process.

3 Methodology

This research used semi structured interviews, socialization and focus group method to explore the views, experiences, beliefs and motivations of the domain experts. Four (4) different software organisations were used for this research. The organisations were selected because of their successes in their past and present software projects. Discussions in the form of key informant interviews were held with four (4) project managers and twelve (12) developers on the experiences and lessons learnt from past projects. Key activities that resulted in project success during the process of software development were elicited. Focus group discussion was used to capture knowledge on the specific activities carried out during software development from the key stakeholders. The interviewees did not grant permission to record the interviews electronically, so the responses were recorded on paper. Each interview session lasted for about

fridaya@uef.fi

one hour and a total of five (5) interviews were conducted over a two-week period. Follow up questions were asked via telephone conversations. The data obtained from the interviews were documented and later transcribed and meaningful knowledge for software development process was extracted using content analysis.

4 Result and Discussion

Four main activities were identified as knowledge entities in a typical software development process irrespective of the life cycle model adopted: Requirements Definition, Design & Coding, implementation and maintenance. These are core human centric activities performed by developers that create opportunities for sharing tacit knowledge during the software development process. The software project manager, Software developers and team members were the knowledge entities identified in the study. According to the software project manager, software teams are constituted at the beginning of a project and in most cases, new developers are assigned to senior developers so that knowledge can be impacted through learning and mentoring. Also, teams meet on regular and event-driven basis to discuss and brainstorm on issues relating to the software process. The experience gained during the process of software development is captured and stored within the organization. This knowledge oftentimes exists in tacit and explicit forms within software development organizations. While the explicit knowledge is documented and shared easily, tacit knowledge exists in the minds of the stakeholders in software development organizations.

The interview responses were transcribed and content analysis was used to analyze the data. Content analysis is the process of turning data obtained from interviews into findings by coding and classifying the interview data into patterns. Analysis was carried out using deductive analysis. Deductive analysis is used to test whether data are consistent with existing assumptions, theories, or hypotheses. This research used deductive analysis by first identifying keywords related to software development process and then grouping the keywords into categories related to requirements definition, coding, implementation and maintenance. The results are presented in Table 1.

Table 1. Organization of harvested knowledge

Software process	Knowledge harvested
Requirement definition	i. **Ethnographic study** of the application domain ii. **Identify** the various categories of **end users** and the tasks performed iii. Interact with and observe users (were possible) on **how tasks are performed** and the outcome iv. **Obtain** users **initial requirements** (using techniques such as interview, observation, document review, questionnaire) and establish a **communication link** (e.g. phone numbers, e-mail addresses, social media) v. **Create** a **directory of experts** from application domain

(*continued*)

fridaya@uef.fi

Table 1. (*continued*)

Software process	Knowledge harvested
Design & Coding	– Use simple modular design tools like **Data Flow Diagrams** to **Walkthrough** the **various process** for easy understanding. – For **Mathematical processes**, use simple flow charts to show program flow. – Consider **implementation/deployment platform** before **choosing a programming language** – Handle first few weeks of **Coding** as a **Retreat** – **Pair wise Code** to transfer knowledge – Create structures for **knowledge retention** within the organization – **Mentoring** through pair-wise coding – **Knowledge sharing and acquisition** through **interaction** with colleagues in team meetings – Create **repository of codes** (UR) for each problem domain that the organization has implemented software systems for – Create **Unique code** for each application domain – Deepen the **logic of the applications** by making it richer and giving it more functionality in each build – Deployment Process- e.g. Use one URL for **homogeneous projects** to cut cost, save time and ease maintenance. – Develop **generic Applications** i.e. using the same code but giving the clients the ability to select features. This reduces cost of development and maintenance. – Try to imagine what the **user** wants without necessarily going to meet him – **Theme code:** develop code for reuse and easy maintenance – **Handling Programmers Blocks** ✓Check requirements **definition and design** ✓ Go to application/problem **Domain** ✓ Discuss with **senior/experienced programmer** ✓ Go to **interactive blogs** e.g. Stack Overflow/Stack Exchange ✓ Use **physical objects** to try to establish a connection. E.g games, exercises ✓ Consult **Teams** with similar job interest ✓ **Brainstorming** ✓ **Programmer** should Switch off – Adopt a **Style** for coding – Avoid **Code Ownership** by version control – Software **development methodology**: determined by the type of project. – **maintainable code** issues by tight coupling, cohesion, proper unit testing, and proper benchmarking – Avoid **technical debt** by code review – Avoid **bus factor** issues by writing clean codes and documenting codes. Achieved through pair -wise coding

(*continued*)

fridaya@uef.fi

Table 1. (*continued*)

Software process	Knowledge harvested
Implementation	– **Phased Changeover** – **File conversion** on first build most times is for users to capture data into the database – **User training** – Prompts and popups messages to help assist users perform their tasks – Complaints are recorded – Most common model adopted similar to An **incremental software** development process ✓ Build initial model ✓ Connect users to a **complaint/requests repository** ✓ Build next increment based on users request
Maintenance	– Allow **customers** take the lead to add extra functionality through the documented complaints and requests collected – Create **URL for codes** from different application domain developed – Maintain from **remote sites**. Once an error is sited, the primary copy is corrected and propagated to all other copies – Additional features (**new builds**) are added to existing ones based on users' **documented complaints** and requests which form the requirements for the new increment. The **new increment** usually interfaced into the existing system and all sites are informed and useful information on how to access the new feature are provided.

5 Knowledge Representation

The harvested knowledge in Table 1 was informally represented in Figs. 1, 2, 3, 4, 5, 6, and 7 using concept maps. A concept map is an informal tool for knowledge representation.

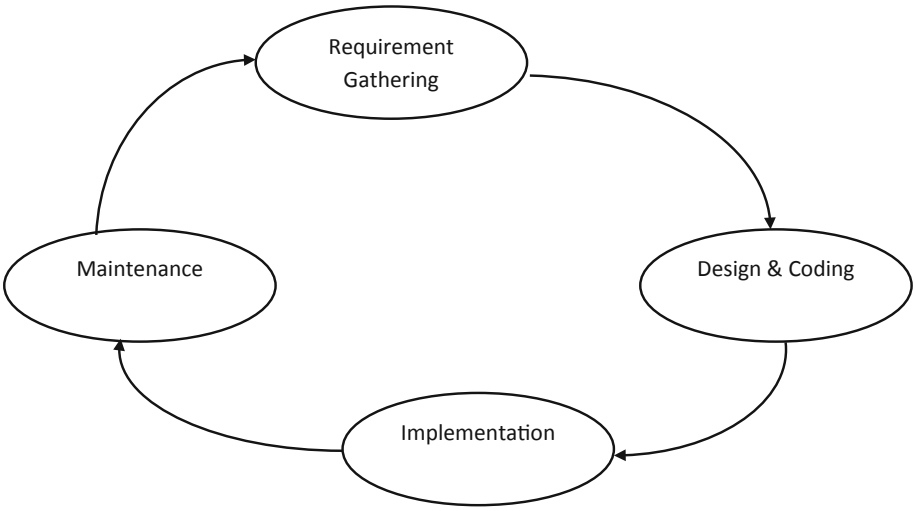


Fig. 1. A high level view of the harvested knowledge

fridaya@uef.fi

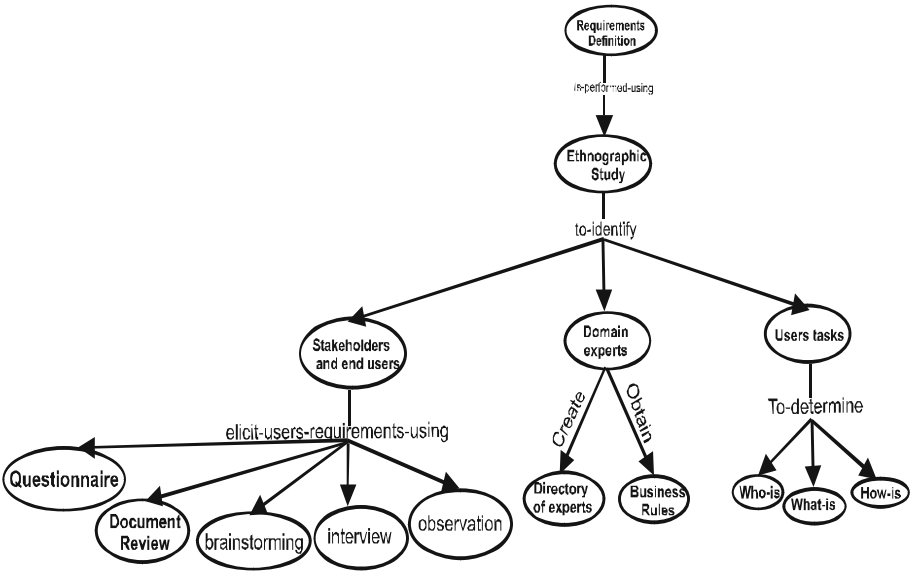


Fig. 2. Requirement definition concept map

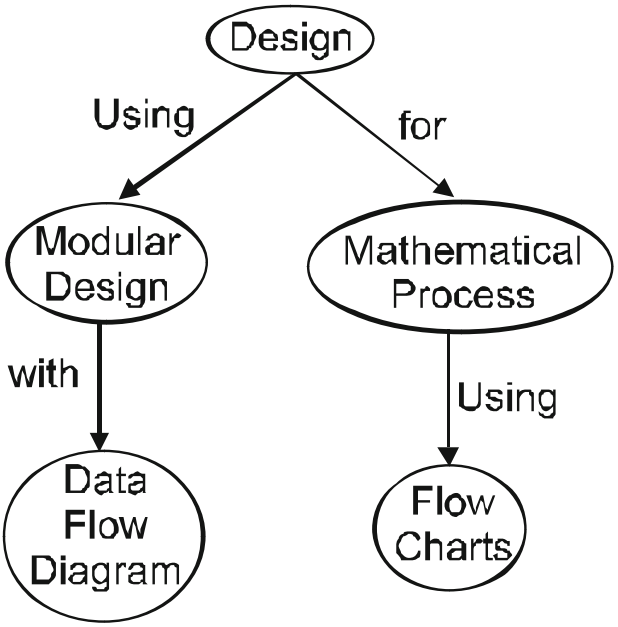


Fig. 3. Design concept map

fridaya@uef.fi

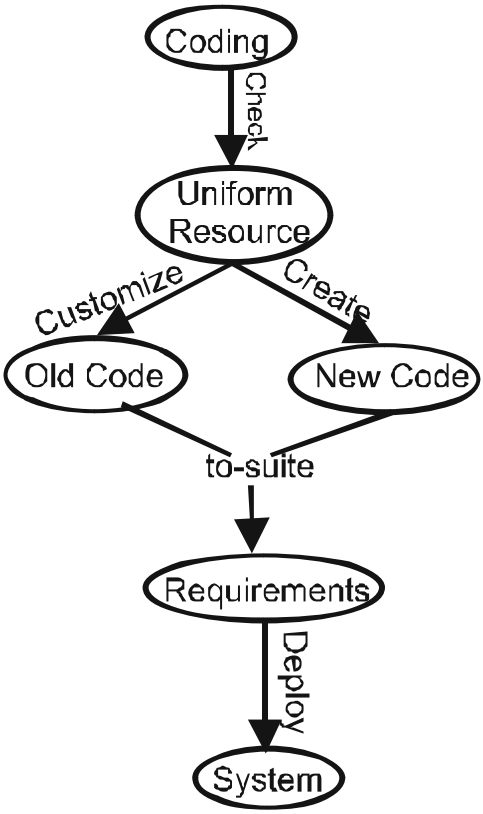


Fig. 4. A concept map showing Coding

6 Discussion

Figure 1 shows a map that depicts the various knowledge entities identified from the case study where knowledge was harvested. For each knowledge entity, the harvested activities are further refined and the relationships between the various entities are represented by the directed edges on the maps in Figs. 2, 3, 4, 5, 6, and 7. As shown in Fig. 2, the main method of knowledge acquisition during requirements definition is through ethnographic study to identify stakeholders/end users, domain experts and users tasks. Intrinsic knowledge relating to user's tasks, business rules and the organizational culture are harvested from: Informal interactions, observation, documents review, brainstorming, interview and questionnaire. Domain experts who can be contacted for clarification and resolution of issues relating to the application in view during the software development process are identified during ethnographic study, and their information is used to create/update the directory of experts. Figure 3 shows that for modular design Data Flow Diagram (DFD), which is used to walkthrough the various process for easy understanding. While for mathematical process, simple flow charts are used to show the program flow. As shown in Fig. 4, the uniform resource

fridaya@uef.fi

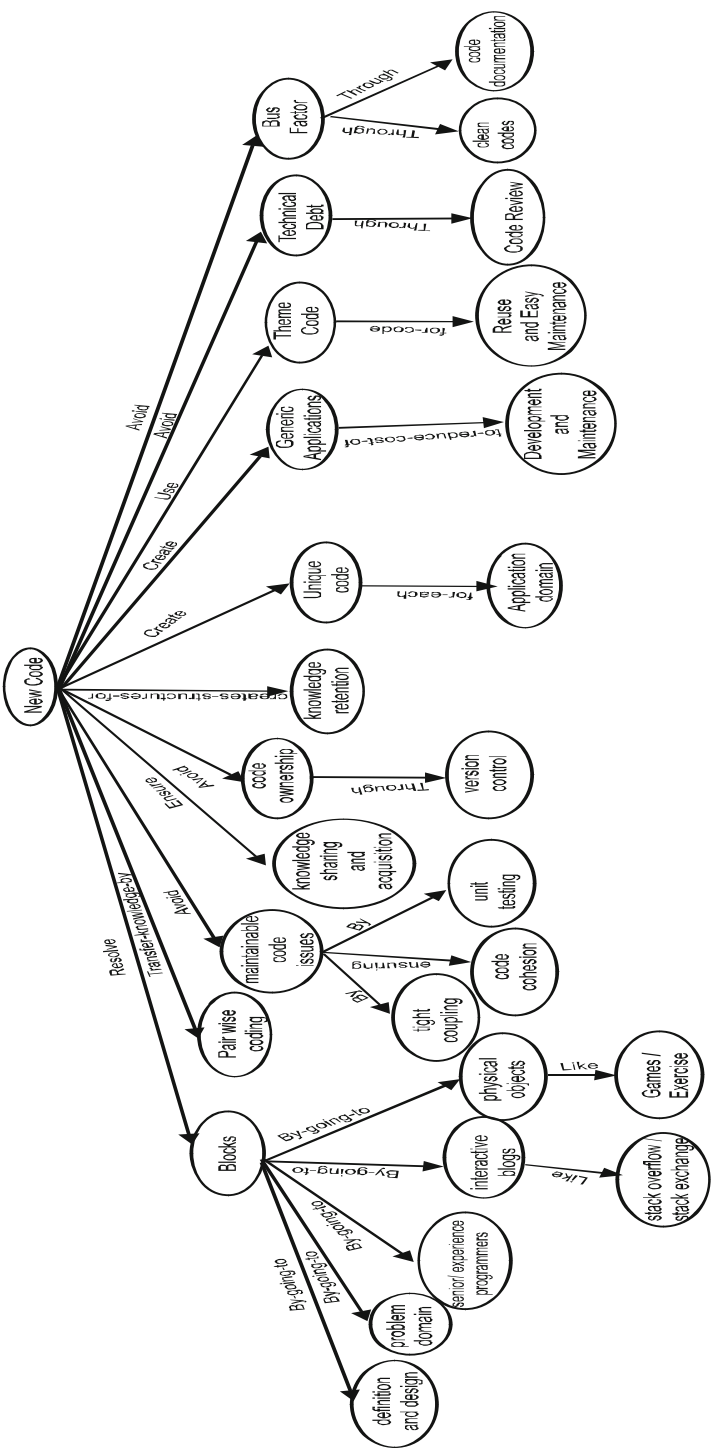


Fig. 5. Creating a new code concept map

fridaya@uef.fi

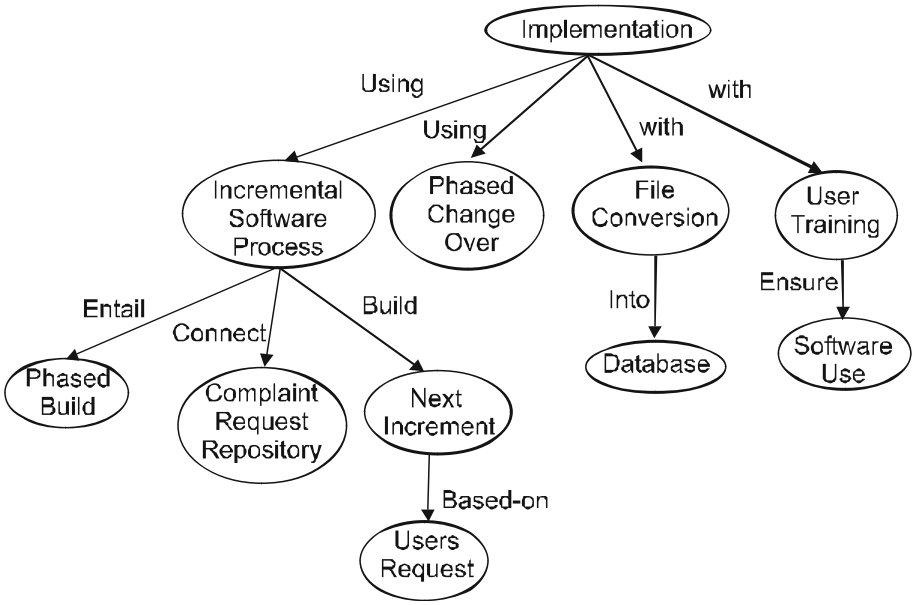


Fig. 6. Implementation concept map

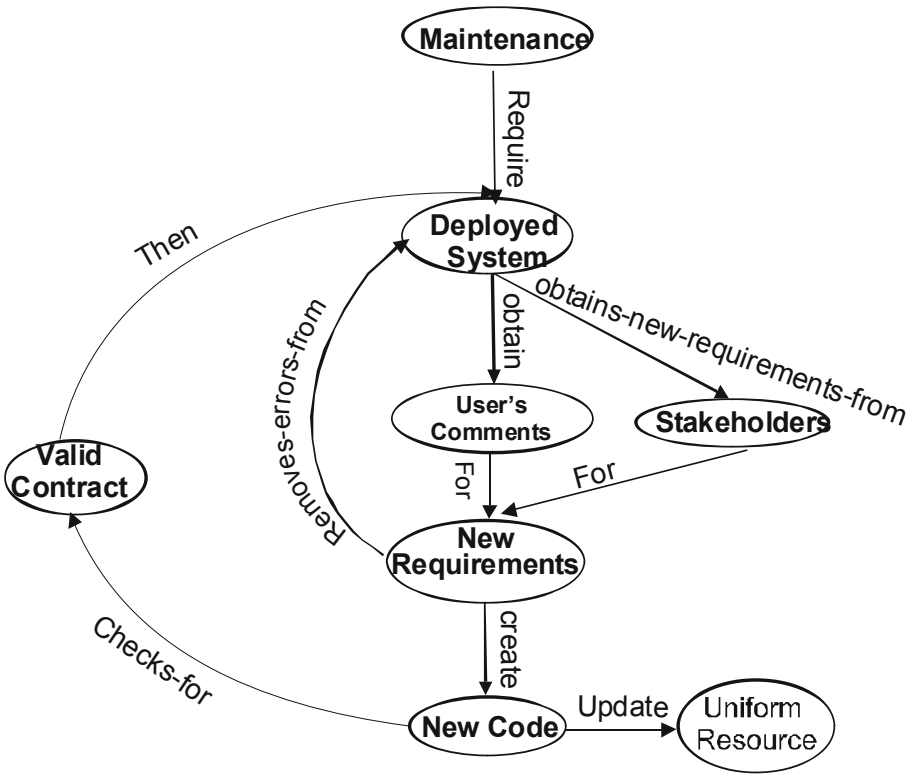


Fig. 7. Maintenance concept map

fridaya@uef.fi

(i.e. repository) is first checked for existing code, in other not to reinvent the wheel. If there is an existing code, it is customized to suite the requirements and then deployed as a system. If there is no existing code in the repository, then a new code is created based on elicited requirements and then deployed as a system. Figure 5 shows things obtainable during writing a new code. It can be broadly grouped into the do's and don'ts of writing a new code. The do's include: pairwise coding for knowledge retention, interaction with colleagues for knowledge sharing and acquisition, creating structures for knowledge retention, creating generic applications to reduce cost of development and maintenance, theme code for code reuse and easy maintenance, creating unique code for each application domain, and resolving blocks by going to: definition and design, problem domain, senior/experience programmers, interactive blogs (like stack-overflow or stack-exchange), physical exercise (like games). The don'ts include avoiding: maintainable code issues (by tight coupling, code cohesion, and unit testing), code ownership through version control, technical debt through code review, and bus factor issues through clean codes and code documentation. Figure 6 shows the implementation concept map which should be done using incremental software process. It entails phased build, connects complaint requests repository for maintenance, and for building the next software based on users request. Implementation should use phased change over. It is also done with file conversion, in order to capture data to the database. Implementation should be done with user training to encourage them to use the system. In Fig. 7, maintenance begins when the system is deployed. It entails removing errors from operational system, obtaining user's comments, and obtaining new requirements from stakeholders with the aim of creating new system by writing new code. A new code/fresh system is only produced if there is a valid contract.

7 Conclusion

Knowledge is an invaluable resource that could give an organization competitive advantage. Knowledge management involves creating, storing and reusing knowledge. The first step to managing knowledge is to harvest knowledge. Having identified software organization to be at Level 1, there was the need to harvest and represent their process activities. Using case study methodology, knowledge considered critical for software development process was harvested from the knowledge entities identified. The harvested knowledge was represented using concept maps. It is further recommended that the harvested knowledge be formally represented and evaluated for reuse.

References

1. Alavi, M., Leidner, D.E.: Knowledge management and knowledge management systems: conceptual foundations and research issues. MIS Q. **25**(1), 107–136 (2001)
2. Aurum, A., Jeffery, R., Wohlin, C., Handzic, M.: Managing Software Engineering Knowledge. Springer, Germany (2003)

fridaya@uef.fi

3. Beazley, H., Boenisch, J., Harden, D.: Continuity Management: Preserving Corporate Knowledge and Productivity When Employees Leave. Wiley, Hoboken (2002)
4. Brands, R. F.: Intelligence lost: The boomers are exiting. Bloomberg Businessweek (2011). www.businessweek.com/management/intelligence-lost-the-boomers-areexiting-08052011.html. Accessed 8 Sept 2018
5. DeLong, D.W.: Lost Knowledge: Confronting the Threat of an Aging Workforce. Oxford University Press, New York (2004)
6. Eisenhart, M.: Gathering knowledge while it's ripe. Knowl. Manag. **4**, 48–53 (2001)
7. Field, A.: When employees leave the company, how can you make sure that their expertise doesn't? Harvard Manag. Commun. Lett. **6**(4), 3–5 (2003)
8. Hanson, T.K., Kararach, G.: The Challenges of Knowledge Harvesting and the Promotion of Sustainable Development for the Achievement of the MDGS in Africa. The African Capacity Building Foundation (ACBF) Harare, Zimbabwe (2011)
9. Kess, P., Haapasalo, H.: Knowledge creation through a project review process in software production. Int. J. Prod. Econ. **80**(1), 49–55 (2002)
10. Malik, K.P., Malik, S.: Value creation role of knowledge management: a developing country perspective. Electron. J. Knowl. Manag. **6**(1), 41–48 (2008)
11. Oveh, R.O., Egbokhare, F.: Preliminary studies on assessment of the level of knowledge management in Nigerian software development organisations. Niger. J. Educ. Health Technol. Res. (NJEHETR) **10**, 93–99 (2018)
12. Rus, I., Lindvall, M.: Knowledge management in software engineering. IEEE Softw. **19**(3), 26–38 (2002)
13. Serrat, O.: Harvesting Knowledge, vol. 81, pp. 1–5. Asian Development Bank, Washington, DC (2010). http://digitalcommons.ilr.cornell.edu/intl. Accessed September 2018
14. Ward, J., Aurum, A.: Knowledge management in software engineering - describing the process. In: Proceedings of the 2004 Australian Software Engineering Conference (ASWEC'04) (2004)

fridaya@uef.fi

A Methodology for Performing Meta-analyses of Developers Attitudes Towards Programming Practices

Thomas Butler(✉)

University of Northampton, Northampton, UK
thomas.butler@northampton.ac.uk

Abstract. Programming practices are often labelled as "best practice" and "bad practice" by developers. This label can be subjective but we can see trends among developers. A methodology for performing meta-analyses of articles discussing any given practice was created to determine programmers overall attitudes towards any given practice while accounting for factors such as whether they considered alternative approaches.

1 Introduction

Programming practices can often be described as *bad practice* or *best practice* depending on whether they have a positive or negative impact on the maintainability of the code in which they are used [1].

A systematic review of the academic literature on "bad practices", "code smells" and "anti-patterns" by Sabir *et al.* [2] identified 56 bad practices that had been recognised in academia. Meanwhile, industry code reviewers have identified practices such as the singleton [3–12] which negatively impact code quality.

Where the singleton pattern is mentioned in academia it is only discussed as having been utilised while developing software rather than discussing whether it should or should not have been used [13, 14].

The singleton pattern does not appear in the systematic review of academic literature despite its widespread [15] recognition as a bad practice among industry developers. There is currently a gap in the literature between industry and academic works when defining programming practices which have negative effects on maintainability.

It is hypothesised that the gap between industry programmers and academics is that industry developers spend every day working on large projects where they are likely to encounter problems that academics focusing mostly on theory will not. Industry experts tend to work on large software projects which require constant maintenance and enhancement for years or even decades. They are able to determine which practices prevent them performing maintenance efficiently.

The practices most commonly cited by articles reviewed by Sabir *et al.* [2] originally come from a book [16] written by industry developers. As *Refactoring: Improving the Design of Existing Code (Object Technology Series)* [16] is two decades old, industry developers have discovered more bad practices while academic works have tended to

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 948–974, 2019.
https://doi.org/10.1007/978-3-030-22868-2_65

fridaya@uef.fi

focus on the bad practices identified back in 1999. This research is a step towards bringing academic literature back in-line with industry.

One problem with research in this area of research is that it can be seen as subjective. An article can be written discussing the negative aspects of a programming practice and on its own this can be viewed as an opinion. However, if developers independently reach the same conclusions then there is a consensus.

This research presents a methodology for performing a meta-analysis of industry developer opinions about a practice to determine if there is a consensus about the programming practice being studied. This will give academics a reproducible methodology for determining the overall opinion about the merits of any given programming practice.

For the purpose of code review and software quality, the analytic rigour of an article is important when calculating developer consensus. Sources will vary in their level of detail, a manual page will demonstrate how to use a practice but will not discuss when or if the practice should be used. Opinion pieces may go into significantly more detail with discussions about pros and cons of the practice, where it's applicable and alternative approaches that can be used to solve the same problem.

A scoring system will be created to grade articles on their analytic rigour. Once articles are graded, it will be possible to compare articles based on their analytic rigour and then perform a meta-analysis of opinions on a specific programming practice which can be used to observe trends such as "If an author analyses the practice in depth they are more likely to recommend using alternative approaches than authors who do not have as robust methodology."

2 Aims and Objectives

1. Create a scoring system which can be used to:
   1. Grade the analytic rigour of an article/book/paper discussing a particular programming practice.
   2. Compare the analytic rigour of different articles for the purposes of meta-analysis.
   3. Compare the overall quality of discussions about a specific programming practice.
2. Calculating the score should not require reading the article in detail to calculate the score and anything used to calculate the score should be a binary choice.
3. With a scoring system in place, perform proof-of-concept meta-analyses on practices which are well known to be described as "good" and "bad" to demonstrate that the meta-analysis framework is fit for purpose.

fridaya@uef.fi

3 Methodology

3.1 Metric for Comparing Analytical Rigour in Programming Articles

Differing methodological rigor in sources is a problem which exits exists when doing any kind of meta-analysis. When performing meta-analysis of clinical trials the Cochrane Collaboration [17] consider methodological rigour an important part of their meta-analysis.

Rather than simply counting the number of trials which show a positive outcome and counting the number of trials which show a negative outcome, they weigh the trials based on methodological rigour. In a meta-analysis of a particular drug it may be found that three trials show that it is an effective treatment while eight show that it is not. Instead of simply counting the numbers on each side, the analytic rigor of each study is considered and used as factor when drawing conclusions about the efficacy of the treatment.

In a meta-analysis of the efficacy of homeopathic treatments [18] it was found that trials of homeopathy with a poor methodology were much more likely to show a positive outcome whereas trials with a robust methodology were much more likely to conclude that homeopathy is no better than placebo.

Methodological rigour can affect the outcome. For example, by placing the most healthy patients in the experimental group and placing the least healthy patients in the control group it's likely that the experimental group will see significant improvement over the control group regardless of whether the drug being tested has any effect [19].

For programming articles, analytic rigour can be plotted against whether the article recommends using or avoiding the practice to create a meta-analysis in a similar manner.

It should be possible to draw conclusions such as *as an article's analytic rigour increases, it is more likely to recommend using the practice in question.*

The created metric was based on the Jadad Scale [20] used for analysis of clinical trials in medicine. The Jadad Scale is a 5 point scale using a 3 question questionnaire which can be used to quickly assess the methodological rigour used in a clinical trial. The questions asked are: *Was the study described as randomized? Was the study described as double blind? and Was there a description of withdrawals and dropouts?*. These are then used to generate a score from zero (very poor) to five (rigorous). By citation count the Jadad Scale is the most widely used method of comparing clinical trials in the world [21].

A new metric was created based on the principles of the Jadad scale to be used in determining the analytic rigour of any article about a programming practice. A seven point scale was chosen with a point awarded if the article does each of the following:

1. Describes how to use the practice
2. Provides a code example of using the practice
3. Discusses potential negative/positive implications of using the practice
4. Describes alternative approaches to the same problem
5. Provides like for like code samples comparing the practice to alternative approaches
6. Discusses of pros/cons of the compared approaches
7. Offers a conclusion on when/where/if the practice is suitable.

fridaya@uef.fi

Using this metric, a manual page that describes a practice and provides a sample of how to use it would score two whereas an article that discussed the pros/cons of different approaches and made a recommendation would score seven.

3.2 Meta-analysis

Clinical trials could be separated by their Jadad score but this alone tells us nothing about the efficacy of the treatment. To produce a conclusion the Jadad Score is plotted against trial outcome.

Programming articles do not produce a result but they can offer a recommendation to use or avoid the practice being discussed. A manual page will not make a recommendation but an opinion piece will discuss if and when the practice being described should be used.

A five point scale was used to model the recommendation made by an article:

1. Always favour this practice over alternatives.
2. Favour this practice over alternatives unless specific (defined*) circumstances apply.
3. Neutral - No recommendation (e.g. a manual page) or no conclusion drawn.
4. Only use this practice in specific (defined*) circumstances.
5. Always favour alternative approaches.

A five point scale was chosen over a three point scale as there may be cases where an article is concluded with a discussion of trade-offs such as where an approach may be faster but less flexible. This meta-analysis will focus on flexibility. If a conclusion is drawn that a practice should be used when flexibility is preferred over performance (or any other consideration) then the article would be awarded a score of 2.

Research focused on factors other than flexibility could use a different focus for the recommendation score.

* For scores 2 and 4, the specific circumstances have to be described rather than alluded to.

For example Buss [22] writes:

> *When designing a system, it's important to pick the right design principle for your model. In many circumstances, it makes sense to prefer composition over inheritance.*

Buss [22] only alludes to when using inheritance is preferable and only provides examples where composition is preferred. In this case the article is given a 5 despite the conclusion saying "many circumstances" rather than "all circumstances".

Ericson [23] says:

> *If you aren't sure if a class should inherit from another class ask yourself if you can substitute the child class type for the parent class type. For example, if you have a Book class and it has a subclass of Comic Book does that make sense? Is a comic book a kind of book? Yes, a comic book is a kind of book so inheritance makes sense. If it doesn't make sense use association or the has-a relationship instead.*

In this instance, the Ericson [23] clearly states a situation where inheritance should be used over composition so would be given a recommendation score of 4.

fridaya@uef.fi

3.3 Collecting Data

Non-academic Sources
As this research is intended to provide a methodology for brining industry experience into academic work, the sources used are developers working in industry. Google was used to find articles about a chosen programming practice. A search returns any articles discussing the practice regardless of whether they are for or against this also acted as a randomisation tool.

Each article was then given a *Jadad* style score from 0–7 and a score for its recommendation.

3.3.1 Additional Considerations
There are several practical issues with collecting data in this manner:

1. To minimise the effect of Google giving user-specific results based on previous searches, results were collected while logged out and using the browser's private browsing mode and closing the browser between each search term.
2. Search results will not be truly random due to the way Google's algorithm works and results will be sorted by *relevance* and the way Google sorts the results may have implicit bias: The most popular links and most cited links will appear first. Although not truly random, this gives a better overview of the zeitgeist than a genuinely randomised sample by putting the most read/cited articles ahead of less read/cited pages. Articles which are widely shared and linked to will be more likely to appear in the first 100 results.
3. A practice may have more than one common name. When this is the case, each name will be searched for and 100 results collected in total. If a practice is known by 4 different names, the first 25 relevant results for each practice were used. If a result lists both names it will only be counted once.
4. Other search engines may yield different results. Google was chosen because of its dominance and likelihood to have indexed more results. Using a search engine such as Qwant [24] which does not offer personalised results would make the results easier to replicate but may not offer as comprehensive results. Regardless of which search engine is used, results will change over time.

Further research is required to determine the extent of which these factors may affect results.

Regardless of these factors, results should be indicative of developers' attitudes towards the programming practice being analysed.

3.4 Test Methodology

To verify that the suggested meta-analysis methodology produces meaningful results, a meta-analysis was performed on two practices where the result can be anticipated with a high degree of certainty. If the methodology works as intended, the following hypotheses should be proven true.

fridaya@uef.fi

3.4.1 Singleton Pattern

The singleton pattern is well known as being considered bad practice among developers [15] and will act as a good benchmark for testing the meta-analysis methodology.

3.4.1.1 Hypothesis

Before the results were collected it was expected that articles which had a higher *Jadad style score* (higher academic rigour) would be more likely to suggest avoiding the practice.

3.4.2 Dependency Injection

Dependency Injection is antithesis to the Singleton Pattern and is much more flexible. Although there are some practical considerations when using Dependency Injection and there is widespread discussion about the best way to implement it, it's widely considered the best approach for flexibility [25].

3.4.2.1 Hypothesis

Dependency Injection is a well-established method of increasing flexibility in code [26]. Because of this, it was expected that there would be few to no negative recommendations and as the *Jadad* style score increases articles should be more likely to suggest favouring dependency injection over alternative approaches.

4 Results

4.1 Singleton

Figure 1 displays the results for the Singleton Pattern. Each line represents an article and the left (orange) bar for each article is the recommendation going from 5: Avoid this practice at all costs (Far left) to 1: Favour this practice over alternatives.

The right (blue) bar for each article is the Jadad style score measuring analytic rigour. A score of seven means the article describes the practice, provide code examples, discusses alternative approaches, provides like-for-like code samples, discusses the pros/cons of each approach and makes a recommendation of which approach should be used.

Article 1 has a recommendation score of 3 and a Jadad style score of 1. It does not go into detail and its recommendation is neutral; it doesn't suggest either avoiding or favouring use of the Singleton Pattern.

Article 99 strongly recommends against using the Singleton Pattern and has an *Jadad* style score of 7, it compares the singleton against alternatives in detail and concludes by strongly recommending against its use (recommendation score of 5).

Raw data is available as Appendix 1.

fridaya@uef.fi

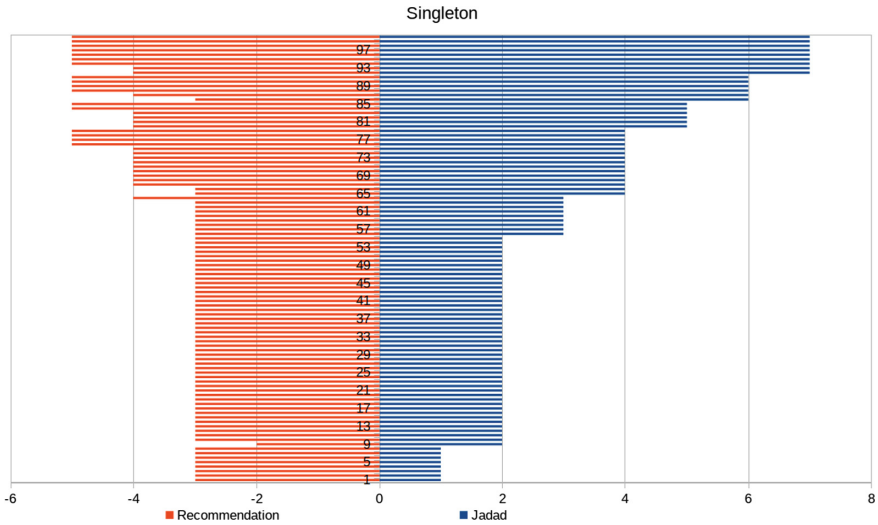


Fig. 1. Singleton – Results

As hypothesised, articles with a high analytic rigour are considerably more like to suggest avoiding the singleton pattern.

Table 1 shows a breakdown of recommendation scores for the articles reviewed. If a simple tally were used, the singleton pattern would appear to have a mostly neutral recommendation score. 65% of articles do not recommend for or against its use.

Table 1. Singleton pattern recommendation score

Recommendation	Number of articles making recommendation
1: Always favour this practice over alternatives	0
2: Favour this practice over alternaitves except in specific circumstances	1
3: Neutral/no recommendation	65
4: Favour alternative approaches except in specific circumstances	16
5: Always favour alternative approaches	18

4.1.1 Key Findings - Singleton Pattern

- The mode recommendation is neutral. If a developer looked through articles about the singleton pattern, 65% of the articles they read would not recommend against using the Singleton Pattern.

fridaya@uef.fi

- The mean recommendation score is 3.5. From this alone it could be inferred that the singleton pattern is generally considered to be neutral, slightly discouraged but not widely avoided.
- When the *Jadad* style score is taken into account, every article which makes a recommendation recommends against using the singleton pattern (recommendation score of 4 or 5).
- Only 22% of articles about the singleton pattern mention alternative approaches that can be used to solve the same problem.
- Of those that recommend against using the pattern, over half say it should be avoided at all cost.
- 55 of the 65 articles which make a neutral recommendation are manual type pages (Jadad style score of 2) which show how to use the pattern but do not weigh in on when, where or if it should be used and do not compare the pattern to alternatives.
- No articles which make a recommendation recommend using the singleton pattern instead of alternative approaches.

4.2 Dependency Injection

Figure 2 shows the results for performing a meta-analysis of Dependency Injection. As hypothesised, Dependency Injection is seen as overwhelmingly positive with zero articles discouraging its general use.

Raw data is available in Appendix 2.

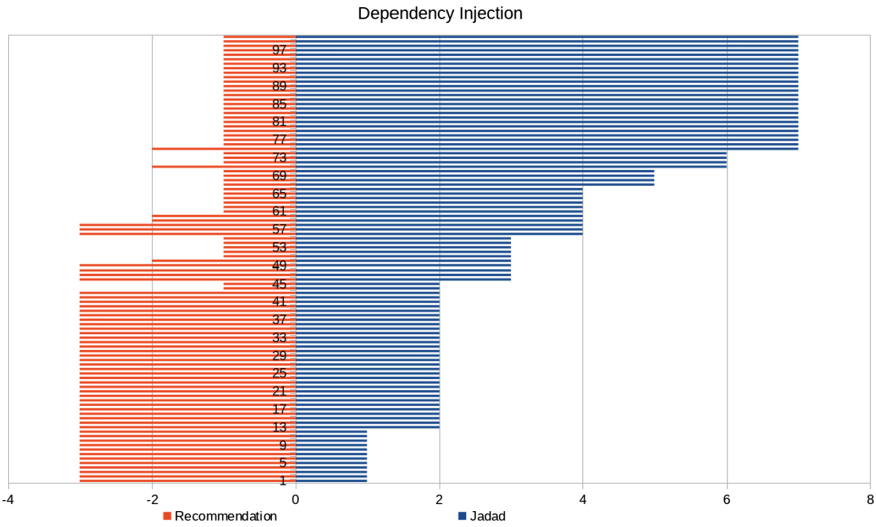


Fig. 2. Dependency Injection - Results

fridaya@uef.fi

Table 2. Dependency Injection recommendation score

Recommendation	Number of articles making recommendation
1: Always favour this practice over alternatives	50
2: Favour this practice over alternaitves except in specific circumstances	5
3: Neutral/no recommedation	45
4: Favour alternative approaches except in specific circumstances	0
5: Always favour alternative approaches	0

Table 2 shows a breakdown of recommendation scores for the articles reviewed. In a tally that doesn't account for analytic rigour 50% of articles would suggest using Dependency Injection in place of alternative approaches.

4.2.1 Key Findings - Dependency Injection

- The mean score is 1.94 which shows that even using a simple tally; the overall recommendation is that Dependency Injection is a favourable pattern among developers.
- Every article with an analytic rigour score of 4 or higher recommends using this practice instead of alternative.
- When the *Jadad* style score is taken into consideration, 47 of the 50 articles with a neutral recommendation are manual style pages which show how the pattern is used but do not discuss when, where or if it should be used.
- Discounting the manual pages, only two of the remaining 53 articles make a neutral recommendation and both of those have a *Jadad* style score 3.
- As the *Jadad* style score increases, the probability that an article will recommend using Dependency Injection over alternatives increases.
- Only 5 of the 55 articles in favour of dependency injection suggest there some specific circumstances where alternatives should be used instead (*Jadad* style score of two).

5 Conclusion

By testing the methodology with practices for which the outcome can be predicted it was possible to validate this meta-analysis methodology.

The methodology produced the expected result. It has been shown that if an author considered alternative approaches they were more likely to recommend against using the Singleton Pattern. The inverse was also true for Dependency Injection.

fridaya@uef.fi

As these were the expected results, the methodology used can be shown to work as intended and provide an overview of the attitudes of developers about any given practice for future meta-analyses of practices for which the result cannot be predicted.

This meta-analysis methodology gives more insight into the overall opinion of programming practices than a simple tally of whether developers encourage or discourage use of a programming practice.

5.1 Additional Findings

1. Although a small sample size of two practices were tested, in both cases roughly half of articles reviewed do not make a recommendation on when or where the practice should be used. For the singleton pattern only 45% of reviewed articles discussed whether the pattern should be used or avoided.
2. Any developer looking for information on a practice will find more information about *how* to use a practice than *when* or *where* the practice is applicable.

5.2 Problems Encountered

Data collection using Google became increasingly difficult after around 80 relevant results. The number of irrelevant articles appearing in search results begin to heavily outweigh the relevant articles and there was a significant issue with duplicated content. Articles had been posted on multiple websites, often without dates or author names, making it difficult to keep track of which articles had already been included in the meta-analysis.

Since Dependency Injection and the Singleton pattern are both widely known and discussed programming practices, finding 100 unique relevant results for lesser known practices may be difficult.

5.3 Future Research

This research could be continued by running the same meta-analysis on different search engines and comparing the results or looking into trends over time using article dates. For example, it may be observed that a practice is seen favourably in articles published in 1990s–2000s and then less favourably as time progresses.

This methodology could be abstracted to and used for a meta-analysis of any widely discussed topic by defining the scales for analytic rigour and recommendation.

fridaya@uef.fi

Appendix 1. Raw Data: Singleton

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
http://www.fssnip.net/7p/title/Singleton-Pattern		1						3	1
https://www.hackerrank.com/challenges/java-singleton/forum		1						3	1
https://gist.github.com/mssola/6138155		1						3	1
http://fortranwiki.org/fortran/show/Singleton+pattern	1							3	1
http://www.adam-bien.com/roller/abien/entry/singleton_pattern_in_es6_and		1						3	1
https://coderwall.com/p/iemfbg/objective-c-singleton-pattern-with-arc		1						3	1
http://cruise.eecs.uottawa.ca/umple/SingletonPattern.html		1						3	1
http://www.netobjectives.com/resources/books/design-patterns-explained/java-code-examples/chapter21		1						3	1
https://www.dotnetperls.com/singleton-static		1					1	2	2
https://www.tutorialspoint.com/design_pattern/singleton_pattern.htm	1	1						3	2
https://msdn.microsoft.com/en-gb/library/ff650316.aspx	1	1						3	2
https://www.javaworld.com/article/2073352/core-java/simply-singleton.html	1	1						3	2
http://www.oodesign.com/singleton-pattern.html	1	1						3	2
https://code.tutsplus.com/tutorials/android-design-patterns-the-singleton-pattern–cms-29153	1	1						3	2
http://www.dofactory.com/javascript/singleton-design-pattern	1	1						3	2

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://www.techopedia.com/definition/15830/singleton	1	1						3	2
https://www.geeksforgeeks.org/singleton-design-pattern/	1	1						3	2
https://dzone.com/articles/singleton-pattern-a-deep-dive	1	1						3	2
https://www.javatpoint.com/singleton-design-pattern-in-java	1	1						3	2
https://developer.salesforce.com/page/Apex_Design_Patterns_-_Singleton	1	1						3	2
https://fullstack-developer.academy/singleton-pattern-in-typescript/	1	1						3	2
http://jargon.js.org/_glossary/SINGLETON_PATTERN.md	1	1						3	2
https://dalibornasevic.com/posts/9-ruby-singleton-pattern	1	1						3	2
http://marcio.io/2015/07/singleton-pattern-in-go/	1	1						3	2
https://basarat.gitbooks.io/typescript/docs/tips/singleton.html	1	1						3	2
http://www.vogella.com/tutorials/DesignPatternSingleton/article.html	1	1						3	2
https://www.nada.kth.se/kurser/kth/2D1359/01-02/contents/forelasningar/lecture10.ppt	1	1						3	2
https://www.linkedin.com/pulse/singleton-pattern-eager-lazy-enum-ramasamy-kasiviswanathan	1	1						3	2
https://www.avajava.com/tutorials/lessons/singleton-pattern.html	1	1						3	2
ftp.mak.com/out/classdocs/vrforces4.4.1/classref/vrv_the_rooted_singleton_pattern.html	1	1						3	2

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
http://www.galloway.me.uk/tutorials/singleton-classes/	1	1						3	2
https://wiki.base22.com/btg/singleton-pattern-3459.html	1	1						3	2
https://coffeescript-cookbook.github.io/chapters/design_patterns/singleton	1	1						3	2
https://blogs.sap.com/2012/04/06/singleton-pattern-in-abap/	1	1						3	2
http://docs.ros.org/indigo/api/typelib/html/group__singleton.html	1	1						3	2
https://learnswiftwithbob.com/course/object-oriented-swift/singleton-pattern.html	1	1						3	2
https://www.visual-paradigm.com/tutorials/singletonpattern.jsp	1	1						3	2
https://medium.freecodecamp.org/lets-talk-about-you-and-the-singleton-design-pattern-bb2e160fa952	1	1						3	2
https://www.htmlgoodies.com/beyond/javascript/implementing-the-singleton-design-pattern-in-javascript.html	1	1						3	2
http://web.science.mq.edu.au/~mattr/courses/object_oriented_development_practices/6/notes.html	1	1						3	2
http://www.jot.fm/issues/issue_2007_03/column2/	1	1						3	2
https://itexico.com/blog/bid/99247/Software-Development-The-Singleton-Design-Pattern-and-other-Creational-Patterns	1	1						3	2
http://code.activestate.com/recipes/52558-the-singleton-pattern-implemented-with-python/	1	1						3	2

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://alvinalexander.com/scala/how-to-implement-singleton-pattern-in-scala-with-object	1	1						3	2
http://wiki.unity3d.com/index.php/Singleton	1	1						3	2
https://locklessinc.com/articles/singleton_pattern/	1	1						3	2
http://www.tothenew.com/blog/singleton-pattern-with-javascript/	1	1						3	2
https://sweetcode.io/singleton-design-pattern-using-java/	1	1						3	2
http://elbenshira.com/blog/singleton-pattern-in-python/	1	1						3	2
https://php.earth/docs/php/ref/oop/design-patterns/singleton	1	1						3	2
https://www.ibm.com/developerworks/library/j-dcl/	1	1						3	2
https://amir.rachum.com/blog/2012/04/26/implementing-the-singleton-pattern-in-python/	1	1						3	2
https://krakendev.io/blog/the-right-way-to-write-a-singleton	1	1						3	2
http://moddb.wikia.com/wiki/Singleton_Pattern	1	1						3	2
http://csharpindepth.com/Articles/General/Singleton.aspx	1	1	1					3	3
https://www.journaldev.com/1377/java-singleton-design-pattern-best-practices-examples	1	1	1					3	3
http://best-practice-software-engineering.ifs.tuwien.ac.at/patterns/singleton.html	1	1	1					3	3
https://refactoring.guru/design-patterns/singleton	1	1		1				3	3

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://www.techrepublic.com/blog/software-engineer/using-the-singleton-pattern-in-java/	1	1	1					3	3
https://www.gamasutra.com/blogs/MattChristian/20101013/88205/OOPsie_Patterns_The_Singleton_Pattern.php	1	1	1					3	3
https://pdfs.semanticscholar.org/presentation/f1dc/667b86aac3f5db63c26a67d30d586d2244a6.pdf	1	1	1					3	3
https://www.implementingquantlib.com/2017/09/odds-and-ends-singleton.html	1	1	1					3	3
https://8thlight.com/blog/josh-cheek/2012/10/20/implementing-and-testing-the-singleton-pattern-in-ruby.html	1	1	1					4	3
http://csc.columbusstate.edu/woolbright/java/singleton.html	1	1	1	1				3	4
https://www.codeproject.com/Articles/307233/Singleton-Pattern-Positive-and-Negative-Aspects	1	1	1			1		3	4
https://sourcemaking.com/design_patterns/singleton	1	1	1				1	4	4
https://www.gofpatterns.com/design-patterns/module3/consequences-effects-singleton-pattern.php	1	1	1				1	4	4
http://wiki.c2.com/?SingletonPattern	1	1	1				1	4	4
https://medium.com/if-let-swift-programming/the-swift-singleton-pattern-442124479b19	1	1	1				1	4	4
http://2ality.com/2011/04/singleton-pattern-in-javascript-not.html	1	1	1				1	4	4
https://ieftimov.com/singleton-pattern	1	1	1				1	4	4

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://addyosmani.com/resources/essentialjsdesignpatterns/book/	1	1				1	1	4	4
https://phpenthusiast.com/blog/the-singleton-design-pattern-in-php	1	1	1				1	4	4
https://www.perl.com/article/52/2013/12/11/Implementing-the-singleton-pattern-in-Perl/	1	1	1				1	4	4
http://www.bogotobogo.com/DesignPatterns/singleton.php	1	1		1			1	5	4
https://www.infoworld.com/article/3112025/application-development/design-patterns-that-i-often-avoid-singleton.html			1	1		1	1	5	4
https://codeburst.io/design-patterns-for-modern-web-development-singletons-bf7bc06bd17d	1		1	1		1		5	4
https://anasshekhamis.com/2017/07/27/the-singleton-design-pattern-es5-and-es2015/	1	1	1				1	5	4
https://tommcfarlin.com/singleton-design-pattern-1/	1	1	1	1			1	4	5
https://php.earth/docs/php/ref/oop/design-patterns/singleton	1	1	1	1			1	4	5
https://torquemag.io/2016/11/singletons-wordpress-good-evil/	1	1	1	1			1	4	5
http://adamschepis.com/2011/05/02/im-adam-and-im-a-recovering-singleton-addict.html	1		1	1		1	1	4	5
http://www.phptherightway.com/pages/Design-Patterns.html	1		1	1		1	1	5	5
http://beust.com/weblog/2011/03/10/rehabilitating-the-singleton-pattern/	1		1	1		1	1	5	5
https://code.tutsplus.com/tutorials/design-patterns-the-singleton-pattern–cms-23073	1	1	1	1	1	1		3	6
http://wiki.freepascal.org/Singleton_Pattern	1	1	1	1		1	1	4	6

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://cocoacasts.com/are-singletons-bad/	1	1	1	1	1		1	5	6
http://robdodson.me/javascript-design-patterns-singleton/	1	1	1	1		1	1	5	6
https://www.vojtechruzicka.com/singleton-pattern-pitfalls/	1	1	1	1		1	1	5	6
https://whydoesitsuck.com/why-the-singleton-pattern-sucks-and-you-should-avoid-it/	1	1	1	1		1	1	5	6
https://theburningmonk.com/2013/09/dart-implementing-the-singleton-pattern-with-factory-constructors/	1	1	1	1	1	1	1	4	7
http://enterprisecraftsmanship.com/2016/05/03/singleton-vs-dependency-injection/	1	1	1	1	1	1	1	4	7
http://gameprogrammingpatterns.com/singleton.html	1	1	1	1	1	1	1	5	7
https://blog.ndepend.com/singleton-pattern-costs/	1	1	1	1	1	1	1	5	7
https://www.michaelsafyan.com/tech/design/patterns/singleton	1	1	1	1	1	1	1	5	7
https://www.zaraffasoft.com/2016/10/14/singleton-pattern-the-light-or-the-dark-side-of-the-force/	1	1	1	1	1	1	1	5	7
https://carlalexander.ca/singletons-in-wordpress/	1	1	1	1	1	1	1	5	7
https://krakendev.io/blog/antipatterns-singletons	1	1	1	1	1	1	1	5	7
https://www.objc.io/issues/13-architecture/singletons/	1	1	1	1	1	1	1	5	7
http://rcardin.github.io/design/programming/2015/07/03/the-good-the-bad-and-the-singleton.html	1	1						3	2

fridaya@uef.fi

Appendix 2. Raw Data: Dependency Injection

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad	
https://msdn.microsoft.com/en-us/library/hh323705(v=vs.100).aspx	1	1	1	1	1	1	1	1	7	
https://hackernoon.com/you-dont-need-to-know-dependency-injection-2e9d2ba1978a	1	1	1			1	1	1	5	
https://www.theserverside.com/news/1321158/A-beginners-guide-to-Dependency-Injection	1	1	1	1		1	1	1	6	
https://angular.io/guide/dependency-injection-pattern	1	1	1	1	1	1	1	1	7	
https://martinfowler.com/articles/injection.html	1	1	1	1	1	1	1	1	7	
https://medium.freecodecamp.org/demystifying-dependency-injection-49d4b6fe6536	1	1	1	1	1	1	1	1	7	
https://documentation.magnolia-cms.com/display/DOCS56/Dependency+injection+and+inversion+of+control	1	1						1	2	
http://www.tutorialsteacher.com/ioc/dependency-injection	1	1	1	1	1	1	1	1	7	
https://www.codeguru.com/csharp/.net/net_asp/mvc/understanding-dependency-injection.htm	1	1	1	1	1	1	1	1	7	
http://krasimirtsonev.com/blog/article/Dependency-injection-in-JavaScript	1	1	1	1	1		1	1	6	
https://www.devbridge.com/articles/dependency-injection-in-javascript/	1	1	1	1	1	1	1	1	7	
https://codecraft.tv/courses/angular/dependency-injection-and-providers/overview/	1	1	1	1	1	1	1	1	7	

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://aspnetboilerplate.com/Pages/Documents/Dependency-Injection	1	1	1	1	1	1	1	1	7
https://webdev.dartlang.org/angular/guide/dependency-injection	1	1	1	1	1	1	1	1	7
https://doc.nette.org/en/2.4/dependency-injection	1	1	1	1	1	1	1	1	7
https://medium.com/makingtuenti/dependency-injection-in-swift-part-1-236fddad144a	1	1					1	1	3
https://android.jlelse.eu/android-mvp-architecture-with-dependency-injection-dee43fe47af0	1	1	1	1	1	1	1	1	7
https://www.dotnettricks.com/learn/dependencyinjection/implementation-of-dependency-injection-pattern-in-csharp	1	1	1				1	1	4
https://nehalist.io/dependency-injection-in-typescript/	1	1	1	1	1	1	1	1	7
https://www.playframework.com/documentation/2.6.x/ScalaDependencyInjection	1	1	1	1			1	1	5
https://www.raywenderlich.com/171327/dependency-injection-android-dagger-2	1	1	1	1	1	1	1	1	7
https://blog.drewolson.org/dependency-injection-in-go/	1	1					1	1	3
https://docs.oracle.com/javaee/6/tutorial/doc/giwhl.html	1						1	1	2
https://stackify.com/dependency-injection/	1	1	1				1	1	4
https://www.swiftbysundell.com/posts/dependency-injection-using-factories-in-swift	1	1					1	1	3

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://matthiasnoback.nl/2018/06/road-to-dependency-injection/	1	1	1	1	1	1	1	1	7
https://codeburst.io/dependency-injection-with-vue-js-f6b44a0dae6d	1	1	1			1	1	1	5
https://www.getopensocial.com/blog/open-source-technology/dependency-injection-php	1	1	1	1	1	1	1	1	7
https://www.tutorialspoint.com/spring/spring_dependency_injection.htm	1	1	1	1	1	1	1	1	7
https://fsharpforfunandprofit.com/posts/dependency-injection-1/	1	1	1	1	1	1	1	1	7
http://fabien.potencier.org/what-is-dependency-injection.html	1	1	1	1	1	1	1	1	7
https://www.bignerdranch.com/blog/dependency-injection-ios/	1	1	1	1	1	1	1	1	7
https://code.tutsplus.com/tutorials/dependency-injection-in-php–net-28146	1	1	1	1	1	1	1	1	7
https://docs.sitefinity.com/use-constructor-dependency-injections-mvc	1		1				1	1	3
https://www.techyourchance.com/dependency-injection-android/	1	1	1				1	1	4
http://andrewembler.com/2018/03/concrete-guide-dependency-injection	1	1	1	1	1	1	1	1	7
https://www.tomasvotruba.cz/blog/2018/05/07/why-you-should-combine-symfony-console-and-dependency-injection/	1	1	1	1			1	1	5
https://codeshare.co.uk/blog/how-to-start-using-dependency-injection-in-mvc-and-umbraco/	1	1	1				1	1	4
https://samueleresca.net/2017/07/inversion-of-control-and-unit-testing-using-typescript/	1	1	1				1	1	4

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://php.earth/docs/php/ref/oop/design-patterns/dependency-injection	1	1	1	1	1	1	1	1	7
https://developer.telerik.com/featured/three-ds-web-development-3-dependency-injection/	1	1	1				1	1	4
https://learnappmaking.com/dependency-injection-swift/	1	1	1	1	1		1	1	6
https://www.guru99.com/angularjs-dependency-injection.html	1	1					1	1	3
https://appliedgo.net/di/	1	1	1	1	1	1	1	1	7
http://rubyblog.pro/2016/10/ruby-dependency-injection	1	1	1	1	1	1	1	1	7
https://vuejs.org/v2/guide/components-edge-cases.html	1	1					1	2	3
https://blog.gojekengineering.com/the-many-flavours-of-dependency-injection-in-go-25aa070d79a0	1	1	1				1	2	4
https://www.eclipse.org/che/docs/guice.html	1	1	1				1	2	4
http://flowframework.readthedocs.io/en/stable/TheDefinitiveGuide/PartIII/ObjectManagement.html	1	1	1	1	1		1	2	6
https://enterprisecraftsmanship.com/2016/05/03/singleton-vs-dependency-injection/	1	1	1	1	1	1	1	2	7
http://best-practice-software-engineering.ifs.tuwien.ac.at/patterns/dependency_injection.html	1	1	1					3	3
http://www.baeldung.com/inversion-control-and-dependency-injection-in-spring	1	1	1			1		3	4
http://www.vogella.com/tutorials/DependencyInjection/article.html	1	1	1					3	3

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad	
https://deviq.com/dependency-injection/	1							3	1	
https://dotnet.github.io/orleans/Documentation/Core-Features/Dependency-Injection.html	1	1						3	2	
https://aurelia.io/docs/fundamentals/dependency-injection/	1	1						3	2	
https://symfony.com/doc/current/components/dependency_injection.html	1	1						3	2	
https://guides.emberjs.com/release/applications/dependency-injection/	1	1						3	2	
https://itnext.io/typescript-dependency-injection-setting-up-inversifyjs-ioc-for-a-ts-project-f25d48799d70	1	1						3	2	
https://www.drupal.org/docs/8/api/services-and-dependency-injection/services-and-dependency-injection-in-drupal-8	1	1						3	2	
https://devdocs.magento.com/guides/v2.2/extension-dev-guide/depend-inj.html	1	1						3	2	
https://loopback.io/doc/en/lb4/Dependency-injection.html	1	1	1					3	3	
https://blog.kotlin-academy.com/dependency-injection-the-pattern-without-the-framework-33cfa9d5f312	1	1						3	2	
http://docs.automapper.org/en/stable/Dependency-injection.html	1	1						3	2	
https://pimcore.com/docs/4.6.x/Development_Documentation/Extending_Pimcore/Dependency_Injection.html	1	1						3	2	
https://www.mvvmcross.com/documentation/fundamentals/dependency-injection	1	1						3	2	

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad	
https://blog.carbonfive.com/2018/03/19/lightweight-dependency-injection-in-elixir-without-the-tears/	1	1						3	2	
https://learn-blazor.com/architecture/dependency-injection/	1	1						3	2	
https://medium.com/scribd-data-science-engineering/weaver-a-painless-dependency-injection-framework-for-swift-7c4afad5ef6a	1	1						3	2	
https://www.javatpoint.com/dependency-injection-in-spring	1	1						3	2	
https://blog.angularindepth.com/angular-dependency-injection-and-tree-shakeable-tokens-4588a8f70d5d	1							3	1	
https://www.yiiframework.com/doc/guide/2.0/en/concept-di-container	1	1						3	2	
http://blog.thecodewhisperer.com/permalink/keep-dependency-injection-simple	1							3	1	
https://docs.phalconphp.com/hr/3.3/di	1	1					1	3	3	
https://docs.silverstripe.org/en/4/developer_guides/extending/injector/	1	1						3	2	
https://struts.apache.org/core-developers/dependency-injection.html	1							3	1	
https://exceptionnotfound.net/using-entity-framework-dbcontext-with-dependency-injection/	1	1	1	1				3	4	
https://dzone.com/articles/optional-dependency-injection-with-spring		1						3	1	
http://picocontainer.com/injection.html	1							3	1	

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad
https://www.pacoworks.com/2018/02/25/simple-dependency-injection-in-kotlin-part-1/	1	1						3	2
https://proandroiddev.com/better-dependency-injection-for-android-567b93353ad		1						3	1
https://getakka.net/articles/actors/dependency-injection.html		1						3	1
http://docs.drush.org/en/master/dependency-injection/	1	1						3	2
https://devdocs.magento.com/guides/v2.2/extension-dev-guide/depend-inj.html	1	1						3	2
https://doc.sitecore.net/sitecore_experience_platform/developing/developing_with_sitecore/dependency_injection	1	1						3	2
https://docs.litium.com/documentation/architecture/dependency-injection	1							3	1
https://arrow-kt.io/docs/patterns/dependency_injection/		1						3	1
http://www.dotnetcurry.com/aspnet-core/1426/dependency-injection-di-aspnet-core	1	1						3	2
https://framework.zend.com/manual/2.4/en/tutorials/quickstart.di.html	1	1						3	2
https://www.future-processing.pl/blog/introduction-to-dependency-injection/	1	1						3	2
https://docs.particular.net/nservicebus/dependency-injection/	1	1						3	2
https://www.oreilly.com/ideas/handling-dependency-injection-using-java9-modularity	1	1						3	2

(*continued*)

fridaya@uef.fi

(*continued*)

URL	Describe	Example	Implications	Alternatives	Comparison code	Pros/Cons	Recommendation	Recommendation scale	Jadad	
https://blog.nrwl.io/essential-angular-dependency-injection-a6b9dcca1761	1	1						3	2	
https://toddmotto.com/angular-dependency-injection	1	1						3	2	
https://blog.mexia.com.au/dependency-injections-on-azure-functions-v2	1	1	1	1	1			3	1	
http://tutorials.jenkov.com/dependency-injection/index.html	1	1	1	1				3	4	
https://chromatichq.com/blog/dependency-injection-drupal-8-plugins	1	1						3	2	
http://jbehave.org/reference/stable/dependency-injection.html		1						3	1	
https://www.infoworld.com/article/2974298/application-architecture/exploring-the-dependency-injection-principle.html	1	1						3	2	

fridaya@uef.fi

References

1. Hevery, M.: Flaw: Constructor Does Real Work (2008). http://misko.hevery.com/code-reviewers-guide/flaw-constructor-does-real-work/
2. Sabir, F., Palma, F., Rasool, G., Guéhéneuc, Y., Moha, N.: A systematic literature review on the detection of smells and their evolution in object-oriented and service-oriented systems. Softw. Pract. Experience **79**(1), 3–39 (2018)
3. Sayfan, M.: Avoid Global Variables, Environment Variables, and Singletons. https://sites.google.com/site/michaelsafyan/software-engineering/avoid-global-variables-environment-variables-and-singletons
4. Densmore, S.: Why Singletons Are Evil (2004). http://blogs.msdn.com/b/scottdensmore/archive/2004/05/25/140827.aspx
5. Radford, M.: Singleton - the anti-pattern. Overload **57** (2003)
6. Yegge, S.: Singleton Considered Stupid (2004). https://sites.google.com/site/steveyegge2/singleton-considered-stupid
7. Ronacher, A.: Singletons and their problems in Python (2009). http://lucumr.pocoo.org/2009/7/24/singletons-and-their-problems-in-python/
8. Brown, W.: Why Singletons are "Bad Patterns" (2013). http://brollace.blogspot.co.uk/2013/04/why-singletons-are-bad-patterns.html
9. Kofler, P.: Why Singletons Are Evil (2012). http://blog.code-cop.org/2012/01/why-singletons-are-evil.html
10. Weaver, R.: Static methods vs singletons: choose neither (2010). http://www.phparch.com/2010/03/static-methods-vs-singletons-choose-neither/
11. Badu, K.: What's so evil about Singleton? (2008). http://www.sitepoint.com/forums/showthread.php?530917-What-s-so-evil-about-Singleton
12. Hart, S.: Why helper, singletons and utility classes are mostly bad (2011). http://smart421.wordpress.com/2011/08/31/why-helper-singletons-and-utility-classes-are-mostly-bad-2/
13. Alipour, G., Sangar, A., Mogaddam, M.: Aspect oriented Implementation of design patterns using metadata. J. Fundam. Appl. Sci. **57**, 66–75 (2016)
14. Liu, H., Cai, C., Zu, C.: An object-oriented serial implementation of a DSMC simulation package. J. Fundam. Appl. Sci. **8**, 816–825 (2011)
15. Knack-Nielsen, T.: What's so bad about the Singleton? (2008). http://www.sitepoint.com/whats-so-bad-about-the-singleton/
16. Fowler, M., Beck, K., Brant, J., Opdyke, W., Roberts, D.: Refactoring: Improving the Design of Existing Code (Object Technology Series). Addison Wesley Longman, Boston (1999). ISBN: 0201485672
17. Cochrane, C.: Cochrane. http://www.cochrane.org/
18. Mathie, R., Frye, J., Fisher, P.: Homeopathic Oscillococcinum® for preventing and treating influenza and influenza-like illness. Cochrane Database Syst. Rev. **12** (2015)
19. Goldacre, B.: Bad Science. Fourth Estate, London (2010). ISBN: 978-0-00-724019-7
20. Jadad, A., Moore, A., Carroll, D., Jenkinson, C.: Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control. Clin. Trials **17**(1), 1–12 (1996)
21. Olivo, S., Macedo, L., Caroline, I., Fuentes, J., Magee, D.: Scales to assess the quality of randomized controlled trials: a systematic review (Research Report). Phys. Ther. **88**(2), 156 (2008)
22. Buss, M.: Interfaces vs Inheritance in Swift (2016). https://mikebuss.com/2016/01/10/interfaces-vs-inheritance/
23. Ericson, B.: Association vs Inheritance (1995). http://ice-web.cc.gatech.edu/ce21/1/static/JavaReview-RU/OOBasics/ooAssocVsInherit.html

fridaya@uef.fi

24. Qwant, Q.: Qwant search engine. https://www.qwant.com/
25. Albert, A.: Why should we use dependency injection? (2013). http://www.javacreed.com/why-should-we-use-dependency-injection/
26. Fowler, M.: Inversion of Control Containers and the Dependency Injection pattern (2004). http://martinfowler.com/articles/injection.html

fridaya@uef.fi

Modelling and Analysis of Partially Stochastic Time Petri Nets Using Uppaal Model Checkers

Christian Nigro, Libero Nigro(✉), and Paolo F. Sciammarella

DIMES, University of Calabria, Rende, Italy
l.nigro@unical.it, p.sciammarella@dimes.unical.it

Abstract. Modelling and analysis of time-dependent concurrent/distributed systems is very challenging when such systems also exhibit a stochastic behavior. In this paper, the partially stochastic Time Petri Net formalism (psTPN) is introduced, which has the expressive power to model realistic complex systems. Each transition owns both a non-deterministic temporal interval and a stochastic duration established by a generally distributed (i.e., not necessarily Markovian) probabilistic function whose samples are constrained to occur in the associated time interval. A psTPN model admits two possible interpretations: the non-deterministic (which ignores stochastic aspects), useful for *qualitative analysis* (establishing that an event *can* possibly occur) and the stochastic one (which considers stochastic behavior), useful for *quantitative analysis* (estimating the probability for an event to actually occur). The paper focuses on the reduction of psTPN onto UPPAAL timed automata which permit non-deterministic analysis through the symbolic exhaustive model checker, and quantitative analysis through the statistical model checker (SMC) which rests on simulations and statistical inference. Although potentially less accurate of analysis techniques based on numerical methods and stochastic state space enumeration, this paper claims that statistical model checking can provide quantitative property assessment which is of practical engineering value. As an example, psTPN is exploited in the paper to model and evaluate a stochastic version of the Fisher timed-based mutual exclusion algorithm.

Keywords: Time Petri Nets · Stochastic behavior · Timing constraints · Fisher's mutual exclusion · Model checking · Statistical model checking · UPPAAL

1 Introduction

Modern society depends more and more on software systems which are both distributed and time-dependent. Such systems can be in control of applications in the domain of economy, engineering, healthcare, biology, science and so forth. The development of such systems is challenging for the necessity of ensuring correctness

C. Nigro—Independent Computing Professional, Italy.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 975–993, 2019.
https://doi.org/10.1007/978-3-030-22868-2_66

fridaya@uef.fi

of both *functional* and *temporal* system requirements. Failing to fulfil some timing constraints can have severe consequences in the practical case. Therefore, the use of formal modelling languages and tools is highly recommended during the early design and analysis phases. Moreover, modelling and analysis can become worst when timing gets combined with a probabilistic/stochastic behavior [1–5]. In these cases, often it is the case that system behavior can be studied in an approximate way by simulation tools [6], being impossible to completely analyze a not decidable system, as occurs, e.g., when non-deterministic timing is used together with general stochastic aspects.

In the work described in this paper, the use of Time Petri Nets [7, 8] extended with stochastic issues is considered. In particular, following similar approaches described in the literature [3, 9], a partially stochastic Time Petri net (psTPN) formalism is adopted. Each transition is associated with a non-deterministic time interval and a probability distribution function (*pdf*) whose samples are constrained to belong to the associated (supporting) temporal interval. A psTPN model admits two time interpretations. In the non-deterministic interpretation, pdfs are ignored, and the model can be studied for qualitative analysis (assessing that an event *can* occur). In the stochastic interpretation, the constrained pdfs are enabled and the goal is to allow a quantitative analysis of system behavior (estimating a probability measure for an event occurrence).

A notable representative of quantitative analysis is based on numerical methods [10] and the enumeration (that is, the construction of a transition system or model state graph) of symbolic *stochastic state classes* [11], which can guarantee an accurate quantitative evaluation of system properties [3, 4]. However, these approaches can suffer of state explosion problems when facing general complex systems, or they cannot exploit specific assumptions (e.g., the existence of regenerative points in the state space [3]).

In this paper, an original, yet practical, approach is proposed for supporting psTPN which is based on (stochastic) timed automata (TA) [12] in the context of popular Uppaal model checkers [13, 14]. As, e.g., in the actor-based modelling described in [5], the proposed approach enables both qualitative and quantitative evaluation of system properties. Qualitative non-deterministic verification is based on exhaustive symbolic model checking [13, 15] where probability/stochastic aspects are temporarily disabled. Quantitative evaluation is instead based on the statistical model checker (SMC) [16] of Uppaal [14] which rests on simulation runs and statistical inference. SMC does not build a model state graph and its memory consumption is linear with the model size. With respect to ad hoc simulation tools, SMC automates the simulation runs and offers a temporal logic language for specifying application tailored properties to check.

The contributions of this paper are: (a) describing the adopted formalism of psTPN which generalizes definitions reported in [3, 9]; (b) showing a reduction from psTPN onto Uppaal; (c) applying psTPN to a stochastic version of the Fisher's timed-based mutual exclusion algorithm [3, 17]; (d) documenting an experimental work with the Fisher's protocol which confirms known properties but it is also capable of revealing new details.

The paper is structured as follows. Section 2 provides basic definitions of the psTPN formalism. Section 3 describes the chosen Fisher algorithm as a modelling example.

fridaya@uef.fi

Section 4 illustrates the reduction process of psTPN onto UPPAAL using the Fisher model to detail the translation process. The section discusses both the qualitative and quantitative property checking of a reduced model. Section 5 reports the experimental work which was carried out on the reduced Fisher protocol model. Section 6 terminates the paper by presenting conclusions and indicating directions of future work.

2 Definition of Partially Stochastic Time Petri Nets

psTPN aims at defining models of concurrent, stochastic and timed systems. One such a model consists of *transitions* (small bars in Fig. 1) which have both a non-deterministic time interval [EarliestFiringTime or EFT, LatestFiringTime or LTF] as in Time Petri Nets [7, 8] and a stochastic duration established by a probability distribution function (pdf). The samples of this PDF are constrained to belong to the transition time interval. Transitions represent activities which can occur in the modelled system. A psTPN model also introduces a set of *places* (circles in Fig. 1) and a set of *arcs* (arrows in Fig. 1) linking places to transitions and transitions to places only. Each place can contain a number of *tokens* (black dots in Fig. 1) and each arc has a natural *weight* which can also be 0. Places are used in *input* and *output* of transitions. Input places, as in classical Petri nets, constrain the *enabling* of a transition. A transition is enabled if each of its input places holds a number of tokens greater than or equal to the corresponding input arc weight. *Inhibitor arcs* permit to input places to contribute to the enabling in the case of absence of tokens. When a transition *fires*, a number of tokens are *withdrawn* from input places as dictated by the input arc weights, and a number of tokens are *deposited* to the output places of the transition, again according to output arc weights. The process of transition firing is *atomic* and obviously can influence the enabling of other transitions.

In the non-deterministic interpretation of a psTPN, pdfs are ignored. Each transition holds a *timer* which is started as soon as the transition gets enabled. The transition cannot conclude its firing if the timer has not reached the EFT, but it has to fire before the timer goes beyond the LFT. A transition can lose its enabling at any time and even at its *last time* LFT. A transition can fire at a given time instant, provided that no other transition would have to be fired before, possibly disabling this transition.

In the stochastic interpretation, at its enabling, a transition samples its pdf for a duration compliant with the associated time interval. A transition can fire when its timer reaches the duration.

2.1 Formal Definitions

A *psTPN* is a tuple $(P, T, B, F, I_{nh}, M_0, ExtE, ExtW, ExtD, EFT, LFT, PDF)$ where P is a set of places, T is a set of transitions with $P \cap T = \emptyset$, but $P \cup T \neq \emptyset$; B is the *backward* incidence function, $B : P \times T \rightarrow \mathbb{N}$, which assigns a weight to each input arc; F is the *forward* incidence function, $F : P \times T \rightarrow \mathbb{N}^+$, which assigns a positive weight to each output arc; I_{nh} is the set of the inhibitor arcs: $I_{nh} \subset P \times T$ where $(p, t) \in I_{nh} \Rightarrow B(p, t) = 0$; M_0 is the initial marking, $M_0 : P \rightarrow \mathbb{N}$ which assigns to each

fridaya@uef.fi

place an initial number of tokens. $ExtE, ExtW, ExtD$ are three *extension functions* which enhance respectively the enabling, the withdraw and the deposit phases of each transition. Such functions (see also, e.g., [9]) are intended to simplify the design of complex models by avoiding intricate arcs and topology cluttering. In addition, they make it possible to handle an abstract notion of places where tokens are high-level data (e.g., an integer, as occurs in the model example in Fig. 1). $EFT : T \rightarrow Q^+$ is a function which associates to each transition a (finite) earliest firing time. Q denotes the set of non-negative rational numbers; $LFT : T \rightarrow Q^+ \cup \{\infty\}$ is a function which associates to each transition a (possibly infinite) latest firing time. It must be $LFT(t) \geq EFT(t)$ for $t \in T$. PDF is a function: $PDF : T \rightarrow pdf$ which assigns to each transition a probability density function pdf(t). Each sample of the pdf is constrained to belong to its time interval:

$$EFT(t) \leq pdf(t) \leq LFT(t)$$

Transition Enabling. A transition t is enabled at marking m, denoted by $m[t >$, when its input places contain sufficient tokens:

$$\forall p \in P, (p,t) \in I_{nh} \Rightarrow M(p) = 0 \wedge$$

$$B(p,t) > 0 \Rightarrow M(p) \geq B(p,t) \wedge ExtE(t)$$

Transition Firability. Under non-deterministic interpretation, a transition t of a psTPN model is *firable* at the global time τ if it is enabled and its *time-to-fire* reflected into its *timer* is within its time interval $[EFT, LFT]$ and no transition had to fire before because it already reached its last time to fire. A transition is *fireable* under the stochastic interpretation when it is enabled and its time-to-fire reflected by its timer has reached its duration.

Transition Firing. When $t \in T$ fires, it transforms current marking m into a new marking m'' thus:

$$m'(p) = (m(p) - B(p,t)) \wedge ExtW(t)(\text{withdraw sub - phase})$$
$$m''(p) = (m'(p) + F(p,t)) \wedge ExtD(t)(\text{deposit sub - phase})$$

Purposely, the atomic firing process is split into two sub-phases which correspond respectively to the withdrawal and the deposit of tokens. m' is the intermediate marking reached after the withdraw sub-phase. A transition $t^\sim$ can change its status (enabled/disabled) or it can remain unaltered since the firing of t. If $m[t^\sim > \wedge m'[t^\sim > \wedge m''[t^\sim >$, transition $t^\sim$ is said to be *persistent* to the firing of t. A transition $t^\sim$ s.t. $m[t^\sim >$ can become disabled into m' or into m'' (e.g., due to inhibitor arcs), in which case t is said *not persistent*. Finally, a transition $t^\sim$ (including t itself) can become enabled into the final marking m''. It is said a *newly enabled* transition.

fridaya@uef.fi

Single Server. For simplicity, when a transition t has multiple enablings, it is assumed (see also page 100 in [18]) that it will fire its enablings one at a time, sequentially. This forbids parallel firings of a same transition.

3 Modelling Example

In the following a psTPN model of the Fisher's time-based mutual exclusion protocol [17] is shown. The model was adapted from [3] and retains the same timing information for validation purposes. In Fig. 1 three processes are assumed, whose id are 1, 2 and 3.

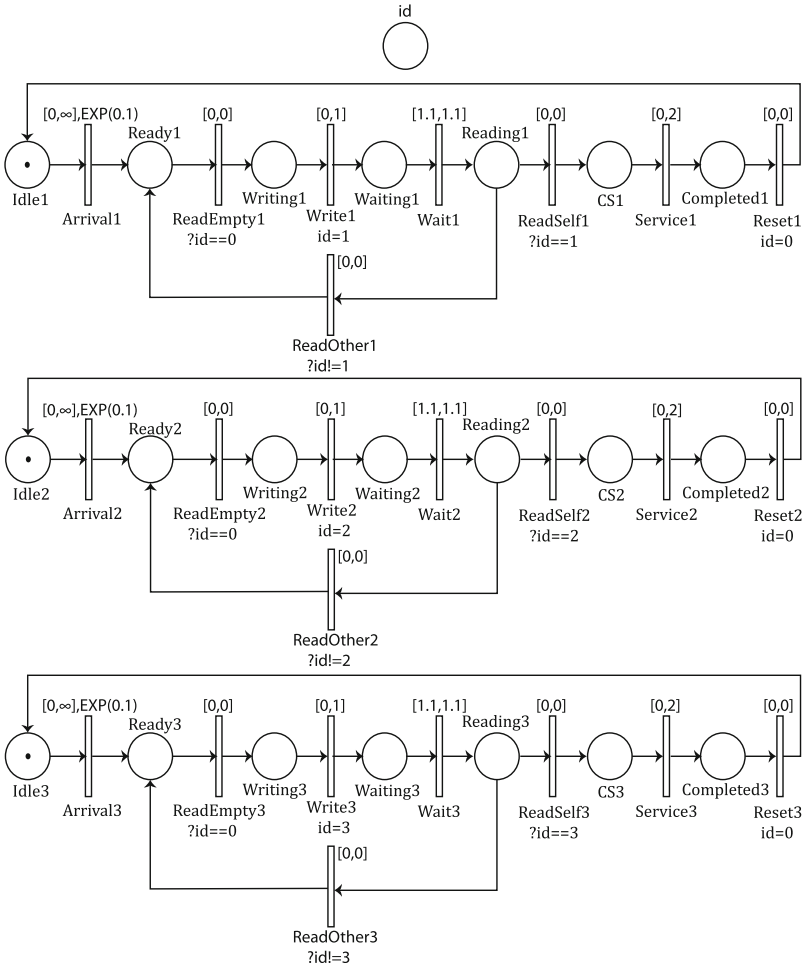


Fig. 1. A psTPN model for the Fisher's timed-based mutual exclusion algorithm, adapted from [3].

fridaya@uef.fi

Shared place *id* holds an integer as its token, whose value can be 0 when no process is interested in entering its critical section. The value of *id* is $i = 1, 2 \, or \, 3$ when the process i intends to enter its critical section. The place *id* is handled through the $ExtE, ExtW$, and $ExtD$ functions.

The $Arrival_i$ transition has an interval $[0, \infty]$ and its pdf is an exponential distribution whose parameter λ is 0.01 s^{-1}. All the other transitions have, by default, a uniform distribution in the associated temporal interval. In particular, $ReadEmpty_i$, $ReadSelf_i$, $Reset_i$, $ReadOther_i$ and $Wait_i$ have a pdf which reduces to a deterministic constant value which is 1.1 for $Wait_i$ and 0 for the other transitions.

Initially, each process is in its non-critical section (NCS) modelled by one token in the $Idle_i$ place. The dwell-time in NCS is arbitrary as reflected by the [0,∞] time interval of the $Arrival_i$ transition. After a firing of $Arrival_i$, the process raises its interest to compete to enter its critical section by putting one token in the $Ready_i$ place. After that, the process continues by reading the value in the *id* place. In the case $id == 0$ the process can go on. If $id \neq 0$, another process assigned its id to the shared place *id*, and the current process has to remain in the ready status (transition $ReadEmpty_i$ is *not* enabled). When $id == 0$ the process goes on by tentatively writing its id to the shared place *id*. The writing time (W) was supposed to be a non-deterministic value in the interval $[0, 1]$. A key point of the protocol consists now in forcing the process to wait for a time just greater than W. In Fig. 1 the waiting time is the constant 1.1. After the waiting phase, the process checks again (reading phase) if its id is still in the *id* place or not. A different value would testify that another process attempted the same writing concurrently and succeeded. Therefore, the process has to come back into its ready status by a firing of the $ReadOther_i$ transition. When its own id is found in *id*, the process (through a firing of $ReadSelf_i$) eventually enters its critical section (one token is deposited in the CS_i place). The duration of the critical section is modelled by the time interval of the $Service_i$ transition. After its firing, $Service_i$ puts one token in the $Completed_i$ place and then, through $Reset_i$, the process frees the shared place *id* by putting 0 in it, and returns to its non-critical section ($Idle_i$ place).

The number of times the $ReadOther_i$ transition fires, represents the cases where this process is overtaken by competing processes. Hopefully, the competing time should be bounded before entering the critical section.

The following properties should be checked when proving the correctness of a mutual exclusion, in general untimed, algorithm (see, e.g., [19]):

1. The algorithm should be deadlock free (*safety*).
2. Only one process at a time can be in its critical section (*safety*).
3. The waiting time due to competing should be bounded (absence of starvation, *bounded liveness*).
4. A not-competing process should not forbid a competing one to enter its critical section (*liveness*).
5. No hypothesis should be made about the relative speed of the processes.

fridaya@uef.fi

It is worth noting that in the literature there are mutual exclusion algorithms which, e.g., satisfy the constraint 2, but not the constraint 3. A notable example is the Dijkstra's untimed mutual exclusion algorithm for N processes [20] which it can be proved it does not satisfy the constraint 3.

The analysis work described later in this paper aims at showing the interplay between non-deterministic behavior and the stochastic one, by confirming, in particular, that the stochastic behavior can positively contribute to the proper operation of the Fisher's algorithm.

4 Reducing psTPN onto Uppaal

The proposed reduction process enhances a similar approach described in [21]. In addition, the immediate transitions defined in [4, 21] and associated probabilistic weights and random switches are avoided, because the work described in this paper addresses the analysis of real-time stochastic systems and not the performance evaluation of general stochastic models.

Uppaal was chosen instead, e.g., of the probabilistic model checker PRISM [22], because it supports (a) clocks which naturally can be exploited to model transition timers; (b) high-level data structures and C-like functions which greatly contribute to compactness and efficiency of a psTPN model; (c) graphical modelling which favors intuitive design; (d) symbolic model checking and statistical model checking for the analysis.

The rationale behind the translation of psTPN onto Uppaal consists in modelling transitions (the active components of a model) as parameterized template processes, that is timed automata, and reproducing model topology by some global data structures. The translation is actually assisted by a customization of the TPN Designer tool [23] which automates the generation of data structures and transition templates of a source model under the non-deterministic interpretation. The achieved Uppaal model is then easily decorated with details corresponding to the stochastic interpretation. Details of the reduction process are provided in the following using the Fisher model described in Sect. 3 as a source.

4.1 Global Data Declarations

For the Fisher model, the following global declarations are generated together with some sub-range types describing the unique identifiers associated with places and transitions (see *pid* and *tid* subtypes).

fridaya@uef.fi

```
const int N=3; //nr of processes
const int T = 24; //nr of transitions
const int P = 21; //nr of places
const int PRE = 1; //max size of a transition preset
const int POST = 1; //max size of a transition postset
const int NONE = -1;
const int INF = -1;
//subrange types
typedef int[0,P-1] pid; //place ids
typedef int[0,T-1] tid; //transition ids
```

In order to keep a close correspondence with the high-level model, place and transition ids are explicitly generated:

```
//place identifiers
const int Idle1=0;
const int Ready1=1;
const int Writing1=2;
const int Waiting1=3;
const int Reading1=4;
const int CS1=5;
const int Completed1=6;
…
const int Completed3=20;

//transition identifiers
const int Arrival1=0;
const int ReadEmpty1=1;
const int Write1=2;
const int Wait1=3;
const int ReadSelf1=4;
const int Service1=5;
const int Reset1=6;
const int ReadOther1=7;
…
const int ReadOther3=23;
```

For a compact representation of a psTPN model, the (constant) matrices B (backward) and F (forward), and (mutable) vector M (marking) are introduced. B and F are TxPRE and TxPOST respectively, instead of being TxP. Their elements are of the following Info type. Matrix B serves to decide transition enabling and withdrawing of tokens at a transition firing. Matrix F is involved in the deposit of tokens at a transition firing.

fridaya@uef.fi

```
typedef struct{
    int index; //place id or NONE for a non-existing place
    int weight; //non-negative arc weight
} Info;

const Info B[T][PRE] = {
    {{Idle1,1}},          //Arrival1
    {{Ready1,1}},         //ReadEmpty1
    {{Writing1,1}},       //Write1
    {{Waiting1,1}},       //Wait1
    {{Reading1,1}},       //ReadSelf1
    {{CS1,1}},            //Service1
    {{Completed1,1}},//Reset1
    {{Reading1,1}},       //ReadOther1
    …
    {{Reading3,1}}        //ReadOther3
};

const Info F[T][POST] = {
    {{Ready1,1}},         //Arrival1
    {{Writing1,1}},       //ReadEmpty1
    {{Waiting1,1}},       //Write1
    {{Reading1,1}},       //Wait1
    {{CS1,1}},            //ReadSelf1
    {{Completed1,1}},//Service1
    {{Idle1,1}},          //Reset1
    {{Ready1,1}},         //ReadOther1
    …
    {{Ready3,1}}          //ReadOther3
};

int[0,1] M[P]={ 1,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0 }; //initial marking
```

Another (constant) matrix I as Tx2 of int is generated, which stores the earliest and latest firing times of each transition: I[t][0] holds the earliest firing time, I[t][1] is the latest firing time. Since Uppaal does not admit rational numbers as bounds in clock constraints, the original bounds of time intervals are scaled (by a factor of 10 for the chosen Fisher model) when building the I matrix.

```
const int I[T][2]={
    {0,INF},    //Arrival1
    {0,0},      //ReadEmpty1
    {0,10},     //Write1
    {11,11},    //Wait1
    {0,0},      //ReadSelf1
    {0,20},     //Service1
    {0,0},      //Reset1
    {0,0},      //ReadOther1
    …
    {0,0}       //ReadOther3
};
```

fridaya@uef.fi

Model translation is completed by the functions enabled(tid), withdraw(tid), deposit (tid) and extension functions ExtE(tid), ExtW(tid) and ExtD(tid) which regulate the enabling, and the token withdrawal/deposit of transitions. Extension functions are transparently called from within respectively the enabled(tid), withdraw(tid) and deposit(tid) functions. For the Fisher model, ExtW(tid) is void. ExtE(tid) checks the value of the *id* place and ExtD(tid) assigns a value to the *id* place as required by the model in Fig. 1.

4.2 Transition Template Processes

The active part of a reduced model rests on two fundamental timed automata *ndTransition* (for a non-deterministic transition) and *stTransition* (for a stochastic transition) plus a *Starter* automaton used for model bootstrap. The firing process is triggered by a global broadcast channel *end_fire* which allows to split the atomic firing process into the two sub-phases of withdrawal and deposit, and forces all the remaining transitions of a model to re-evaluate their status just after each of the sub-phases. All of this is a key for proper management of persistence/not persistence of transitions at the time of firing of a particular transition.

ndTransition Automaton. Figure 2 shows the ndTransition template automaton, which has one single parameter: const tid t. In addition, the template process introduces a local clock x assisting the firing process.

An informal argument of correctness of the ndTransition follows. A relative time model is used. The transition begins in the not enabled location N, from which it reaches the F (under firing) location when another transition in the model completes its firing and this transition finds itself enabled. In the edge from N to F, clock x is reset. Location F is constrained by an invariant on the upper bound (latest firing time). In the case the latest firing time is infinite, the transition can stay in F an arbitrary time. At any time, the transition moves from F to N as it detects itself to be disabled. As the clock reaches the lower bound (earliest firing time) I[t][0], the transition can conclude its firing by starting the token withdrawal and moving to the committed W location (in Uppaal a committed location has to be abandoned without a time passage). From W, two subsequent broadcast synchronizations through the end_fire are generated. The

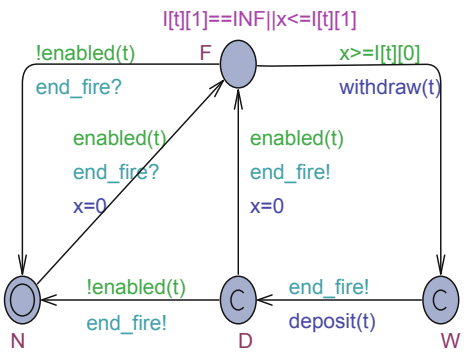


Fig. 2. The ndTransition automaton.

fridaya@uef.fi

first one serves to enter the D location by making the deposit of tokens in the output places. The second synchronization can either moves the transition from D to N if it is no longer enabled, or it could go again into F would the transition be still enabled, by resetting the clock x. All of this complies with the single-server semantics of transition firing.

A subtle point in Fig. 2 concerns the effects of the first end_fire synchronization. At the time of this synchronization, all the remaining transitions of the model (excluding this transition t) are forced to re-evaluate their status on the basis of the just carried withdrawal of tokens. At the time of second end_fire synchronization, model transitions will check again their status as it was influenced by the just carried deposit of tokens.

The automaton in Fig. 2 correctly behaves even when a transition loses its enabling at the *last time* of its temporal interval. It is worth noting that transition firing at a same time is non-deterministic and that transition disabling has greater priority than transition firing due to the end_fire signal being generated from a committed location.

stTransition Automaton. Figure 3 shows the stTransition automaton (with parameter: const tid t) which enforces a stochastic behavior to the psTPN model. Many aspects of transition firing are identical to those explained for the ndTransition in Fig. 2.

The construction in Fig. 3 follows the pattern suggested in the Uppaal SMC tutorial [14]. This pattern aims at building an activity whose duration is established by sampling a user-defined probabilistic distribution function. Two clocks are used: *delay*, which gets initialized with the sampled duration, and the clock *x* which measures the elapsed time until the delay is reached. In the F location the stopwatch *delay* is stopped by forcing its first derivative to be 0. Exiting from the F location toward W is guarded by the clock *x* which has reached the *delay* value.

As one can see from Fig. 3, an auxiliary, urgent and broadcast channel force_firing is used to assist the exiting from F. The force_firing signal is fictitiously sent but it is received by no other automata. Uppaal SMC would not otherwise force exiting from F in all the cases the *delay* is reached, e.g., when the *delay* is 0 or a constant value.

The global function f(t) replaces the I matrix when the psTPN model is used according to the stochastic interpretation. In this case, Uppaal SMC allows the use of

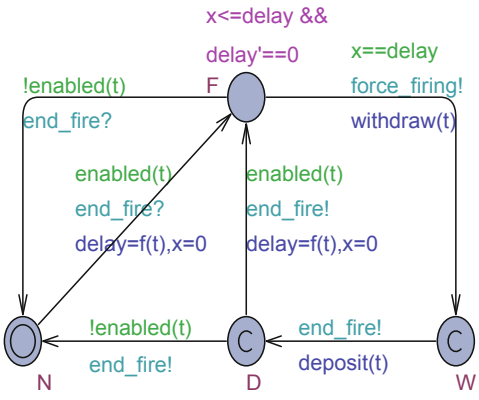


Fig. 3. The stTransition automaton

fridaya@uef.fi

```
double f( const tid ID ){
    if( ID==Arrival1 || ID==Arrival2 || ID==Arrival3 ){
        return -(1.0/0.1)*ln(random(1.0)); //the support interval is [0,oo]
    }
    if( ID==ReadEmpty1 || ID==ReadEmpty2 || ID==ReadEmpty3 ||
        ID==ReadSelf1 || ID==ReadSelf2 || ID==ReadSelf3 ||
        ID==ReadOther1 || ID==ReadOther2 || ID==ReadOther3 ||
        ID==Reset1 || ID==Reset2 || ID==Reset3 ){//constant 0
        return 0.0;
    }
    if( ID==Write1 || ID==Write2 || ID==Write3 ){//UNI[0,1]
        return random(1.0);
    }
    if( ID==Wait1 || ID==Wait2 || ID==Wait3 ){//DET(1.1)
        return 1.1;
    }
    if( ID==Service1 || ID==Service2 || ID==Service3 ){//UNI[0,2]
        return random(1.0)*2.0;
    }
    return -1; //never executed
}//f
```

Fig. 4. The body of the function f(tid) for the Fisher model.

double variables. In Fig. 4 the body of this function for the Fisher model is shown. Note that now the exact timing information of Fig. 1 are used. It should be recalled that Uppaal uses by default the *uniform distribution* on a transition provided only of the time interval. The exponential pdf of the Arrival transitions implicitly comply with the support interval $[0,\infty]$.

The following are the global channel declarations used by a reduced psTPN model:

```
broadcast chan end_fire;
urgent broadcast chan force_firing;
```

Note that end_fire is not required to be urgent because it is always used on edges exiting from committed locations (see e.g. Fig. 2).

Starter Automaton. In Fig. 5 it is depicted the simple Starter automaton which broadcasts a first end_fire synchronization thus asking each transition of the model to evaluate their enabling status in the initial marking. All the enabled transitions will then move to the F location. After the first synchronization, the Starter takes no longer part to the model behavior.

Fig. 5. The starter automaton.

fridaya@uef.fi

4.3 System Composition

The composition of the system has to be adapted to the non-deterministic or stochastic interpretation of the model. For example, for non-deterministic analysis the following line can be used:

```
// Place template instantiations here.
// List one or more processes to be composed into a system.
system Starter, ndTransition;
```

ndTransition must be replaced by stTransition when the model has to be analyzed stochastically. As a final remark, the use of the sub-range type tid is a key factor for a default instantiation of the transition automata during system composition. T instances of ndTransition/stTransition are automatically created for a reduced psTPN model.

5 Analysis of the Fisher's Protocol

The Fisher's algorithm in Fig. 1 was used in [3] for an in-depth analysis of its stochastic behavior, using numerical methods and stochastic state enumeration. The reduction of the psTPN version of the Fisher model in Uppaal can be thoroughly analyzed under both the non-deterministic interpretation and the stochastic interpretation. Non-deterministic analysis ignores stochasticity and exploits the symbolic model checker of Uppaal which builds the (hopefully finite) model state graph and permits to investigate exhaustively the properties specified in the subset of the TCTL temporal logic language supported by the toolbox [13]. Stochastic analysis depends on the Uppaal Statistical Model Checker (SMC) which does not build the state graph and uses simulation runs and statistical inference for quantitatively estimating probability measures of event occurrences. The SMC rests on a weighted extension of the Metric Interval Temporal Logic (MITL) for formally specifying system properties.

5.1 Non-deterministic Analysis

The non-deterministic version of the psTPN Fisher model, adapted so as to convert temporal interval bounds of transitions to integer constants (see subsection 4.1 where the original timing information are scaled by a factor of 10), was extensively studied for *qualitative* behavior: detecting that something *can* happen. Since the three processes are identical, properties can be checked with a reference to one process, e.g. the process 1.

A first concern was to check for the absence of deadlocks in any execution state, with the query:

```
A[] !deadlock
```

The property is satisfied. Then, the mutual exclusion property was verified as follows:

```
A[] M[CS1]+M[CS2]+M[CS3] <= 1
```

fridaya@uef.fi

This property too is satisfied. Other equivalent queries could be used to assess the mutual exclusion, e.g.:

A[] M[CS1]==1 imply M[CS2]==0 && M[CS3]==0

which ensures that if process 1 is in its critical section, necessarily the other processes are *not* in their critical section. The property is satisfied.

As a further check, the following query gets not satisfied:

E<> M[CS1]+M[CS2]+M[CS3]>1

The last queries need to be accompanied by checks which can guarantee that effectively each process can possibly enter its critical section:

E<> M[CS1]==1

Such existential query is satisfied. The next query checks the property that processes which are in the non-critical-section do not forbid another process to enter its critical section:

E<> M[CS1]==1 && M[Idle2]==1 && M[Idle3]==1

The query is satisfied. Finally, absence of starvation was checked with the queries:

E<> M[Ready1]==1	satisfied
M[Ready1]==1 --> M[CS1]==1	not satisfied

The last query makes use of the *leads-to* operator – > , which asks if starting from a marking where process 1 is competing, inevitably follows that the process effectively enters its critical section. Unfortunately, Uppaal says the property is *not* satisfied. Therefore, a competing process can experiment unbounded waiting. This property was not verified in [3].

Considering that during model checking Uppaal investigates exhaustively *all* the process interleavings, it can be concluded that the Fisher model, under non-deterministic analysis, satisfies all the requirements stated at the end of Sect. 3, except for the constraint 3 about starvation which is not fulfilled.

It is useful to note that although each ndTransition owns its local clock, the complexity of model checking does not depend on the total number of clocks but, at any time, on the number of *active* clocks. An active clock is one that, after being reset, is actively referenced in guards or invariants. Therefore, it is the parallelism degree, i.e., the number of simultaneous under firing transitions, which really affects the efficiency of the model checking procedure on the psTPN model.

5.2 Stochastic Analysis

Whereas the non-deterministic analysis can indicate that something *can* happen, the stochastic analysis can be used to quantify a probability measure of event occurrence. In particular, since the chosen Fisher model does not exclude the starvation of a waiting/competing process, it is interesting to have a quantitative measure of the waiting time for processes wishing to enter their critical section.

fridaya@uef.fi

As discussed in [3], the stochastic aspects of a model like that of Fisher can positively contribute to improve the properties of the algorithm.

As a preparatory phase of a model to be studied with statistical model checking (SMC), the duration of each experiment (a simulation run) should be estimated. SMC will execute, and also adapt, a number of experiments in order to ensure a certain degree of accuracy in the results. Roughly speaking, the SMC evaluates the number of experiments in which an event occurs and extrapolates a probability measure by a proportion. As soon as the event occurs, the experiment is interrupted.

A first query was used to estimate the mean arrival time of a process that is the mean time to abandon the non-critical section (place $Idle_i$ in Fig. 1). Such a time has obviously to tend to the mean time of the exponential distribution of *rate* 0.1, which is $1/0.1 = 10$ time units. The Fisher model was first *decorated* by adding a double variable s which accumulates the samples generated by the distribution (in the $f(t)$ function) and also counting the number of samples in another variable ns. The decoration function $mean()$ was then prepared to furnish, time by time, the ratio s/ns (at the time 0, when $ns == 0$, $mean()$ returns 0). The following query shows the trend of the $mean()$ value:

simulate 1 [<=50000] { mean() }

The picture in Fig. 6 was directly proposed by Uppaal SMC and confirms that $5x10^4$ is a sufficient time for the exponential to converge to its mean time. As a consequence, $5x10^4$ was used as the limit time of subsequent experiments.

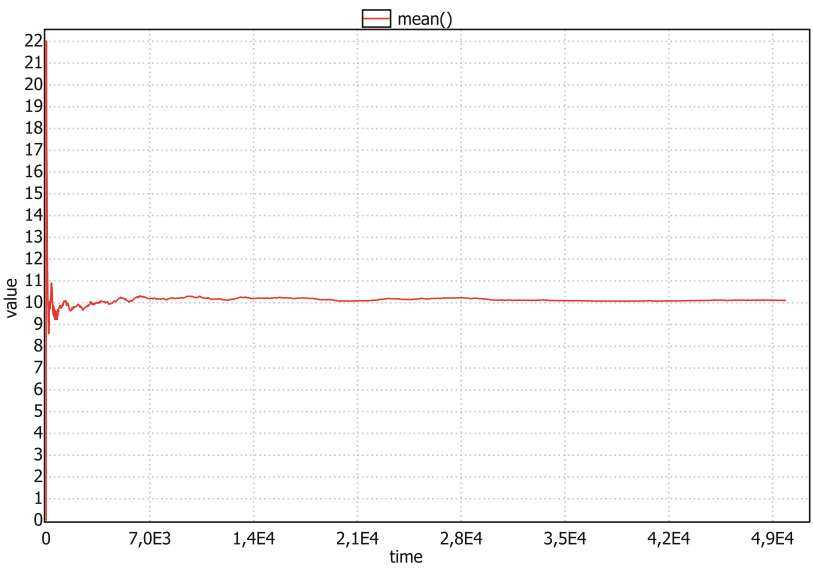


Fig. 6. Monitored mean arrival time of processes.

fridaya@uef.fi

Liveness and safety properties of the Fisher model were re-checked with SMC as follows:

Pr[<=50000] (<>M[CS1]==1)

Uppaal SMC, using 299 runs, proposes a confidence interval (CI) for the event "process 1 possibly enters its critical section" of [0.990031,1] with confidence 0.95. The result was achieved by setting the option $\varepsilon = 0.005$ as the error in the evaluation of a probability (the default value is 0.05). The result confirms the event is "almost certain". This option was kept also for the other queries.

Pr[<=50000] (<>M[CS1]+M[CS2]+M[CS3]>1)

Again, using 299 runs, Uppaal SMC proposes a CI of [0,0.00996915] 95% confidence, to suggest the event is "almost impossible" to occur.

To specifically check the fairness of the model, the following decorations were introduced in the model:

- a clock z which is reset at each firing of the $Arrival_1$ transition and each firing of the corresponding $ReadSelf_1$ (when process 1 definitely enters its critical section). Such reset operations are purposely accomplished in the deposit(tid) function;
- a double variable *wt* to annotate the waiting time when process 1 is competing; more in particular, at each firing of $ReadSelf_1$, the value of clock z is assigned to *wt* (this type of operation is only allowed in the SMC) before being reset;
- a double variable *mwt* which is updated, during each experiment, with the maximum value of the observed waiting time held by *wt*.

The query

E[50000;100] (max:mwt)

asks the SMC to estimate the maximum value of *mwt*, using 100 experiments each lasting 50000 time units. It is worth noting that in this case each experiment executes completely for the 50000 time units. Uppaal suggests a value of about 15.868 ± 0.433615. This result implicitly emerges also in the work described in [3]. It confirms that the stochastic behavior of the Fisher protocol helps toward fairness and starvation absence.

Fairness was also studied from a different perspective. The model was further decorated by adding a boolean variable *competing* which takes the value *true* at the firing of $Arrival_1$, and the value *false* at the firing of $ReadSelf_1$. The following query estimates (using extended weighted MITL) the probability of the event that within the time window $[0, 50000]$, it there exists an instant from which *competing* remains true for at most 13 consecutive time units:

Pr(<>[0,50000] ([] [0,13] competing))

Uppaal proposes a CI of $[0.897154, 0.997154]95\%$ confidence.

Other experiments were carried out to observe how the estimation of the maximum waiting time (*mwt*) changes when the parameter λ of the exponential distribution of the

fridaya@uef.fi

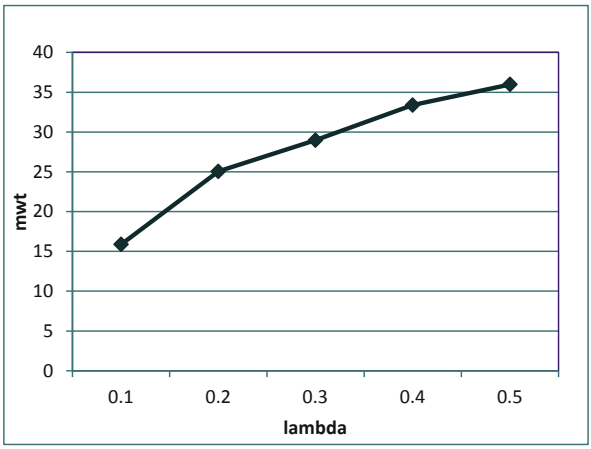


Fig. 7. Maximum waiting time vs. λ

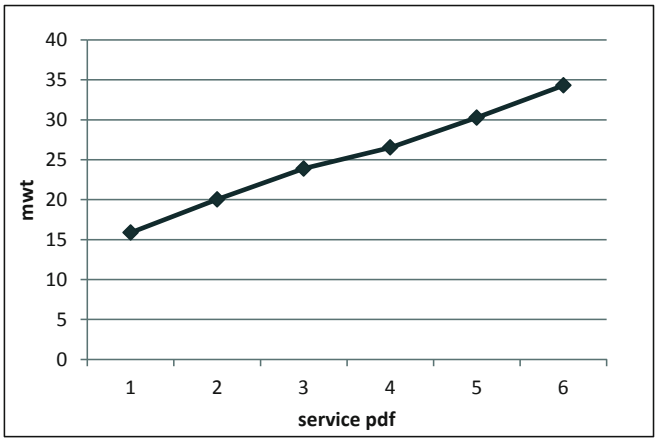


Fig. 8. Maximum waiting time vs. service pdf.

Table 1. Service pdf types used in Fig. 8.

Service type	Service uniform pdf
1	[0, 2]
2	[0, 3]
3	[0, 4]
4	[1, 4]
5	[2, 4]
6	[3, 4]

fridaya@uef.fi

Arrival transitions, is varied in the list $0.1, 0.2, 0.3, 0.4$ to 0.5. As expected, a greater process contention caused by higher values of λ, augments the waiting time as shown in Fig. 7.

Figure 8 shows how ***mwt*** varies when the service uniform pdf, which provides the sojourn time in the critical section, is varied as indicated in the Table 1. In this scenario too, augmenting the average service time causes an increase in ***mwt***.

In the light of the reported stochastic analysis, it appears that the chosen Fisher model now fulfills all the requirements of a mutual exclusion algorithm.

6 Conclusions

This paper proposes an approach based on partially stochastic Time Petri Nets (psTPN) [3, 9] for modelling and analysis of stochastic real-time systems. Transitions of a psTPN model host both a time interval as in classical Time Petri Nets [7, 8] and a generally (not necessarily Markovian) distributed probability function, whose samples provide transition duration.

An original reduction process of psTPN onto Uppaal (stochastic) timed automata is proposed which enables both a non-deterministic analysis and a stochastic analysis where both timing and probability issues are enabled. The non-deterministic analysis can be very important to check preliminarily general, time-dependent properties of a system model, where non-determinism replaces stochastic durations.

As a case study, the paper considers the Fisher's time-based mutual exclusion algorithm proposed in [3]. Whereas in [3] a numerical method with enumeration of the stochastic classes of a stochastic state graph [11] is advocated, this paper claims that the use of psTPN together with Uppaal [13, 14], although potentially less accurate, it is more general and it is capable of providing probability measures which are in any case of great value from the practical point of view. In addition, the proposed approach can be applied to systems where the method in [3] is not applicable due to state explosion problems.

Prosecution of the research aims at:

- experimenting with the psTPN/Uppaal approach to complex time-dependent and stochastic systems;
- comparing the effectiveness of psTPN with respect to the higher-level modelling and verification method based on probabilistic and timed actors proposed in [5]; and
- improving the automatic translation of a psTPN model to Uppaal through a customization of TPN Designer [23].

References

1. Jafari, A., Khamespanah, E., Sirjani, M., Hermanns, H.: Performance analysis of distributed and asynchronous systems using probabilistic timed actors. AVoCS 2014, The Netherlands (2014)

fridaya@uef.fi

2. Jafari, A., Khamespanah, E., Sirjani, M., Hermanns, H., Cimini, M.: PTRebeca: modeling and analysis of distributed and asynchronous systems. Sci. Comput. Program. **128**, 22–50 (2016)
3. Paolieri, M., Horváth, A., Vicario, E.: Probabilistic model checking of regenerative concurrent systems. IEEE Trans. Soft. Eng. **42**, 153–169 (2016)
4. Horvath, A., Paolieri, M., Ridi, L., Vicario, V.: Probabilistic model checking of non-Markovian models with generally distributed timers. In: Quantitative Evaluation of Systems (QEST) (2011)
5. Nigro, L., Sciammarella, P.F.: Qualitative and quantitative model checking of distributed probabilistic timed actors. Simul. Model. Pract. Theory **87**, 343–368 (2018)
6. Jafari, A., Khamespanah, E., Kristinsson, H., Sirjani, M., Magnusson, B.: Statistical model checking of timed Rebeca models. Comput. Lang. Syst. Struct. **45**, 53–79 (2016). https://doi.org/10.1016/j.cl.2016.01.004
7. Merlin, P.M., Farber, D.J.: Recoverability of communication protocols: implications of a theoretical study. IEEE Trans. Commun. **24**(9), 1036–1043 (1976)
8. Berthomieu, B., Diaz, M.: Modeling and verification of time dependent systems using Time Petri Nets. IEEE Trans. Soft. Eng. **17**(3), 259–273 (1991)
9. Carnevali, L., Paolieri, M., Tadano, K., Vicario, E.: Towards the quantitative evaluation of phased maintenance procedures using non-Markovian regenerative analysis. Lect. Notes Comput. Sci. Springer **8168**, 176–190 (2013)
10. Younes, H.L.S., Kwiatkowska, M., Normaln, G., Parker, D.: Numerical vs. statistical probabilistic model checking. Int. J. Softw. Tools Technol. Transf. **8**(3), 216–228 (2006)
11. Vicario, E., Sassoli, L., Carnevali, L.: Using stochastic state classes in quantitative evaluation of dense-time reactive systems. IEEE Trans. Soft. Eng. **35**(5), 703–719 (2009)
12. Alur, R., Dill, D.L.: A theory of timed automata. Theoret. Comput. Sci. **126**, 183–235 (1994)
13. Behrmann, G., David, A., Larsen, K.G.: A tutorial on Uppaal. In: Bernardo, M., Corradini, F. (eds.) Formal Methods for the Design of Real-Time Systems. Lecture Notes in Computer Science (LNCS), vol. 3185, pp. 200–236. Springer-Verlag, Berlin (2004)
14. David, A., Larsen, K.G., Legay, A., Mikucionis, M., Poulsen, D.B.: Uppaal SMC tutorial. Int. J. Softw. Tools Technol. Transfer **17**, 1–19 (2015). Springer, updated on 26 January 2018, www.it.uu.se/research/group/darts/papers/texts/uppaal-smc-tutorial.pdf
15. Clarke, E.M., Grumberg, O., Peled, D.A.: Model Checking. MIT Press, Cambridge (2000)
16. Agha, A., Palmskog, K.A.: Survey of statistical model checking. ACM Trans. Model. Comput. Simul. **28**(1), 6:1–6:39 (2018). http://doi.acm.org/10.1145/3158668
17. Gafni, E., Mitzenmaker, M.: Analysis of timing-based mutual exclusion with random times. SIAM J. Comput. **31**(3), 816–837 (2001)
18. Bucci, G., Fedeli, A., Luigi Sassoli, L., Vicario, E.: Timed state space analysis of real-time preemptive systems. Trans. Softw. Eng. **30**(2), 97–111 (2004)
19. Cicirelli, F., Nigro, L.: Modelling and verification of mutual exclusion algorithms. In: Proceedings of IEEE/ACM 20th International Symposium on Distributed Simulation and Real Time Application, pp. 136–144 (2016)
20. Dijkstra, E.W.: Solution of a problem in concurrent programming control. Comm. ACM **8** (9), 569 (1965)
21. Cicirelli, F., Nigro, C., Nigro, L.: Qualitative and quantitative evaluation of stochastic Time Petri nets. In: Proceedings of 2nd International Workshop on Cyber-Physical Systems (IWCPS'15), Lodz, Poland, pp. 775–784 (2015)
22. Kwiatkowska, M.Z., Norman, G., Parker, D.: PRISM 4.0: Verification of Probabilistic Real-Time Systems. In Proceedings of CAV 2011, pp. 585–591 (2011)
23. Carullo, L., Furfaro, A., Nigro, L., Pupo, F.: Modelling and simulation of complex systems using TPN Designer. Simul. Model. Pract. Theory **11**(7–8), 503–532 (2003)

fridaya@uef.fi

Strengthening Virtual Learning Environments by Incorporating Modern Technologies

Kashif Laeeq(✉) and Zulfiqar Ali Memon

National University of Computer and Emerging Sciences, Karachi, Pakistan
Kashiflaeeq@yahoo.com

Abstract. The current Virtual Learning Environments (VLEs) are significantly successful in empowering the organization of e-learning, however, lacking in giving personalization in learning. The plan of most VLEs is the course and instructor driven and not intended for students to build up individual pathways in learning. New transformative services, administrations, and modern technologies are now being used with learning. This paper gives a model of how the Personal Learning Environment (PLE) and VLE can mesh together for an enhanced e-learning environment with the support of modern technologies. The model has introduced some novel concepts to enhance existing VLE such as Fellows-Inspired Learning (FIL), the Extent of Friendship (EoF) scheme for friends' suggestion, real-time querying support, and Scavenge (an enhanced search system) with e-learning environment. A prototype implementation of the proposed learning environment is also presented in this paper.

Keywords: Virtual learning environment · Personal Learning Environment (PLE) and VLE integration · Enhance VLE · Modern technologies in learning

1 Introduction

The rich application of various technologies in learning is known as electronic-learning (e-learning). In an e-learning environment, learning platforms are considered one of the most vibrant tools, also known as the Learning Management System (LMS), or Virtual Learning Environment (VLE). The major utilization of VLEs is to provide instructors and learners with a set of tools for the fulfillment of various needs of the learning process and also managing them. Nowadays, the importance of the VLEs is vital in education especially in a higher educational setting and even in large business organizations [1, 2]. However, in spite of this large utilization, the existing VLEs have not resulted in the improvements in education which might be expected [3]. The five main reasons for this issue have been derived from literature as (1) limited features of VLEs are used [2, 4], (2) no support provided to informal and PLE [1], (3) limited opportunity for collaboration [5], (4) not evolved with modern technologies [6], and (5) poor query-support [7].

Today's learners have a wide range of personal links, collaboration tools and resources, both outside and within the institution to acquire personal learning objectives. Many modern technologies, including social networking sites (SNSs), Artificial Intelligence (AI) and PLEs are available to enhance existing VLEs. This paper explores

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 994–1008, 2019.
https://doi.org/10.1007/978-3-030-22868-2_67

fridaya@uef.fi

how modern technologies can leverage the learning activities in e-learning platforms. The RAL learning environment is being proposed in this paper. The RAL (articulate as real) is determined by taking the common letters from VLE and PLE and created as RAL. The proposed learning environment provides the features of PLE within VLE with additional support of the AI-based software agent (AI-assistant) to do the various tasks to enhance the current e-learning environment. The model also presents a new method of searching and maintains a search history of each student in the form of metadata that further utilizes by a recommender system. AI-based assistant (a software agent) performs many activities to facilitate students during e-learning activities. The novel concept of Fellows-Inspired-Learning (FIL) is also introduced in this model that is an initiative to encourage different learners to know about new learning features use by other learners and also know the activities of other learners in VLE. A novel method of Extent of Friendship between two students is introducing for the recommendation of collaboration. The RAL learning environment is an effort to resolve the aforementioned five issues that hinder the usability of existing VLEs. The proposed system provides an enhancement in existing e-learning environment and increases students' engagement, retention, and collaboration with e-learning.

2 Theoretical Background

The learning theory mentions a framework that discusses the change in learning happens [8]. The most vibrant learning theories are Cognitivism, Behaviorism, Connectivism, and Constructivism. The Connectivism theory is considered the digital age learning theory [9]. In literature, some discussions can be observed regarding the debate which learning theory is the best suited for PLE. As in [11] the author describes that both learning theories, Connectivism and Constructivism, are suitable for PLE. According to [12], the Connectivism learning theory is the best suitable theory for PLE. As in [13], "The online environment is one application that has been important for the development of Connectivism." In [14] the author suggested considering Connectivism as a principal learning theory of personalization in learning. By considering the experts view about which learning theory is well-matched with PLE, it can be stated that the PLE is based on the principles of Connectivism learning theory. The Connectivism principles of learning theory adequately support the proposed learning environment. The basic principles of the Connectivism learning theory are [10]:

- Development and sustaining connections is essential to continual learning.
- Many information sources stimulate the learning process
- Learning may live in non-human entities.
- The core learning skill is to develop connections between ideas, field, and concept.
- The purpose of all Connectivism learning activities to acquire accurately, and current knowledge.

fridaya@uef.fi

3 Learner-Centered vs. Teacher-Centered

The learner-centered systems are more personalized and collaborative. Nowadays, the learner-centered approach is more popular in open and distance learning because of its flexible nature. While learner-centered systems increase attention, the traditional and teacher-centered approaches have reduced in reputation. The main reason for these declining trends in the teacher-centered approach is an inflexible learning environment. In this context, e-learning platforms are incapable of delivering many of the interactive web 2.0 features to learners.

Today's learner demands and requirements are changed with respect to the change in technologies. Now learners want a more flexible learning environment, instead of the legacy learning system, such as PLE. It is an innovative idea in which students can accomplish their own learning process, communicate and collaborate with others and achieve personalized learning. In addition, the PLEs enable students to continue learning-process after completed the formal courses and established lifelong learning. The students could advantage from not only one source of knowledge, but also from plentiful of web 2.0's technologies, extensions, and add-ons. As in [15], the PLE doesn't provide any application, but it is a new approach that supports new technologies for learning. Similarly, in [17], the author explains that the features of the PLE may achieve by integrating the existing technologies, applications, devices, and services. In contrast to the traditional VLEs, the PLE embrace a student-centric learning methodology, allowing students to enhance their learning process via various available entities. As in [16], the PLEs are digital-spaces in which learners have the abilities to creates, aggregate, stores, access and share all the resources that help in the learning process. The learning resources include such artefacts as notes, lectures, assignments, wikis, forums, blogs, and other social networking sites (SNS).

4 RAL Learning Environment

The RAL Learning Environment (RLE) is being proposed in this paper. Java Agent Development Framework (JADE) was used for the design and development of agent-based applications. A JADE is a FIPA compliance specification for the interoperable multi-agent system (MAS) [31]. The system used Artificial Intelligence Markup Language (AIML) as a query language and rule-based support. The Firebase Real-time Database (FRD) used for the cloud-hosted and real-time support that structured as a JSON tree. For Java programming, the Eclipse Integrated Development Environment (IDE) is used. Figure 1 shows the architecture of the proposed RAL-learning environment.

The proposed learning environment doesn't a substitute for the existing VLE, but expands it, to focus on six dimensions of core components. All these components must be built to strengthen the existing e-learning platforms. Each component of the RAL learning environment is being discussed separately in the next section.

fridaya@uef.fi

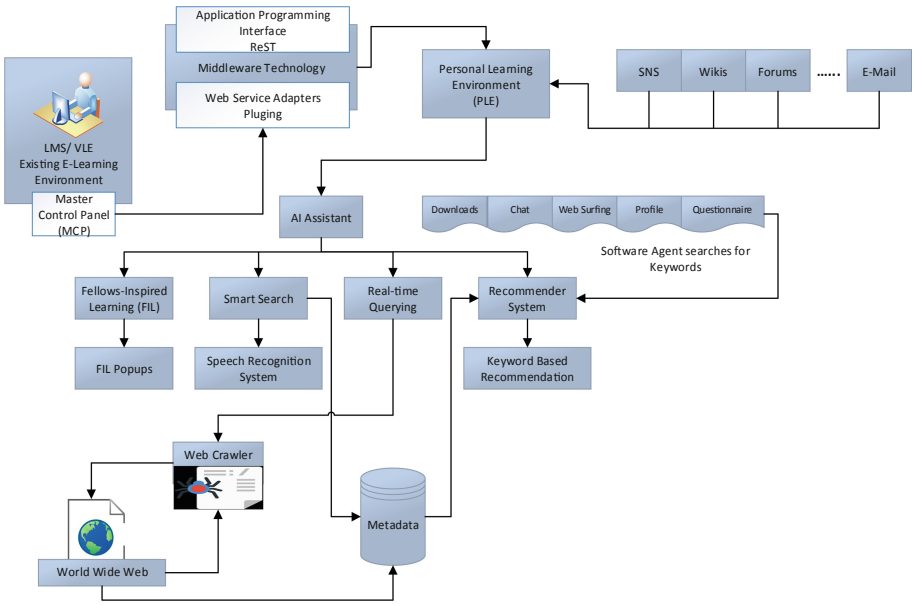


Fig. 1. The architecture of the proposed RAL learning environment

4.1 Personalized Learning Through SNS

A Virtual Learning Environment (VLE) is an accumulation of incorporated software components, empowering the administration of e-learning. The quality of VLE is indispensable for the accomplishment of e-learning goals and responsible to engage students during learning. But the in-hand VLEs are not altered according to modern technologies, thus the attributes of existing VLEs do not completely meet the requirements of today's e-learning. The students' interest, participation, and engagement with web-based e-learning platforms are on a declining trend. In contrast, the learners' engagement and participation are vibrant in many modern technologies, including social networking sites (SNS) and web 2.0 tools. Many studies have highlighted various inadequacies in existing VLEs and also suggested to include the features of modern learning support technologies in an existing e-learning environment. Design and development of new VLEs with all features of learning support modern technologies, including SNSs, online tools, and web 2.0 supportability could be one of the options. Another smart solution for an effective e-learning environment could be the integration of PLE with existing VLE. The open APIs for various technologies are not available so that the smooth integration of external technologies with e-learning platform can be one of the challenging issues. The integration of SNSs and external tools with VLE has followed ReST architecture [18]. This architecture allows VLE to work without any intermediaries, with the application programming interface (API) of various services. The model allows using some adapters where ReSTful APIs were not applicable and also design new adapters with middleware technology. The adapters are software programmed that specially designed to incorporate various external learning

fridaya@uef.fi

support technologies that enhance VLE operations and provided support where ReSTful APIs are not applicable. Other external applications that follow Content Management Interoperability Services (CMIS) standard [19] are integrated with the VLE. For example, a study [20] discussed how to design and develop middleware technologies for an integrated model, these steps are followed by designing middleware layer in the proposed model. To incorporate the cloud-data storage (cloud computing) with e-learning, the model is used, Microsoft Azure AD to bring user access, enable Single Sign-On (SSO) and a provision user account with Dropbox. The Microsoft Azure also provides the data stored in Dropbox can also be viewed and edit with other Microsoft tools, including office online, Word and Excel. The main idea of cloud computing integration with an e-learning environment is followed by the scheme discussed by [21]. Figure 2 shows the architecture of the integration of VLE and PLE.

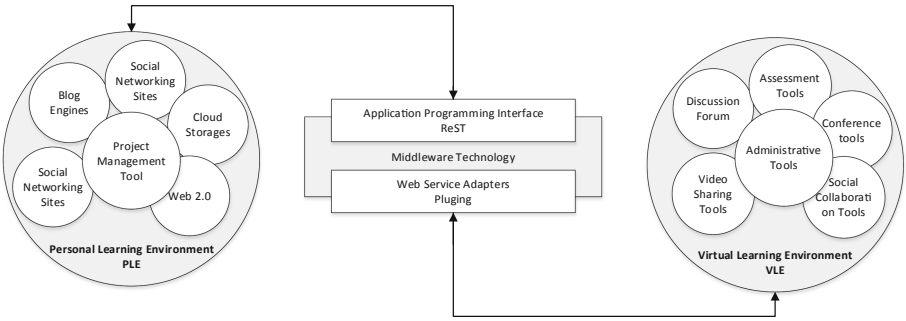


Fig. 2. Incorporation of PLE with existing VLE

4.2 Fellows-Inspired-Learning

As in the early discussion of this paper we have highlighted that limited features of LMS have been using by instructors and learners. Many important features remain unexplored throughout the learning period. The Fellows-Inspired-Learning (FIL) is a novel method to overcome this problem. The basic idea of FIL method is to encourage learners to utilize various unexplored features of the VLE by getting inspiration to other friends or fellows. Figure 3 shows the basic concept of Fellows-Inspired-Learning.

A list of basic and advanced features have provided to the AI-based software agent. Whenever any user uses a feature that belongs to the advance-feature list, the agent immediately pop-up (FIL-Pop-ups) to other users of LMS to use that feature. One of the inspirations of FIL-Popups is from the viral contents that are spread rapidly through online sharing that is a great way to bring attention to many other users. In Fig. 3, we explain how a FIL-Agent will find a new feature that used by a VLE's user and broadcast it to other users of that VLE. Suppose a user A using a new advance-feature X of VLE, at the same time, a dedicated software agent, FIL-Agent, will observe that a new advance-feature of VLE is being used by someone. The FIL-Agent immediately broadcast FIL-Popups of X to all other users. The user has right either open the pop-ups and further explore that feature or click the option to send it later. In this way, the

fridaya@uef.fi

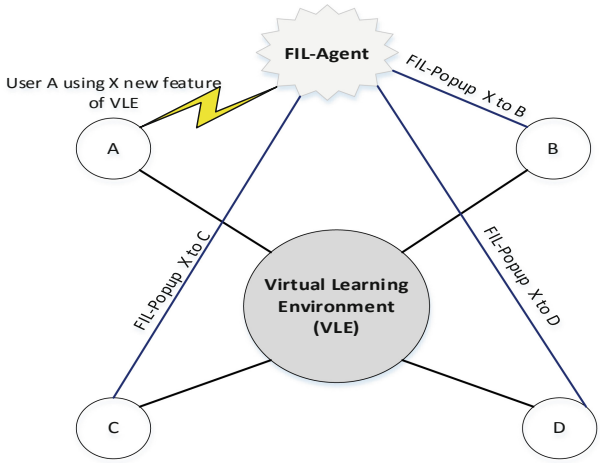


Fig. 3. Fellows-Inspired-Learning (FIL)

probability to explore and use many hidden features (options) of VLEs will be increased.

4.3 Smart Searching

In RAL learning environment we have introduced a search strategy, inspired by Microsoft Cortana [22], a virtual assistant for Windows 10. We named the search system as Scavenge, a virtual assistant to recognize the natural voice without the requirement for keyboard input and answer questions using information from the popular search engines. The Scavenge is a smart search system that can search either within the LMS environment or use the external search system to get the search result. The Scavenge window appears within the VLE and learners can search either by speaking or writing, the speech recognition system converts the speech into words or

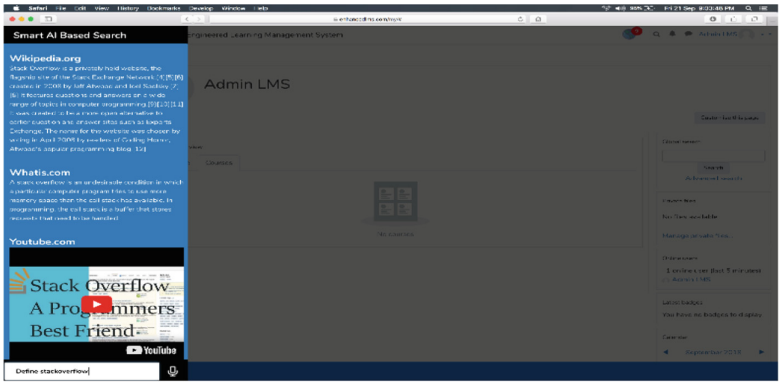


Fig. 4. Screen-shot of Scavenge search in the Moodle environment

fridaya@uef.fi

sentence to search the required result. The AI-based software agent will maintain a search history in a form of metadata document for individual users. This metadata will further utilize by a recommender system, to generate keywords for search history and to develop a collaboration of the same minded learners. The prototype of Scavenge is developed that currently available only in the English language, but it can be further extended to understand other languages too. Figure 4 is a screen-shot of the Scavenge searching within Moodle VLE.

4.4 Real-Time Querying Support

Literature suggested that real-time query support may increase students' engagement and motivation with e-learning activities. The proposed RAL learning environment provided real-time querying support. We designed a software agent (guided web crawler) named it as Query-Assistant (QA). The QA is a program that browses the web in an automated manner for the purpose of providing an instant and relevant answer to the question. The QA designed to ensure the learners to fetch the best answers within no time. The QA search and fetch answers from multiple sites but more focused on the sites that have provided sensible answers to the questions. In prototype design, initially, the QA has powered by five best sites to fetch answers to various queries. The five most powerful sites that have listed in QA search list including Stackoverflow.com, ASK.com, Answers.com, Answerbag.com, and Answers.yahoo.com. The students submit their queries to QA that searches one by one among five listed sites, collect the answers and show them within VLE (Moodle).

4.5 Recommendation for Extent of Friendship (EoF)

The proposed system has introduced a keyword-based recommendation system for better group formulation and collaboration of the like-minded students. The method of recommendations based on learner's keyword similarity index calculation. The AI-based software agents are dedicated for fetching the particular keywords from students' log file that is maintained in VLE such as, the Moodle maintains the activity log file in the Moodle log file discussed in [26]. The students' activities including search history, downloads, chat forum discussions, profile and agent-based questionnaire support to mining the keywords. At the time of course registration, students have to fill a list of 10 questionnaires, the agent further extracts some keywords from answers that describe the students' preferences and interest. The keywords can be extracted using a number of ways; here we utilized AlchemyAPI [23] to extract keywords from multiple sources. The AlchemyAPI uses machine learning or deep learning to extract the desired knowledge. For matching the similar data from the list we used Fuzzy Matching Algorithm discussed in [30]. The screen-shot of EoF scheme shows in Fig. 5.

The software agent calculates the similarity index between the student 'A' and student 'B' as follows:

- Create a set of keywords (KWs) for each student.
- Then, the agent looks for the common KWs between student 'A' and student 'B'. This can be completed using the intersection concept in set theory.

fridaya@uef.fi

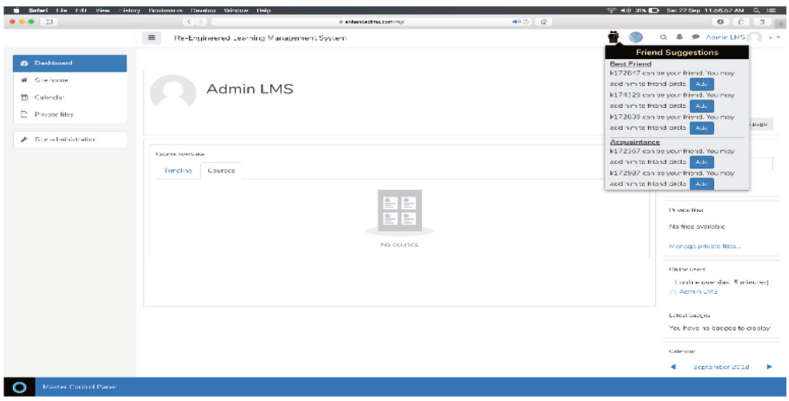


Fig. 5. Screen-shot of Extent-of-Friendship (EoF) scheme

- The similarity index (SI) of student A with student B can be calculated as:

$$SI\ (U_A, U_B) = \left(\frac{\text{Intersection}\ (U_A, U_B)}{\text{KWs}((U_A)}\right) * 100$$

where KWs (U_A) is the number of all KWs (actions) performed by student A in a period of time and Intersection (U_A, U_B) is the common KWs (actions) performed by student A and B. The minimum threshold is one week to calculate rational SI between two students therefore, SI can be calculated at-least in a week or more to decide the students' similarity. The recommendation of a friend, acquaintance, and stranger according to the following rule:

- If SI > 50% then: student 'A' can be a good friend of student 'B'.
- If 10% $\leq$ SI $\leq$ 50% then: student 'A' can be an acquaintance of student 'B'.
- If SI < 10% then: student 'A' is a stranger for student 'B'.

The system recommends a list of friends and acquaintances to student 'A' to add them to his/her learning group.

4.6 Master Control Panel

The integration of PLE, software agent and other components with existing VLE are smooth and not create a mess for users. A user can only see a small panel on the main screen of a VLE, we named it as Master Control Panel (MCP). MCP is the main intermediary component that is responsible to connect students' PLEs and other components with VLE. A user can also be enabling/disabling any features, and changing accessibility through this panel that enhances personalization concept. This makes the learning environment as a more flexible and user-centric learning environment. As a prototype implementation, we use limited features with MCP. Figure 6 is a snapshot of MCP and Fig. 7 shows the proposed RAL learning environment implemented with Moodle VLE.

fridaya@uef.fi

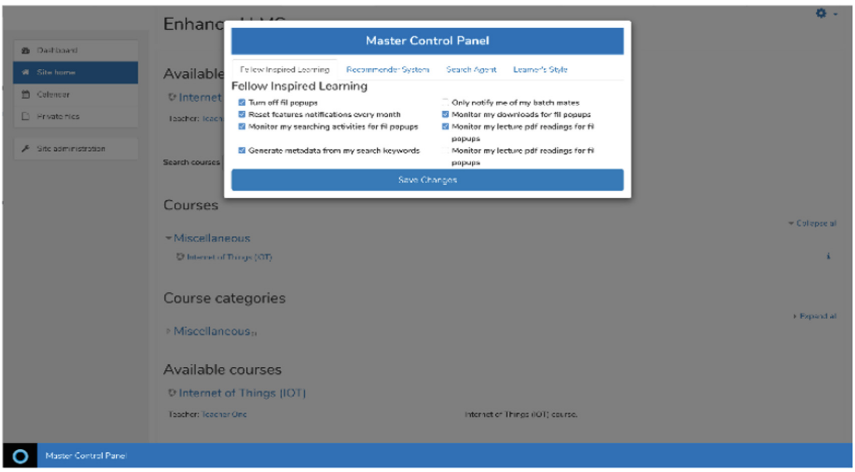


Fig. 6. Screen-shot of Master Control Panel

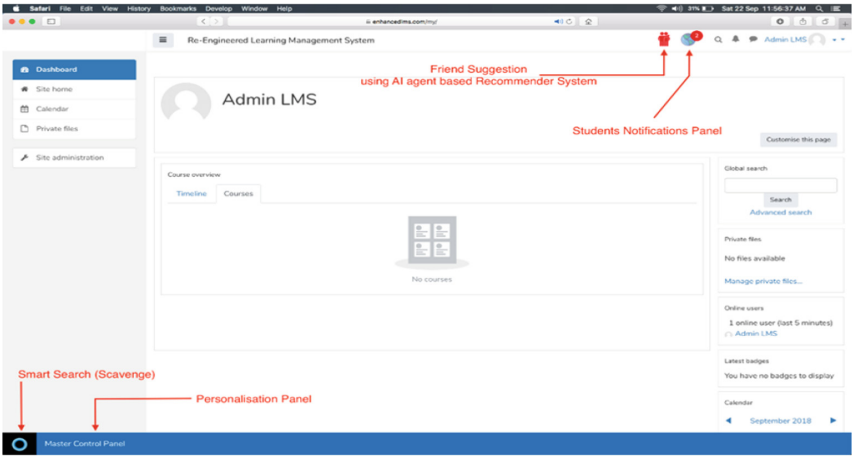


Fig. 7. Screen-shot of the proposed RAL learning environment with Moodle

5 System Usability Testing

The usability testing conducted in this paper based on ISO/IEC 9126-4 usability metrics that are effectiveness, efficiency, and satisfaction [24]. The effectiveness can be defined as the completeness and accuracy to achieve the desired goal (task). The efficiency is the resources (time) used in relation to the completeness and accuracy with which users achieve their goal. The satisfaction is a mental status that comes with suitability (comfortability) and ease of use [25]. We can evaluate effectiveness as a percentage of tasks completed or achieved per unit of time. Efficiency is a percentage of

fridaya@uef.fi

error or time to finish one desire goal. The satisfaction evaluates as a percentage of users' favorable and unfavorable comments or rating scale [24].

We have selected fifty students from the fourth semester of a bachelor of computer science program and randomly placed them in two different groups with their class ID, each group consists of twenty-five students (girls and boys). One group of students utilize existing Moodle (Moodle 3.4) learning environment, we call it as MG-group. Other group makes use of the proposed learning environment (RAL learning environment), we represented this as RG-group. After asking permission from teachers we have recommended RG-group to use our proposed learning environment and rest of class (MG-group) remains using existing Moodle learning environment. After one week of class activities, we have conducted a usability test in a lab environment. This usability test was based on task scenarios that have been described to both groups before testing.

5.1 Task Scenario

The data collected for the usability test were investigated in the form of multiple task scenarios. These tasks were design by taking common characteristics of a learning environment. Five common and easy to understand tasks were given to the students of both groups. The students asked to complete different tasks one by one. The time to finish one task was 10 min. The log of activities in this usability test was saved in the Moodle log file [26]. This action logs recorded by both learning environments have been used as metadata for further usability metrics analysis. Five task scenarios are as follows:

- Task 1: Search the course "fundamental of programming".
- Task 2: Download the first lecture of "fundamental of programming".
- Task 3: Open the chat forum.
- Task 4: Go-to the assignment page and submit a file.
- Task 5: Go-to the last week lecture and download it.

5.2 Usability Metrics for Effectiveness

The effectiveness can be defined as the completeness and accuracy to achieve the desired goal. We have evaluated effectiveness as a percentage of tasks completed or achieved per unit of time. One of the ways to evaluate the completion rate is to assign binary value 1 if the task completed and assign binary 0 if the goal does not finish. The formula as represented by 1 is used to calculate effectiveness [27].

$$Effectiveness = \frac{Number\ of\ tasks\ completed\ successfully}{Total\ number\ of\ tasks\ undertaken} \times 100\% \quad (1)$$

We analyzed the activity log file of both group and perform some calculation. Initially, the MG-group has been selected and observed the activities to find the effectiveness. We observed that among twenty-five students, only fifteen were able to achieve (finished) all five given tasks. While the rest of 10 students were unable to

fridaya@uef.fi

complete all given tasks. The overall effectiveness of this group can be calculated with the help of formula-1, as follows:

$$Effectiveness = \frac{15}{25} \times 100\% = 60\%$$

Similarly, we observed the activities of RG-group and find that twenty students of this group have finished the given tasks at a satisfactory level. Only five students were unable to achieve the desired goal. The effectiveness of this group can be calculated as:

$$Effectiveness = \frac{20}{25} \times 100\%$$

5.3 Usability Metrics for Efficiency

The efficiency can be defined as, the resources used to complete a task. These resources may be the time that a user takes to finish (successfully complete) a given task. One of the ways to find time-based efficiency is to calculate the task time by subtracting the start time of a task with task completion time. The following formula is best suited to calculate the efficiency [27]:

$$Efficiency = \frac{\sum_{y=1}^{p} \sum_{x=1}^{q} \frac{R_{xy}}{t_{xy}}}{pq} \tag{2}$$

Where:

p = Total number of participants (students).
q = Total number of goals (given tasks).
R_{xy} = Result of goal 'x' by student 'y' (R_{xy} = 1, if a student completes a goal, R_{xy} = 0, if a student does not complete the goal).
t_{xy} = Time consumed by student 'y' to complete goal 'x' or quits the goal

Table 1 shows a summary of the task completes, time to complete a task and how long a student trying to complete the task or quit without finishing the task. The value of R_{xy} place one if a student finishes the desired goal(s), and if a student does not achieve the desired goal or quit the task without completion then R_{xy} considers as zero.

For finding the efficiency we can calculate the task-1 efficiency of the MG group by putting the values from Table 1 to formula 2 as follows:

Taking P (the total number of participants) as 25 and q (total number of goals) as 1.

$$Efficiency = \frac{\left(\frac{1}{320} + \frac{1}{400} + \frac{0}{410} + \ldots + \frac{1}{278} + \frac{0}{440}\right)}{1 \times 25}$$
$$= 0.0023 \text{ goals/sec} = 0.138 \text{ goals/min}$$

fridaya@uef.fi

Table 1. Summary of task-1 completed time (in second) of MG-group

MG-group	
Student 1: R_{xy} = 1 and T_{xy} = 320	Student 14: R_{xy} = 1 and T_{xy} = 248
Student 2: R_{xy} = 1 and T_{xy} = 400	Student 15: R_{xy} = 0 and T_{xy} = 430
Student 3: R_{xy} = 0 and T_{xy} = 410	Student 16: R_{xy} = 1 and T_{xy} = 340
Student 4: R_{xy} = 1 and T_{xy} = 300	Student 17: R_{xy} = 1 and T_{xy} = 398
Student 5: R_{xy} = 1 and T_{xy} = 305	Student 18: R_{xy} = 0 and T_{xy} = 455
Student 6: R_{xy} = 1 and T_{xy} = 278	Student 19: R_{xy} = 0 and T_{xy} = 520
Student 7: R_{xy} = 1 and T_{xy} = 400	Student 20: R_{xy} = 1 and T_{xy} = 340
Student 8: R_{xy} = 0 and T_{xy} = 430	Student 21: R_{xy} = 1 and T_{xy} = 321
Student 9: R_{xy} = 1 and T_{xy} = 240	Student 22: R_{xy} = 0 and T_{xy} = 489
Student 10: R_{xy} = 1 and T_{xy} = 389	Student 23: R_{xy} = 1 and T_{xy} = 210
Student 11: R_{xy} = 1 and T_{xy} = 256	Student 24: R_{xy} = 1 and T_{xy} = 278
Student 12: R_{xy} = 0 and T_{xy} = 467	Student 25: R_{xy} = 0 and T_{xy} = 440
Student 13: R_{xy} = 1 and T_{xy} = 231	

Similarly, the same calculation has been performed with the remaining tasks of both groups. The final values of this calculation have been presented in Table 2.

Table 2. Summary of task completion time (in minutes)

Tasks	MG-group (goals/min)	FG-group (goals/min)
Task-1	0.138	0.870
Task-2	0.140	0.8
Task-3	0.320	0.38
Task-4	0.212	0.500
Task-5	0.178	0.830
Average tasks (goals/min)	0.198	3.38

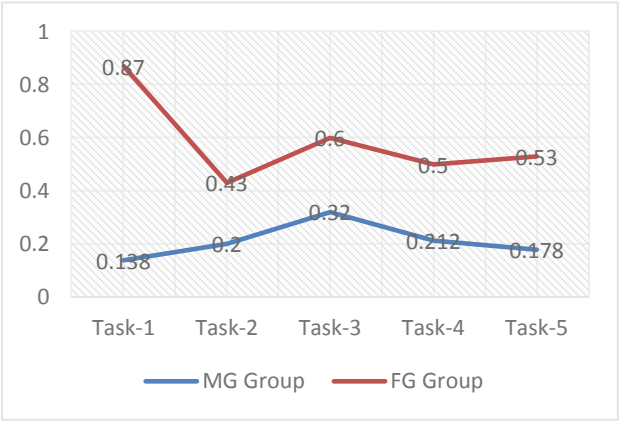


Fig. 8. The number of tasks completed in one minute

fridaya@uef.fi

5.4 Usability Metrics for Satisfaction

A set of 10 questionnaires have been designed based on the System Usability Scale (SUS) (Brooke J. 1986) [28]. These questionnaires were distributed to students of both groups to know their views about the system. This help to measure students' impression of the overall ease of use of the system (existing and proposed) being tested. SUS is also preferred in the research community because it has been found to give very precise results. Also, it consists of a very easy scale that is simple to manage the participants, thus making it perfect for usage with a small group of the participant (sample sizes). The scales were designed by utilizing five points Likert Scales, namely, (1) Strongly disagree, (2) Disagree, (3) Neutral/no opinion, (4) Agree, (5) Strongly agree [29]. Figure 8 shows the 10 questionnaires of SUS.

We performed some calculations on the data set received by the students of both groups. To make an easy and more understandable form we have calculated the overall average of all selected values of the Likert scale for both groups. The overall mean value of the MG group is derived as 2 and FG group as 4. The result clearly shows that students of FG-group were more satisfied with their virtual learning environment as compared to the students of FG-group. The complete usability test result is presented in Table 3.

Table 3. A complete usability test result of this study.

Usability attributes	Measuring instrument	Measurement	Traditional Moodle	Enhanced learning environment
Effectiveness	Task scenario	Task completion rate	60%	80%
Efficiency	Task scenario	Total number of goals/min	0.198 goals/min	3.38 goals/min
Satisfaction	Task scenario	Students' satisfaction level	02	04

6 Conclusion

The RAL learning environment has been proposed in this paper with many innovative ideas to enhance the usability of the existing VLEs. It has also been introduced many novel schemes to increase students' learnability, motivation, engagement and collaboration with e-learning activities. The modern trends and technologies in learning were highly involved in the design and development of the proposed learning environment. The RAL learning environment was an initiative taken by this paper to enhance the existing VLE. A prototype implementation and student test results validate the effectiveness of the proposed system and ensure that it can be extended in large-scale. Usability metrics effectiveness, efficiency, and satisfaction were investigated. The system usability assessment revealed that the proposed system significantly enhanced the legacy e-learning system.

fridaya@uef.fi

References

1. Melo Filho, I.J.d., Carvalho, R.S., Tavares, E.L.C., Gomes, A.S.: Towards the formative accompaniment in e-learning: conception of a social mechanism between the educational contexts LMS and PLE. In: Proceedings of World Conference on E-Learning, New Orleans, LA, USA, pp. 1330–1339 (2014)
2. García-Peñalvo, F.J., Conde, M.A., Alier, M., Casany, M.J.: Opening learning management systems to personal learning environments. J. Univ. Comput. Sci. J. UCS **17**(9), 1222–1240 (2011)
3. de oliveira, C., Jose, C., Keiko, M.: Learning management systems (LMS) and e-learning management: an integrative review and research agenda. J. Inf. Syst. Technol. Manage. (2016). http://dx.doi.org/10.4301/S1807-17752016000200001
4. Sakai-Pilot: Sakai Pilot Evaluation Final Report. University of North Carolina (2009)
5. Wesch, M.: From Knowledgeable to Knowledgeable: Learning in New Media Environments (2009)
6. Pilli, O.: LMS vs. SNS: can social networking sites act as a learning management systems? Am. Int. J. Contemp. Res. **4**(5), 90–97 (2014)
7. Son, J., Kim, J.D., Na, H.S.: A social learning management system supporting feedback for incorrect answers based on social network services. Educ. Technol. Soc. **19**(2), 245–257 (2016)
8. Smith, M.K.: Learning Theory. The Encyclopaedia of Informal education (1999)
9. Siemens, G.: Connectivism: a learning theory for the digital age. Elearnspace (2004). http://www.elearnspace.org/Articles/connectivism.htm
10. Hussain, F.: E-learning 3.0 = e-learning 2.0 + web 3.0? In: IADIS International Conference on Cognition and Exploratory Learning in Digital Age (CELDA) (2012)
11. Skrabut, S.: Personal Learning Environments: The Natural Way of Learning (2008). http://www.uwyo.edu/skrabut/docs/ADED5050_project.pdf
12. Drexler, W.: The networked student model for the construction of personal learning environments balancing teacher control and student autonomy. Australas. J. Educ. Technol. **26**(3), 369–385 (2010)
13. Kop, R., Hill, A.: Connectivism: learning theory of the future or vestige of the past? IRRODL **9**(3) (2008). http://www.irrodl.org/index.php/irrodl/article/view/523/1103
14. Baxi, V.: Connectivist and Constructivist PLEs (2010). http://learnos.wordpress.com/2010/09/21/connectivist-and-constructivist-ples/
15. Kop, R.: "The design and development of a personal learning environment" Researching the learning experience. Paper H4 32 presented at the European Distance and E-learning Network Annual Conference, Valencia, Spain, 9–12 June (2010)
16. Archee, R.: Reflections on personal learning environments, theory and practice. Procedia – Soc. Behav. Sci. **55**, 419–428 (2012). https://doi.org/10.1016/j.sbspro.2012.09.520
17. Halimi, K., Seridi-Bouchelaghem, H., Faron-Zucker, C.: An enhanced personal learning environment using social semantic web technologies. Interact. Learn. Environ. **22**(2), 165–187 (2014)
18. Fielding, R.T., et al.: Refections on the REST architectural style and 'principled design of the modern web architecture'. In: Proceedings 11th Joint Meeting Foundations of Software Engineering (ESEC/FSE 17), pp. 4–14 (2017). https://research.google.com/pubs/archive/46310.pdf
19. Bergstedt, S., Wiegreffe, S., Wittmann, J., Moller, D.: Content management systems and e-learning systems - a symbiosis? In: Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies, pp. 155–159. IEEE (2003)

fridaya@uef.fi

20. Alario-Hoyos, C., Bote-Lorenzo, M.L., Gómez-Sánchez, E., Asensio-Pérez, J.I., Vega-Gorgojo, G., Ruiz-Calleja, A.: GLUE!: an architecture for the integration of external tools in virtual learning environments. Comput. Educ. **60** (2013). http://doi.org/10.1016/j.compedu.2012.08.010
21. Zheng, H.: A virtual learning community based on cloud computing and web 2.0. Int. J. Comput. Sci. Issues (IJCSI) **9**(6), 361–366 (2012)
22. Goksel-Canbek, N., Mutlu, M.E.: On the track of artificial intelligence: learning with intelligent personal assistants. Int. J. Human Sci. **13**(1), 592–601 (2016)
23. https://github.com/AlchemyAPI
24. Abran, A., et al.: Usability meanings and interpretations in ISO standards. Softw. Qual. J. **11** (2), 325–338 (2003)
25. Seffah, A., et al.: Usability measurement and metrics: a consolidated model. Softw. Qual. J. **14**(2), 159–178 (2006)
26. Estacio, R.R., Raga Jr., R.C.: Analyzing students online learning behavior in blended courses using Moodle. Asian Assoc. Open Univ. J. **12**(1), 52–68 (2017)
27. Mifsud, J.: Usability Metrics – A Guide to Quantify the Usability of Any System. Usability Geek (2015)
28. Brooke, J.: SUS - a quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)
29. Allen, I.E., Seaman, C.A.: Likert scales and data analyses. Qual. Prog. **40**(7), 64–65 (2007)
30. Platenius, M.C., et al.: Imprecise matching of requirements specifications for software services using fuzzy logic. IEEE Trans. Software Eng. **43**(8), 739–759 (2017)
31. Bellifemine, F., Poggi, A., Rimassa, G.: JADE–A FIPA-compliant agent framework. JADE (2017)

fridaya@uef.fi

Knowledge Management System: Designing a Virtual Community of Practice for Higher Education

Gabriela Ziegler(✉)

Department Chair, Computer Information Science, Davenport University, Grand Rapids, MI, USA
gabriela.ziegler@davenport.edu

Abstract. This paper investigated if participating in a virtual community of practice (VCoP) would influence faculty's knowledge contribution, knowledge seeking, and the development of a knowledge repository. The research problem in this study would be of theoretical interest by comparing knowledge seeking, knowledge contribution, and developing a repository of knowledge by group (experimental and control group). The research method was a quasi-experimental design study. The results of the study are beneficial in extending the body of knowledge by quantitatively validating the importance of VCoP in knowledge seeking and contribution. Results of the study brought forward recommendations for future research, such as expanding the focus of the research to include measuring the quality of the knowledge transfer, impact of information and communications technologies (ICTs), etc.

Keywords: Knowledge management · Community of practice · Knowledge creation · Information and communications technologies · Quasi-experimental design

1 Introduction

The purpose of knowledge management is to improve the processes of creation, storage and sharing of information [57]. Knowledge-based economies worked under the assumption that people tend to work "smarter" and not "harder"; so knowledge has become one of the most important assets for the organization. Over the years knowledge had expanded at such a fast rate, it had become impossible for individuals to acquire and possess all knowledge. Faculty are not experts in all areas and networking and collaboration could be effective tools for finding information, network with experts, professional development, and to recruit and retain faculty [6, 7, 19, 25, 28] and [49].

Faculty generates large amounts of information, increasing the necessity for colleges to manage knowledge as a critical task. Universities played an important role in knowledge creation and dissemination through their research and publication activities [19, 25] and [27]. Faculty in higher education may participate in virtual communities of practice (VCoPs) to expand their knowledge base of their areas of expertise and help students achieve their academic goals [19]. Learning communities could be found

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1009–1029, 2019.
https://doi.org/10.1007/978-3-030-22868-2_68

fridaya@uef.fi

across universities around the world. College courses were led by instructors supported by a learning base. VCoPs had become learning vehicles used by faculty in higher education making knowledge sharing fundamental to effective learning. Additionally, information technology played a critical role in supporting knowledge sharing [19, 48] and [56].

This research study made three important contributions to the fields of knowledge management (KM) and communities of practice (CoPs): (1) a source for researchers to examine how VCoP could help faculty become more effective creating and sharing knowledge, (2) It aimed to extend the previous knowledge by quantitatively validating the importance of VCoP to seek and contribute knowledge in a systematic manner, and (3) give leaders an insight on recognizing the important factors that enable participation, collaboration, and knowledge sharing in communities to create new knowledge and make it available for participants.

2 Research Problem

Despite all the research supporting collaboration for knowledge creation, full-time and adjunct faculty were not collaborating and sharing their knowledge and best practices, and found themselves creating the same knowledge most of the time [6, 19, 25] and [49]. The specific problem on which this research concentrated was determining if participating in a VCoP influenced knowledge sharing, knowledge contribution, and the development of a repository of knowledge within the faculty at a private university in the mid-west of the United States of America.

2.1 Research Objective

The purpose of this Quantitative post-test only quasi-experimental study was to determine if, after participating in a VCoP, full-time and adjunct faculty shared their knowledge and experience. Participation in the VCoP was measured in the survey by assessing if participants in the VCoP used the forum for knowledge seeking, knowledge contribution, and if the participants perceived the online system was capable of tracking their participation.

2.2 Scope

The population consisted of active full-time and adjunct faculty teaching at the Lansing and Kalamazoo campuses at Davenport University [16]. The sites of the study, the Lansing and Kalamazoo campuses, were 2-year and a 4-year degree-granting and postgraduate institution. The purpose of the study was to determine if after participating in a VCoP, faculty shared their knowledge and experience, and created a repository of knowledge. This study used a control group to compare posttest results with the experimental group. Participants in the experimental group were required to participate in a VCoP. The experimental and the control group completed two posttest surveys to yield data about knowledge seeking, knowledge contribution, and creating a repository of knowledge.

fridaya@uef.fi

3 Literature Review

The theoretical framework for this research study discussed three major areas: knowledge management, communities of practice, and collaboration. The study was based on Nokata's and Takeuchi's knowledge creation theories and processes [42, 44] and [43]; and Wenger's CoP concepts [58–60] and [61].

CoPs were defined as groups of people informally coming together by their shared experience and interest on a cause [58] and [61]. When a group of people participated in an ongoing basis to share problems, concerns, knowledge, and expertise in certain areas, the group was considered a CoP [58] and [61]. A CoP could be a loosely organized and informal gathering of people with the purpose of sharing ideas during coffee breaks, lunch; etc. The main purpose for the community for a knowledge management system was to share information, experiences, documents, and knowledge regarding best practices [25, 27, 39] and [49]. CoPs had an important role for learning organizations by enhancing collaboration, sharing knowledge, and facilitating change. According to Wenger [58], CoPs had three critical elements: (1) domain, their shared identity; (2) community, the people who cared about the shared identity; and (3) practice, the knowledge developed, shared and maintained by the community [61].

CoPs had evolved over the past two decades. This evolution reflected the interactive flexibility of CoPs and their increased popularity across disciplines and sectors [41, 61] and [33].

Knowledge had been categorized in many ways, [35] categorized knowledge as tacit and explicit knowledge. Tacit knowledge as being undocumented knowledge and explicit knowledge as documented knowledge. The authors further argued that there was a need for organizations to create knowledge management systems to gather and document their tacit knowledge. Knowledge was fundamentally tacit with two dimensions: (1) the technical dimension, the "know-how" of doing things; and (2) the cognitive dimension, beliefs, and perceptions to be taken for granted. These two dimensions of knowledge made organizations "living organisms" that share what they knew [43].

The conceptual phase of CoPs, knowledge embraced the ability to put the acquired knowledge into practice. Experience at work created its own knowledge [45]. Knowledge creation and learning had become very important to acquire and maintain competitive advantage [46]. CoPs were a great repository of existing beliefs, values, and practices [32].

Technology infrastructure had given VCoPs an array of information and communication technologies (ICT) to create and share knowledge [57]. Many recognized VCoPs had been developed in the WWW and Web 2.0 working as virtual forums for seeking and contributing knowledge, and information for knowledge management [53]. Web 2.0–3.0 technologies had already transformed knowledge management and sharing in the business world and were being used to build communities of knowledge systems where organizations created and shared knowledge [53].

It is important to discuss ICTs and their impact on KM. Many agreed that the Internet was the communication channel allowing knowledge sharing on the WWW [11]. The explosion of the Internet had impacted information and knowledge distribution.

fridaya@uef.fi

Many recognized that VCoPs had been developed in the WWW and new ICTs tools such as Web 2.0/3.0, collaborative technologies 2.0, social networking tools, wikis, internal blogging, etc. had facilitated knowledge seeking, knowledge contribution, and impacted KM [53, 54] and [34]. Technologies such as web-based technologies, artificial intelligence, expert-based and decision-support systems, object oriented and relational databases, semantic networks, etc. have disrupt the KM field and are being used for knowledge creation [34].

The effects of ICT in KM and innovation were used to develop a framework for analysis [26]. This framework supported the KM processes of: socialization, exteriorization, combination and interiorization. The author described the ICTs tools supporting the different KM processes at Zara case study: socialization process supported by real-time information systems (intranet, email, video conferencing and others); exteriorization process supported by a management system based on electronic communication using the internet (intranet, email, video conferencing and others); combination process supported by integrated technology applicable to KM (management system based on electronic communications processed over the internet); and interiorization process supported by E-commerce, extranet. Researchers classified the available ICTs on their support to KM processes [8, 47] and [26]. ICTs technologies such as: e-mail; groupware, chat and blog support the transfer of knowledge, tacit or explicit [18].

Although CoPs were a good forum to participate and acquire knowledge, academics were reluctant to share their practices and decision-making processes in the CoPs. These experienced faculty members would need more than participating in the CoP to understand the importance of sharing their tacit knowledge [25]. The main reason for faculty not collaborating was that faculty in higher education had self-centered practices, and because of their primary loyalty to their department and not to the organization as a whole [25]. Research had shown that leadership was a key factor in the success of a CoP [1, 13] and [49]. Learning communities such as VCoPs and CoPs fostered leadership because of decision-making and leadership challenges when creating and managing new knowledge [13, 25, 51] and [49]. Leadership in higher education was critical to promote knowledge sharing and contribution by providing opportunities to share and transfer their knowledge [25].

4 Research Method and Design

The posttest only quasi-experimental design was the best suited method to complete a quantitative study in which researchers would test if an intervention made a difference for participants [9, 15] and [52]. The posttest only quasi-experimental design method was also a suitable method for cases in which specific knowledge existed about how the dependent variables behaved [52] and [9]. The quasi-experiment was designed to minimize outside factors and run for a length of a 15-week participation period to provide sufficient time to show significant results. This quasi-experimental design was a posttest only design because the university did not have a VCoP, making it impossible to gather pretest data. This posttest only quasi-experimental design was strengthened by adding a control group [9] and [52].

fridaya@uef.fi

A review of current literature revealed that the case study method was the most predominant method to study CoPs and VCoPs. After surveying the different research methods, it was decided that a quantitative method was more appropriate for the purpose of this research study. Before proposing the quasi-experimental design as the adequate research design for this research study, a correlational design was evaluated as a possible research design. It was determined that using a correlational design was not effective because correlational research was better suited when there was a need to relate two or more variables to assess the influence in each other, which was not the purpose of this proposed research study.

This quantitative posttest only quasi-experimental study was entirely quantitative, and measured whether participating in a VCoP (treatment) influenced seeking knowledge, contributing knowledge, and developing a repository of knowledge (dependent variables) [9] and [52]. The research study used two groups, one experimental group and one control group; and a posttest to establish the values for the dependent variables of: seeking knowledge, sharing knowledge, and creating a repository of knowledge. The diagram for the research design, Fig. 1, showed "NR" as non-randomized, "X" as treatment or the application on the independent variable, and "{01A}" referred to the posttest treatment construct [9] and [52].

Selected Control Group (Non-randomized Assignment)	No treatment	Posttest
NR		{O1A}
Selected Experimental Group (Non-randomize Assignment)	Experimental Treatment	Posttest
NR	X	{O1A}

Fig. 1. Non-randomized posttest only quasi-experimental design with one posttest.

The independent variable, participating in an online forum, was an online forum tool to support participant knowledge sharing, contribution, and developing a repository of knowledge. Researchers used Blackboard Learning 9.0. as the platform to develop the VCoP.

The experiment for this quasi-experimental design was the creation of a VCoP based on recommendations for developing a successful virtual forum: (1) the VCoP had a community leader, who led the knowledge sharing-contribution processes, (2) start up content: there was a structure to start with and it was changed as the dialog continued, (3) the VCoP had a common purpose to form community, share knowledge, and seek knowledge, (4) Scope: the end was not perfectly clear; participants should be able to continually use the VCoP after the experiment was over, (5) value of participation: contributors knew that their contributed knowledge was going to be used and that they got credit for sharing the knowledge, (6) social setting: Participants in this VCoP had the opportunity to meet face to face because they taught at the same campus, and the opportunity to build trust based on credentials, (7) appropriate technology with low barriers on entry, common to everyone participating in the study and easy to use, (8) the experiment used the Intellectual Property (IP) sharing model of community development [1, 3–5, 10, 22, 23, 27, 29, 30, 31, 36, 40, 49, 53, 55] and [63].

fridaya@uef.fi

After the selected 15-weeks of participation (Winter academic semester, 2014), and after faculty participated in the online forum, the effect of participating in the forum on the dependent variables of knowledge contribution, knowledge seeking, and the development of a repository of knowledge was assessed with two online surveys.

4.1 Research Questions

This posttest only quasi-experimental design study focused on the following research questions:

RQ1: How did participating in a VCoP affect seeking knowledge among faculty?
RQ2: How did participating in a VCoP affect contributing knowledge among faculty?
RQ3: How did participating in a VCoP affect the development of a repository of knowledge?

4.2 Hypotheses

This posttest only quasi-experimental study was guided by the following research hypotheses with their independent variables (IV) and dependent variables (DV):

$H1_0$ stated that faculty in the experimental group who participated in the online forum (IV) did not seek more knowledge (DV) than faculty in the control group that did not participate in the online forum.
$H1_a$ stated that faculty, in the experimental group, who participated in the online forum (IV) sought more knowledge (DV) than faculty in the control group, that did not participate in the online forum.
$H2_0$ stated faculty, in the experimental group, who participated in the online forum (IV) did not contribute more knowledge (DV) than faculty in the control group that did not participate in the online forum.
$H2_a$ stated that faculty, in the experimental group, who participated in the online forum (IV) contributed more knowledge (DV) than faculty in the control group, that did not participate in the online forum.
$H3_0$ stated that faculty, in the experimental group, who participated in the online forum (IV) did not create a repository of knowledge (DV) and faculty in the control group that did not participate in the online forum, created it.
$H3_a$ stated that faculty, in the experimental group, who participated in the online forum (IV) created a repository of knowledge (DV) and faculty in the control group that did not participate in the online forum, did not create it.

Population. The population was the active full-time and adjunct faculty from the Colleges of Technology, Business, Health Professions and Arts and Science that volunteered to participate in the VCoP and in the surveys. Faculty in this group were distributed by college as follow: College of Technology, 13 adjunct faculty and 2 full-time faculty; College of Business, 23 adjunct faculty and 4 full-time faculty; College of Health Professions, 3 adjunct faculty and 1 full-time faculty; and College of Arts and Science, 19 adjunct faculty and 7 full-time faculty.

fridaya@uef.fi

Sampling Frame. The sampling technique used in this research study was the convenience sampling, in which the researcher selected the participants because they were available and willing to participate. Researcher used a p = 0.5 as the statistical level of significance to use in testing group comparison hypothesis, and .80 as the power needed. The effect size was .90 as the standard for differences in the means between the two groups. Using the Lipsey table, this research study needed and had 20 faculty members for each of the two groups, the experimental and control groups [12, 15]. Because of the small size of the population, the researcher in this study requested the entire population to participate in the study. The total population recruited for participation in this quasi-experimental study were 72 adjunct and full-time faculty; 46 faculty members from the Lansing campus (experimental group) and 26 faculty members from the Kalamazoo campus (control group).

Instrumentation. The research study used two testing instruments as posttest in the subject matter of measuring knowledge seeking, knowledge contribution, and the development of a repository of knowledge. The selection/evaluation of the posttest followed [15] and [12] recommendations to effectively select/evaluate a survey instrument. A thorough instruments' validation was conducting following Moore and Benbasat recommendations (as cited in [48]. Measurement invariance was evaluated using LISREL v8.80 [48]. The authors also used Partial Least Square (PLS) to test the measurement model and structural model. Acceptable thresholds of Cronbach's Alpha greater than 0.70 criterion; Composite Reliability greater than 0.7 criterion, and Average Variance Extracted greater than 0.5 criterion were reported, as recommended for a reliable survey by [48]. External validity was addressed by performing statistical tests, confirming that their questionnaire would collect nonbiased responses [48]. Internal validity was addressed by assuring that the data collected from participants was accurate. Threats to the internal validity were based on the changes in dependent variables caused by changes in the independent variables.

The posttest used in this quasi-experimental design, was the online surveys designed by [48] to measure community members' knowledge contribution, knowledge seeking, and the development of a repository of knowledge within the VCoP. These online surveys collected data using two different questionnaires: one focusing on knowledge seeking and the other one focusing on knowledge contribution in an online community discussion forum. The surveys were field tested [48].

In this research study, [48] survey items were changed from a 7 Likert scale to a 5 Likert scale to make it shorter and easier to answer and tabulate. The survey used for the experimental group was changed to include the specific VCoP created for the experiment. The survey used for the control group was modified to reflect the faculty community (FC) at the Kalamazoo campus. The questions presented in both surveys gave relevant information about seeking and contributing knowledge in two different groups: experimental group with access to a VCoP, Lansing campus; and the control group without access to a VCoP, faculty community at the Kalamazoo campus.

Data Collection. All participants in the experimental and control groups of this research study had the opportunity to answer the posttest surveys. After 15-week of participation and after the experimental group participated in an online forum, participants from the experiment group and control group were invited to participate in the

fridaya@uef.fi

surveys. The participation of the experimental group in the online forum were recorded during a period of 15-weeks of participation. The online survey for the experimental group and the online survey for the control group were formatted using Google Docs forms, and distributed using Google email.

Validity and Reliability. Researcher addressed internal validity by the changes in dependent variables were not caused by changes in the independent variables. Additionally by ensuring that the activities experienced by all participants were the same, all participants came from each of the selected campuses. The selection process for both groups were very similar. This research design and analysis used 100% of the full-time and adjunct faculty teaching at the Lansing and Kalamazoo campuses. The same online forum was used for all participants in the experimental group, and the same forum leader. Both surveys were administered only once as posttests. Threats to external validity were addressed by making the experimental treatment convenient and available to all participants in the study regardless of gender, race, personality groups, social and economic status, or teaching department.

5 Data Analysis

The results from the experimental group were compared with the results of the control group which did not receive treatment. The hypotheses ($H1_{a,}$ $H2_a$, $H3_a$) were tested to measure invariance. A Mann-Whitney test was selected over the more common *t* test for independent means due to the small sample size ($N = 40$) [24]. A Mann-Whitney test was conducted to measure whether significant differences existed in the mean scores of two independent sample groups on both surveys answers between the control group, that only answered the survey, and the experimental group that participated in the VCoP [15].

The central problem investigated in this research study was that of an intentionally created VCoP would help faculty in higher education to share and create knowledge. Despite all the research supporting collaboration for knowledge creation, full-time and adjunct faculty were not collaborating and sharing their tacit knowledge and best practices, and found themselves creating the same knowledge most of the time [6] and [7]. The purpose of this research was to determine if, after participating in a VCoP, faculty shared their knowledge and experience more than faculty not participating in an online forum. The posttest only quasi-experimental design served as the framework for gathering, analyzing, synthesizing, and evaluating the participants' responses. The results of the study did not support the general research assumption examined, in which faculty participation in a VCoP would seek more knowledge, contribute more knowledge and create a repository of knowledge than faculty not participating in an online forum.

The results of the Mann-Whitney test showed no significant differences between the two faculty groups (p .20) for knowledge seeking; ($p = 0.29$) knowledge contribution, and ($p = 0.39$) for knowledge repository (reference Table 3).

fridaya@uef.fi

Table 1 displays the frequency counts for the following selected variables: Group, Years of Experience, Gender and Age. The study findings showed that equal numbers of faculty members did or did not participate in the online forum. Years of experience ranged from 1–2 years (5.0%) to 7 years or more (52.5%). Most in the sample (70.0%) were male. The ages of the faculty members ranged from 19–29 years (2.5%) to 52 years or older (45.0%) with the median age being 46 years (Table 2).

Table 1. Frequency counts for selected variables (N = 40)

Variable	Category	*n*	%
Group	No online forum	20	50.0
	Online forum	20	50.0
Years of experience	1–2 Years	2	5.0
	3–4 Years	6	15.0
	5–6 Years	11	27.5
	7 Years and more	21	52.5
Gender	Male	28	70.0
	Female	12	30.0
Age[a]	19–29	1	2.5
	30–40	6	15.0
	41–51	15	37.5
	52 or older	18	45.0

[a]Age: *Mdn* = 46 years.

Table 2 displayed the psychometric characteristics for the six aggregated scale scores. These scores were based on a 5-point metric: 1 = *Strongly Disagree* to 5 = *Strongly Agree*. The total score had a mean of $M = 3.40$ ($SD = 0.53$). Among the five subscale scores, the lowest mean was for knowledge contribution ($M = 2.65$, $SD = 1.03$) while the highest mean scale score was for usage ($M = 3.70$, $SD = 0.64$). Cronbach alpha reliability coefficients for the six scores ranged in size from $\alpha = .89$ to $\alpha = .96$ with the median sized alpha being $\alpha = .93$. This suggested that all scale scores had acceptable levels of internal reliability [12] (Table 2).

Answering the Research Questions. Research Question one asked, RQ1: How did participating in a VCoP affect seeking knowledge among faculty? and the related null hypothesis predicted that, $H1_0$ faculty, in the experimental group, who participated in the online forum (IV) did not seek more knowledge (DV) than faculty in the control group that did not participate in the online forum. To answer this question, Table 3 displays the results for the Mann-Whitney test comparing the two groups of faculty members. A Mann-Whitney test was selected over the more common *t* test for independent means due to the small sample size ($N = 40$) [24]. No differences were found between the two faculty groups ($p = .20$) for knowledge seeking which provided support to retain the null hypothesis (Table 2).

fridaya@uef.fi

Table 2. Psychometric characteristics for the aggregated scale scores (N = 40)

Scale score	Number of items	*M*	*SD*	Low	High	α
Knowledge seeking	2	3.15	0.94	1.00	5.00	.92
Knowledge contribution	2	2.65	1.03	1.00	4.50	.93
Knowledge repository	12	3.51	0.48	2.33	4.58	.89
Usage	12	3.70	0.64	2.42	5.00	.94
Leadership	4	3.59	0.69	2.25	5.00	.95
Total score	32	3.40	0.53	2.44	4.44	.96

Note. Ratings were based on a 5-point metric: 1 = *Strongly Disagree* to 5 = *Strongly Agree*.

Research question two asked: RQ2: How did participating in a VCoP affect contributing knowledge among faculty? and the related null hypothesis predicted that, $H2_0$ faculty, in the experimental group, who participated in the online forum (IV) did not contribute more knowledge (DV) than faculty in the control group that did not participate in the online forum. To answer this question, Table 3 displayed the results for the Mann-Whitney test comparing the two groups of faculty members. A Mann-Whitney test was selected over the common *t*-test for independent means due to the small sample size (N = 40) [24]. No differences were found between the two faculty groups (p = 0.29) for knowledge contribution which provided support to retain the null hypothesis (Table 3).

Research question three asked: RQ3: How did participating in a VCoP affect the development of a repository of knowledge? and the related null hypothesis predicted that, $H3_0$ faculty in the experimental group, who participated in the online forum (IV) did not create a repository of knowledge (DV) and faculty in the control group that did not participate in the online forum, created it. To answer this question, Table 4 displayed– the results for the Mann-Whitney test comparing the two groups of faculty members. A Mann-Whitney test was selected over the common *t*-test for independent means due to the small sample size (N = 40) [24]. No differences were found between the two faculty groups (p = .39) for developing a repository of knowledge which provided support to retain the null hypothesis (Table 3).

5.1 Additional Findings

Also in Table 3 were the Mann-Whitney tests for three additional scale scores. No significant between-group differences were found for usage (p = .94), leadership (p = .45) or the total score (p = .56) (Table 3).

fridaya@uef.fi

Table 3. Mann-Whitney tests comparing scale scores based on participation in the online forum (N = 40)

Scale score	Online forum	*n*	*M*	*SD*	r_s	*z*	*p*
Knowledge seeking					.20	1.27	.20
	No	20	3.33	0.99			
	Yes	20	2.98	0.87			
Knowledge contribution					.17	1.06	.29
	No	20	2.83	1.05			
	Yes	20	2.48	0.99			
Knowledge repository					.14	0.86	.39
	No	20	3.58	0.46			
	Yes	20	3.45	0.50			
Usage					.01	0.08	.94
	No	20	3.69	0.68			
	Yes	20	3.71	0.61			
Leadership					.12	0.75	.45
	No	20	3.63	0.55			
	Yes	20	3.55	0.83			
Total score					.09	0.58	.56
	No	20	3.44	0.54			
	Yes	20	3.35	0.53			

Note. Ratings were based on a 5-point metric: 1 = *Strongly Disagree* to 5 = *Strongly Agree*.

Table 4 displayed the Spearman correlations for the six aggregated scale scores with four demographic variables (faculty group, experience, gender, and age). Due to the small sample size ($N = 40$) and the exploratory nature of this study, findings that were significant at the $p < .10$ level were noted to suggest possible avenues for future research [24]. Four of the resulting twenty four correlations were significant at the $p < .10$ level. Specifically, female faculty members had higher scores for knowledge repository ($r_s = .28$, $p = .08$) and leadership ($r_s = .30$, $p = .07$). In addition, older faculty members had higher scores for knowledge contribution ($r_s = .41$, $p = .008$) and the total score ($r_s = .27$, $p = .09$) (Table 4).

Table 4. Spearman correlations for the scale scores with selected variables (N = 40)

Scale score	Group[a]	Experience	Gender[b]	Age
Knowledge seeking	−.20	.09	−.06	.30
Knowledge contribution	−.17	−.02	.00	.41***
Knowledge repository	−.14	.02	.28*	.08
Usage	.01	.04	.04	.25
Leadership	−.12	.21	.30*	.19
Total score	−.09	.07	.13	.27*

*$p < .10$. **$p < .05$. ***$p < .01$.

[a]Group: 1 = *no online forum* 2 = *online forum.*

[b]Gender: 1 = *male* 2 = *female.*

fridaya@uef.fi

5.2 Limitations

The sample population was based on full-time faculty and adjunct faculty from two campuses at Davenport University, and did not include all campuses, and other universities mirroring the total population. Using convenience sampling, two campuses at Davenport University were selected as the population for the study. Because of the small expected sample size 46 faculty members in the experimental group, and 26 faculty members in the control group), the researcher in the study decided to use the complete population for the study. The power statistical test was used to determine the sample and effect size [15] and [12]. This study used power of 0.80, with a $p = 0.5$, therefore this research study needed 20 faculty members for each group (according to Lipsey's table). The study received 20 survey responses for each group, toward the final data analysis and findings. The data gathered may not represent the full scope of experiences related to the use of VCoP for knowledge creation, knowledge contribution and the creation of a repository of knowledge, if the study is implemented in a different context than the original in which the study took place. Because of the scope and focus of this research study, caution should be exercised when generalizing the results to different contexts than a learning community in higher education.

Another limitation included was the threat to internal validity called History in posttest quasi-experimental design. History events that can happen during the experiment, might affect the experiment outcomes. Because of this internal threat to internal validity, a statistical conclusion may also be reduced if implementation varies from location to location. It is important to point out that people's live and circumstances change and it is very difficult to control these natural changes [9, 52] and [12].

Finally this research study did not evaluate the quality of knowledge transfer between members of the community. The study measured if members in an intentionally created VCoP, participated to seek and contribute knowledge, and to create a repository of knowledge. Measuring the quality of the knowledge transferred was outside the scope of this study.

5.3 Findings

Researchers suggested a top benefit of participation in VCoPs was to empower participants to share and contribute their knowledge. Also, researchers have concluded that organizations were more effective when members collaborated [4, 27, 37] and [58]. Research also suggested that there are advantages of using intentionally created VCoP in organizations to foster knowledge creation and collaboration [6, 7, 10, 49, 53] and [59]. Research studies based VCoPs' participation to create knowledge vary in results [2, 5, 25, 56] and [62]. This current research study did not statistically confirm faculty participating in a VCoP would seek and contribute more knowledge, and create a repository of knowledge, than faculty not participating in an online forum. There were no differences between the experimental and the control group for knowledge seeking, knowledge contribution and creating a repository of knowledge.

The researcher in this study did not find a quasi-experimental study available that supported that members participating in a VCoP would contribute and seek more knowledge, and create a repository of knowledge than members not participating in an

fridaya@uef.fi

online forum. There was abundant research supporting the use of VCoP to enhance participation and knowledge creation in higher education. To mention only a few of them, eBerea case study was a collaboration research between China and Europe based on a social network eBerea IRSIS [62]. eBerea's case study purpose was to support advanced research in higher education between Europe and China. Data was collected using a web questionnaire. Results from the survey showed that in general, the eight universities participating in the eBerea case study, collaborated in research. The author concluded that the eBerea network has been developed over the years, and participation and collaboration have been developed. Collaboration was stronger inside the origin university, and the interactions with other universities were limited even though they were part of a larger network [62]. A regional clinical fellow developed a CoP with the purpose of creating a collaborative environment in which members: (1) could learn from each other, (2) produced quality medical research, and (3) shared resources increasing the quality of education. This CoP was centered on teaching and learning to encourage faculty to share their best practices and improve quality [38]. A case study was presented which the authors developed a model for creating knowledge management systems in higher education. The authors argued that universities needed to provide a culture of knowledge sharing by offering and encouraging faculty and staff to participate in seminars and conferences related to the topic [21]. The study of a CoP used to mentor and guide faculty in their academic activities was presented [55]. The researcher concluded that new faculty that participated in well-established CoPs would be able to distinguish between conformist practices and innovative practices.

These finding supports the conclusion that although institutions of higher education had a culture of sharing, this culture was individualist, and self-serving. The collaboration tended to be within departments of the same institution. The first and second null hypotheses stated that faculty in the experimental group, who participated in the online forum, would not seek, and would not contribute more knowledge than faculty that did not participate in the online forum. The Mann-Whitney test comparing both groups presented no differences between the two faculty groups. The study also supported the third null hypothesis that participating in the online forum would not provide faculty with a knowledge repository [25].

The Spearman Correlation showed that female population had a higher agreement with creating a repository of knowledge and leadership than the male population (Table 4). Testing this relationship would have provided varying results. Another recommendation is to investigate the relationship between older faculty members and knowledge contribution. The Spearman Correlation showed that older population had higher agreement with knowledge contribution (Table 4). Testing this relationship would have provided varying results. Investigating the mentioned variables may provide more specific data for effect on participation in VCoP to seek and contribute knowledge, and to create a repository of knowledge.

Additionally, the research location was a limiting factor, as the study was conducted within the context of Davenport University, and it might not hold true in other contexts. The age, gender, and experience of the sample population did not represent the whole population. Although this study did not show a correlation between participating in a VCoP to encourage seeking and contributing knowledge, and creating a repository of knowledge, the expansion of the number of locations included in the study may provide additional data for a comparison based on gender, age and experience.

fridaya@uef.fi

6 Conclusions and Recommendations

This research study results presented a focus on the relationship of full-time faculty and adjunct faculty participating in a VCoP, and sharing their knowledge and best practices. The general question asked was if participating in a VCoP affected knowledge seeking, knowledge contribution and the development of a knowledge repository among faculty at Davenport University. None of the three hypotheses supported the ideas that faculty participating in a VCoP will seek more knowledge, contribute more knowledge, and create a knowledge repository compared to those faculty members that did not participate in the VCoP.

The study instruments focused on measuring knowledge contribution, knowledge seeking and the development of a repository of knowledge. These surveys were previously used to measure the same variables, knowledge contribution, knowledge seeking, development of a repository of knowledge and usability of the system [28]. Questions related to participating in the VCoP provided data collection for participating in the VCoP as the independent variable. Questions related to knowledge seeking, knowledge contribution, and development of a repository of knowledge provided data collection for knowledge seeking, knowledge contribution, and development of a repository of knowledge as dependent variables. Based on the data analysis, no relationship was found between faculty participating in the VCoP to seek more knowledge, and contributing more knowledge, and creating a repository of knowledge than faculty not participating in a VCoP.

Based on previous research, the challenge of effective VCoPs was creating an environment in which those communities could flourish. Creating VCoPs in which members participate to share and contribute knowledge could be based on many variables, such as: (1) leadership role in the online community, (2) content development, and quality of the knowledge transfer, (3) common purpose to join the community, (4) value of participation, (5) organizational culture of knowledge sharing and collaboration, (6) a developed technology infrastructure, and (7) integration of the VCoP in the organizational structure [1, 3, 5, 10, 20, 22, 27, 29, 30] and [63]. The current research did not provide data related to specific questions on content development and quality of knowledge transfer, usability of the technology, organizational culture, and integration with the organization structure. The results of the study are beneficial in extending body of knowledge in the topics by quantitatively validating the importance of VCoPs to seek and contribute knowledge, and developing a repository of knowledge in a systematic manner.

6.1 Recommendations for Future Research

The findings from this research study suggest further research may be needed. The benefit of using two different groups, an experimental and a control group did not specifically measure the quality or intensity of the knowledge transfer. The quality of knowledge transfer was important to knowledge creation in organizations [17]. Additionally, the study results brought two main recommendations for future research: first, to expand the study to test the relationship between the variables of gender and knowledge repository and leadership; and the relationship between the variables of age

fridaya@uef.fi

and knowledge contribution. Expanding this study to test these variables may better determine the relationship between these two categories and seeking and contributing knowledge; and creating a repository of knowledge. The Spearman correlations (Table 4) for the mentioned variables were significant at the $p < 0.10$ level suggesting possible avenues for future research [12]. The second recommendation is to expand the study scope to include different contexts than higher education so the study results can hold true in other contexts. Expanding the scope by including a more diverse group of faculty based on gender, age and experience.

Recommendations for future research on the role of leadership in developing VCoPs and nurturing them, include a study of a substantial number of VCoPs, in different organizational environments, using different research designs to test the importance of leadership and managerial practices [5]. Leadership and managerial practices are needed in intentionally formed VCoPs to create a knowledge sharing culture [49]. It is also recommended to include building leadership programs for future practitioners to promote knowledge sharing and collaboration, and understanding the benefits in using VCoPs and/or social networks as Web 2.0–3.0.

Further research should also include an investigation related to the impact of ICTs, specifically Web 2.0–3.0, in knowledge sharing and contributions. Web 2.0–3.0 has changed the way people interact with each other, search, share, and create knowledge [53].

Another recommendation is to expand the sample population. The sample population should include different age groups, gender, and experience to gain the full perspective on knowledge seeking, knowledge contributions, and creation of a repository of knowledge when participating in a VCoP. Various organizational forms, such as business, government, sports teams, and volunteer organizations, should be considered.

Two other recommendations for future research are included: One important recommendation is to build professional development programs for future practitioners to promote knowledge creation and collaboration and the understanding of the benefits of using VCoP and/or social networks. Designing training programs, either as part of career development and advancement in the organization, or as part of requirements to improve collaboration and knowledge sharing in their areas. Additional training should address the use of ICT, specifically Web 2.0–3.0, to promote knowledge sharing and collaboration. The second recommendation is to conduct future research studies to test the importance of leadership, managerial practices, and ICT in knowledge creation and collaboration. There may be external factors that were not considered in this study, such as managerial practices, trust in technology, and benefits in implementing and aligning an intentional VCoP with the current organizational structure of the organization.

Based on the literature, there are benefits in creating a VCoP to encourage knowledge sharing and collaboration. Technology infrastructure gives VCoPs a diverse set of ICT options to create and share knowledge [57]. It is important for VCoPs to create and share knowledge and best practices [50]. VCoPs support participation and foster collaborative learning [5]–[25] and [49]. VCoPs with an ingrained collaboration purpose would encourage learning, quality of the knowledge transfer and resource sharing; increasing the overall quality of education at a regional clinical fellow community [38]. However there are limited research studies that quantitatively validate the

fridaya@uef.fi

importance of VCoPs to seek and contribute knowledge in a systematic manner. Therefore future studies should quantitatively address the factors that contribute to seeking and contributing knowledge, and creating a repository of knowledge in different types of organizations.

6.2 Implications and Significance to Leadership

The results of this research study contributed to the literature as previous research studies are limited to support knowledge sharing based on participation in CoPs/VCoPs in case studies. There might be many variables that contribute to knowledge sharing, such as leadership and system usability. Future research is suggested to specify what could affect participation in social communities within institutions of higher education.

The goal of the study was to inform leadership in higher education, that participating in a VCoP would provide university settings a forum in which faculty would share their knowledge and experience. Leadership is a critical factor in developing successful collaboration within the community, and it is also fundamental for building, managing, and sustaining communities [1, 14, 49] and [22]. The results of this research study suggests further investigation is required but might be useful in providing training or leadership consideration when creating and maintaining virtual communities in which sharing knowledge is the main purpose.

The first research suggestion for future research is to conduct an investigation to specify how organizations can help VCoPs be successful vehicles for participating and sharing knowledge. Additional research parameters could include the correlation between managerial support and success of the project. Investigation is suggested to study a diverse group of VCoPs, using different research methods, in different organizational environments to determine the importance of the managerial role for the success of VCoPs. Leadership plays a critical role in creating and maintaining a healthy and effective VCoPs [14] and [49]. Although leaders in higher education promote an environment and culture for sharing, faculty are not sharing tacit knowledge because of their self-centric interest and loyalty to their academic department and not to the organization [25]. One of the main reasons for VCoPs not contributing knowledge is the lack of leadership involvement [5] and [49]. The second recommendation suggests to identify what is required to improve participation and knowledge sharing in a VCoP. Research has reported that 70% of participants considered encouraging participation in VCoPs the most challenging facilitating task [55]. It is critical for leaders to recognize the strategic importance of determining factors influencing participation, knowledge sharing and collaboration in VCoPs [11]. This may include communication programs development to manage the use of technology in VCoPs and the impact of factors such as trust in the organization and risks of losing corporate information. ICTs play a critical role in knowledge sharing and collaboration when using a VCoP for knowledge creation [21] and [49]. Improvements of ICTs, specifically Web 2.0 technology, are especially effective for collaboration and communication in any topic. Although Web 2.0 had great potential for knowledge creation, organizations are not exploiting all its capabilities for knowledge creation. Future research is needed to examine motivational factors for using Web 2.0 by organizations for knowledge creation [53].

fridaya@uef.fi

6.3 Research Reflections

Limited time and financial resources available to the researcher were limiting factors. More time to participate in the VCoP would have been beneficial before data collection which could had led to varying results. Other institutions beside Davenport University, would have also provided varying perspectives from the sample population.

The research design was also a limiting factor. The posttest only quasi-experimental study was appropriate for the research study, but the results were limited to a survey form. The benefit in participant interaction and feedback would have provided varying results. Participants shared their perspectives about collaboration and knowledge sharing in the virtual forum, advantages and limitations. The input was outside the scope of this research study but provided input for further research outside the implementation of a VCoP to seek and contribute knowledge, and to develop a knowledge repository.

Changes in technology are exponential, and particularly the impact of Web 2.0–3.0 in knowledge sharing and contribution can lead to many areas of study. The use of technology is beneficial at all levels in the organization, and are impacting how people interact with each other in different organizational contexts, such as educational, organizational, and political. The expectation in the initial phase of this research was that knowledge sharing and collaboration may be affected. The results proved interesting as proof was not established of faculty seeking more knowledge, contributing more knowledge, and creating a knowledge repository by participating in a VCoP than faculty not participating in an online forum, based on the data analysis.

6.4 Conclusion

There are many challenges with implementing a VCoP in which members participate in sharing their knowledge and best practices. Many research studies support the implementation of VCoPs, and others explore the challenges with implementation. These challenges include adequate leadership as a critical factor for implementing and maintaining an active and healthy VCoP. Organizations need to have a knowledge creating culture that enables members to freely participate. VCoPs need to be established in a well-developed technology infrastructure. Organizational structures formal or informal need to create an environment supporting participation and knowledge sharing. Organizations should also have a mechanism to measure the effectiveness of the VCoP for collaboration and knowledge sharing [1, 10, 25, 27, 29, 30, 49] and [53]. This current research study added another debate for the benefits of VCoP as tools to foster collaboration and knowledge sharing. The purpose of this research study was to compare knowledge seeking, knowledge contribution, and development of a repository of knowledge by full-time and adjunct faculty in two different campuses. The study's results supported the purpose of the study by comparing surveys' results from two groups of participants.

Acknowledgment. I thank Davenport University for supporting this research project.

fridaya@uef.fi

References

1. Albinsson, L., Curtin, G., Forsgren, O., Wall, M.: Creating and sustaining successful knowledge management in purposeful communities-summary of key experiences from pioneers. Wiley InterScience **25**, 615–626 (2008). https://doi.org/10.1002/srs.936
2. Bansler, J., Havn, E.: Exploring the role of network effects in IT implementation. Inf. Technol. People **17**(3), 268–285 (2004). https://doi.org/10.1108/09593840410554184
3. Barnard, Z., Rensleigh, C.: Investigating online community portals for enhanced alumni networking. Electron. Libr. **26**(4), 433–445 (2008)
4. Barnett, S., Jones, S.C., Caton, T., Iverson, D., Bennett, S., Robinson, L.: Implementing a virtual community of practice for family physician training: a mixed-methods case study. J. Med. Internet Res. **16**(3), e83 (2014)
5. Bourthis, A., Dube, L.: Structuring spontaneity: investigating the impact of management practices on the success of VCoP. J. Inf. Sci. **26**(2), 175–193 (2010)
6. Buckley, S., Giannakopoulos, A.: Sharing knowledge the CoP way. In: International Conference of Information Management and Evaluation, 72-VII. Reading: Academic Conferences International Limited (2011)
7. Buckley, S., Toit, A.: Sharing knowledge in universities: communities of practice the answer? Educ. Knowl. Econ. **3**(1), 35–44 (2009). https://doi.org/10.1080/17496890902858249
8. Butler, T.: From data to knowledge and back again: understanding the limitations of KMS. Knowledge and process management. J. Corp. Transform. **10**(4), 144–155 (2003)
9. Campbell, D., Stanley, J.: Experimental and Quasi-Experimental Designs for Research. Houghton Mifflin, Boston (2015)
10. Chalkiti, K., Sigala, M.: Information sharing and knowledge creation in online forums: the case of the Greek online forum "DIALOGOI". Curr. Issues Tour. **11**(5), 381–406 (2008). https://doi.org/10.1080/13683500802316006
11. Chang, H., Chuang, S.: Social capital and individual motivations on knowledge sharing: participant involvement as a moderator. Inf. Manage. **48**(1), 9–18 (2011). https://doi.org/10.1016/j.im.2010.11.001
12. Christensen, L.B., Johnson, R.B., Turner, L.A.: Research Methods, Design and Analysis. Pearson, Upper Saddle (2014)
13. Chrisentary, J., Barret, D.: An exploration of leadership in virtual communities of practice. Management **77**, 25–34 (2015). https://doi.org/10.7595/management.fon.2015.0027
14. Correia, A., Paulos, A., Mesquita, A.: Virtual communities of practice: investigating motivations and constraints in the process of knowledge creation and transfer. J. Knowl. Manage. **8**(1), 11–20 (2010)
15. Creswell, J.W.: Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. Prentice Hall, Upper Saddle River (2015)
16. Davenport University: http://www.davenport.edu. July 2014
17. Donate, M., Canales, I.: A new approach to the concept of knowledge strategy. J. Knowl. Manage. **16**(1), 22–44 (2011). https://doi.org/10.1108/1367327/1211118927
18. Dorn, E., Sahinyan, A.: Effects of Information and Communication Technologies on Knowledge Transfer: An Employee Perspective (2011). https://gupea.ub.gu.se/bitstream/2077/26737/1/gupea_2077_26737_1.pdf. Accessed 12 Jan 2018
19. Dukic, G., Jozinovic, T.: Knowledge management in university: preparedness of master of information sciences graduates from the faculty of humanities and social sciences in Osikek for the labor market. Technical J. **10**(1–2), 22–28 (2016). ISSN 1848-5588

fridaya@uef.fi

20. Easterby-Smith, M., Lyles, M.: The evolving field of organizational learning and knowledge management. In: Handbook of Organizational Learning and Knowledge Management. Wiley, Chichester (2011)
21. Eftekharzade, B., Mohammadi, B.: The presentation of a suitable model for creating knowledge management in educational institutes. Proc.-Soc. Behav. Sci. **29**, 1001–1011 (2011). https://doi.org/10.10161/jsbspro.2011.11.334
22. Erat, P., Desouza, K., Schafer-Jugel, A., Kurzawa, M.: Business customer communities and knowledge sharing: exploratory study of critical issues. Eur. J. Inf. Syst. **15**(5), 511–524 (2006). https://doi.org/10.10057/palgrave.ajis.3000643
23. Fan, S., Cooling, N., Radford, J., Fabian, D., Brown, J.: Building on the evaluation of stars: using online repositories to support the general practice learning community. Eval. J. Australas. **14**(2), 25 (2014)
24. Field, A.: Discovering Statistics Using IBM SPSS Statistics. Sage, Thousand Oaks (2017)
25. Fullwood, R., Rowley, J., Delbridge, R.: Knowledge sharing amongst academics in UK universities. J. Knowl. Manage. **17**(1), 123–136 (2013). https://doi.org/10.1108/13673271311300831
26. García-Álvarez, M.T.: Analysis of the effects of ICTs in knowledge management and innovation: the case of Zara Group. Comput. Hum. Behav. **51**, 994–1002 (2014)
27. Goel, L., Mousavidin, E.: A proposed framework for designing sustainable community for knowledge management systems. Int. J. Knowl. Manage. **4**(3), 82–100 (2008)
28. Heinrich, K., Oberleitner, M.: How a faculty group's peer mentoring of each other's scholarship can enhance retention and recruitment. J. Prof. Nurs. **28**(1), 5–12 (2012)
29. Hsu, M., Chang, C., Yen, C.: Exploring the antecedents of trust in virtual communities. Behav. Inf. Technol. **30**(5), 587–601 (2011). https://doi.org/10.1080/0144929X.2010.549513
30. Iverson, J., McPhee, R.: Communicating knowing through communities of practice: exploring internal communicative processes and differences among CoPs. J. Appl. Commun. Res. **36**(2), 176–199 (2008). https://doi.org/10.1080/00909880801923738
31. Jiménez Zarco, A.I., González González, I., Saigí Rubió, F., Torrent, S.: The Co-learning process in healthcare professionals: assessing user satisfaction in virtual communities of practice. Comput. Hum. Behav. **51**, 1303–1313 (2014)
32. Jonsson, A.: How to share knowledge through logic and practice. Knowl. Manage. Res. **13**(1), 45–58 (2015)
33. Kislov, R., Harvey, G., Walshe, K.: Collaboration for leadership in applied health research and care: lessons from the theory of CoP. Implement. Sci. **6**, 64 (2011)
34. Klimova, A.: Systematic Literature Review of Using Knowledge Management Systems and Processes in Green ICT and ICT for Greening (2018). https://doi.org/10.1007/978-3-319-60137-3_16
35. Laudon, K.C., Laudon, J.P.: Management information systems: managing the digital firm. Pearson, Upper Saddle River (2015)
36. Levy, Y., Ellis, T.: A guide for novice researchers on experimental and quasi-experimental studies in information systems research. Interdisc. J. Inf. Knowl. Manage. **6**, 151–160 (2011)
37. Li, L., Grimshaw, J., Nielsen, C., Judd, M., Coyte, P., Graham, I.: Evolution of Wenger's concept of community of practice. Implement. Sci. **4**, 11 (2009). https://doi.org/10.1186/1748-5908-4-11
38. Little, D., Butcher, K., Atkinson, S., Still, D., Vasant, J.: A regional teaching fellow CoP. Clin. Teach. **11**, 516–519 (2014). https://doi.org/10.1111/tct.12229
39. Mark, B.E., Hoholm, T., Maaninen-Olsson, E., Aanestad, M.: Changing practice through boundary organizing: a case from medial R&D. Hum. Relat. **65**, 263288 (2012)

fridaya@uef.fi

40. Mendizabal, G.A., Solinís, R.N., González, I.Z.: Hobe, a case study: a virtual community of practice to support innovation in primary care in Basque public health service. BMC Fam. Pract. **14**(1), 168 (2013)
41. Murillo, E.: Communities of practice in the business and organization studies literature. Inf. Res. **16**(1), 1–18 (2011)
42. Nonaka, I., Nishiguchi, T.: Knowledge Emergency: Social, Technical, and Evolutionary Dimensions of Knowledge Creation. Oxford University Press, New York (2001)
43. Nonaka, I., Takeuchi, H.: The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press, New York (1995)
44. Nonaka, I., Konno, N., Toyama, R.: Emergence of "Ba". In: Nonaka, I., Nishiguchi, T. (eds.) Knowledge Emergence. Social, Technical, and Evolutionary Dimensions of Knowledge Creation. Oxford University Press, New York (2001)
45. Omidvar, O., Kislov, R.: The evolution of the communities of practice approach: toward knowledgability in a landscape of practice: an interview with Ettiene Wenger-Trayner. J. Manage. Inquiry **23**(3), 266–275 (2013). https://doi.org/10.1177/10564492612505908
46. Peschl, M., Fundneider, T.: Designing and enabling spaces for collaborative knowledge creation and innovation: from managing to enabling innovation as socio-epistemological technology. Comput. Hum. Behav. **37**, 346–359 (2014). https://doi.org/10.1016/j.chb.2012.05.027
47. Phang, M., Foong, S.: Information communication technologies (ICTs) and knowledge sharing: the case of professional accountants in Malaysia. World J. Sci. Technol. Sustain. Dev. **7**(1), 21–34 (2010)
48. Phang, C., Kankanhalli, A., Sabherwal, R.: Usability and participation in online communities: a comparative study of knowledge seeking and contribution. J. Assoc. Inf. Syst. **10**(10), 721–747 (2009)
49. Rivera, G., Rivera, I.: Design, measurement and knowledge management model in the context of a Mexican university. Innovar **26**(59), 21–34 (2013)
50. Ruberg, L., Cumming, M., Piecka, D., Ruckman, C., Seward, R.: A logical approach to supporting professional learning communities. Knowl. Manage. E-learn.: Int. J. **3**(4), 599–620 (2011)
51. Ryman, S., Richardson, R.: Creating and sustaining online learning communities: designing environments for transformative learning. Int. J. Pedagogies Learn. **5**(3), 46–58 (2009)
52. Shadish, W., Cook, T., Campbell, D.: Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Cengage, Belmont (2002)
53. Sigala, M., Chalkiti, K.: Investigating the exploitation of Web 2.0 for knowledge management in the Greek tourism industry: an utilization-importance analysis. Comput. Hum. Behav. **30**, 800–812 (2014). https://doi.org/10.1016/j.chb.2013.05.032
54. Soto-Acosta, P., Cegarra-Navarro, J.: New ICTs for knowledge management in organizations. J. Knowl. Manage. **20**(3), 417–422 (2016)
55. Tarmizi, H., de Vreede, G.J., Zigurs, I.: Leadership challenges in communities of practice: supporting facilitators via design and technology. Int. J. e-Collab. **3**(1), 18–39 (2007)
56. Tawse, G.: An exploration of the assessment experiences of new academics as they engage with a community of practice in higher education. Br. J. Educ. **39**(3), 433–442 (2009). 10.1111./j.1467-8535.2008.0084.x
57. Trusson, C., Doherty, N., Hiscop, D.: Knowledge sharing using information technology service management tools: conflicting and incompatible practices. Inf. Syst. J. **24**(4), 347–371 (2014)
58. Wenger, E., Snyder, W.: Communities of practice: The organizational frontier. Harv. Bus. Rev. **78**(1), 139–145 (2000)

fridaya@uef.fi

59. Wenger, E.: Community of practice learning as a social system. Syst. Think. **9**(5), 2–3 (1998)
60. Wenger, E.: Communities of practice and social learning systems: the career of a concept. In: Blackmore, C. (ed.) Social Learning Systems and Community of Practices, pp. 179–198. Springer, London (2010)
61. Wenger, E.: Community of Practice: A Brief Introduction (2015). http://wenger-trayner.com/wp-content/uploads/2015/04/07-Brief-introduction-to-communities-of-practice.pdf
62. Wijingaert, L.A. Van de.: Knowledge sharing, culture and media usage in an academic network: a network analysis of the EBERE/IRSE case. In: 18th Pacific Asia Conference in Information Systems (PASIS2014) Chengdu, June 24–28 June 2014, Chengdu, China
63. Yang, H.: From a website to a web based research support system for collaborative research-eBerea case. Master Thesis, Faculty of Technology Policy and Management, Tu Delft, The Netherlands (2011)

fridaya@uef.fi

The Influence of Online and Traditional Computer Laboratories on Academic Performance of Students

Ahmed O. A. Ismail[1(✉)], Ahmad K. Mahmood[2], Ammar E. Babiker[3], and Abdelzahir Abdelmaboud[1]

[1] Department of Information Systems, King Khalid University, Abha, Saudi Arabia
ibn_omer@hotmail.com, aelnour@kku.edu.sa
[2] Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Seri Iskandar, Malaysia
kamilmhpet@gmail.com
[3] Department of Information System, Al-Baha University, Al Bahah, Saudi Arabia
ammaralyas@gmail.com

Abstract. This paper focuses on online computer laboratories and traditional computer laboratories attempting to improve the academic performance of students' practical skills. It also explores the differences of participants in online computer laboratory compared to traditional stated laboratory. The method used is a laboratory experiment with two groups. The first group uses online computer laboratory on asynchronous communication mode and the second group uses traditional computer laboratory. The results show that there are significant differences between the two groups in an online laboratory compared to traditional laboratory. In addition, the time spent by each group has significant variation in the two different environments. Moreover, the study shows that the online laboratory is better than traditional laboratory in terms of time variation, benefits, accessibility, flexibility, total acquisition, operation, maintenance cost, and achievement performance. In addition, the impact of online laboratory on the students' academic performance is higher than traditional laboratory.

Keywords: Online laboratory · Traditional laboratory · Student academic performance · Online application links · Blackboard learning system

1 Introduction

An online computer laboratory is a challenging issue to the trainees, students, and instructors. It adds new methods, approaches, and techniques to combine the software programs to an online teaching environment [1]. It offers greater flexibility and allows access to more students within a given timeframe while reduces the total acquisition, operating and maintenance cost [2]. One of the advantages of an online computer laboratory is the time management with no location constrains in teaching and learning.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1030–1046, 2019.
https://doi.org/10.1007/978-3-030-22868-2_69

fridaya@uef.fi

Both students and instructors are free from the restricted laboratory hours carried out in a single location [3]. The online laboratory helps in the distance learning setting where learners implement a laboratory-oriented course or exercise at home. The collaboration among the learners will be high in various educational institutions through executing their laboratory exercises as if the laboratory is brought online to the classroom [2]. In addition, students are able to submit their posts to the instructors via course mail or discussion forums. This flexibility is useful for the learners especially for disabled students who cannot physically attend the traditional laboratory [3]. Hence, it reduces students' excuses and helps students perform their assignments and exercises anytime and anywhere.

Online laboratory assignments are conducted under asynchronous communication mode. It is a non-real time interaction of students participation in which information is exchanged with many tools namely; e-mail, mailing lists, discussion forums, course mail messages, and newsgroups [4]. These tools have no time restrictions and location constrains since the participants do not have to be online at the same time allowing more flexibility for the students. The benefits of an online laboratory compared to traditional laboratory are many. For example, the instructor can design assignments, exercises, and quizzes managed by software to share among instructors and students. Furthermore, as there is no physical attendance taken, students can choose the preferred time to do their work and to complete their assignments [4]. In addition to that the development made for an online laboratory allows World Wide Web students to gain practical experience and obtain all knowledge similar to a traditional laboratory through internet technology [5]. Finally, from an economic point of view, an online laboratory can solve the problem of space that could accommodate the presence of many students. It will save the undesirable overhead costs associated with logistics planning and instructor time. As the numbers of distant learners and distant programs increase, the demand for an online laboratory experiences will also increase [2]. All these issues are important and must be considered in both online and traditional laboratories [6]. On the other hand, in traditional laboratory the instructor has to consider different practical styles in designing and managing assignments as well as preparing the assignments on sheets or board. Attendance will be taken and students should complete their laboratory assignments or exercises in a limited time [3].

In many situations, the traditional laboratory session may not help instructor to manage all students effectively. In this work, two groups of participants directed for studying an online laboratory versus traditional laboratory using an online application links developed by [10] at the Department of Information System, King Khalid University, Saudi Arabia. In computer science, there is often a need to develop practical experiences as a key to becoming an effective student and to be able to transfer theoretical concepts into applicable knowledge [6]. In order to address this need of practical experience, this study uses an online application links to complement and to enhance the delivery of online courses on topic of C++ programming as well as to develop learners' practical experiences. The goal of this article is to apply new methods, approaches, and techniques for combining the software programs to an online teaching environment as well as serving a large number of students distributed geographically in different locations. Also, to extend the literature with new approach of applicable knowledge in an online environment [5].

fridaya@uef.fi

The online application called Real Time Narrator (RTN) is a quantitative program, which is designed for students. It explains how to carryout C++ program step by step. The RTN explains basic programming concepts, programming codes, flow charts, loops, and matrix. The same application uses simplified examples on the concept of programming. It also helps in tracking the body of main function. In addition, students are able to compile and run the program code and get the result in an online environment. The RTN links synchronously between a program statement, output screen, memory status, and simplified textual explanation. It aims at making a competent generation in programming language since programming is highly essential in the information technology. The scope of RTN offers simplified explanation for the C++ language such as announcement of variables, rules of priority in the implementation of the operation, and control the output.

The objectives of the RTN can be summarized in six points. First, RTN provides new learning environment for learners to get knowledge. Second, it consolidates the basic programming concepts via C++ language. Third, it uses interactive learning environment that enables learners to give and take in learning situations. Fourth, it raises learners' motivation via attractive learning environment. Fifth, it develops learners' self-directed learning. Finally, it draws attention of teachers and developers to the huge potentials and creative skills that can be gained through using e-learning system [10].

The learning skills provided by the RTN are classified into three: first, students gain experience by using different methods to perform their practical work in an online laboratory compared to traditional laboratory. Second, students develop their skills by participating in discussion forums and course mail messages. Finally, students improve their relationships by exchanging comments in discussion board. The main aim of this study is to investigate the influence of computer laboratories on academic performance in online learning environment compared to traditional learning environment.

Many studies conducted researches on laboratories in different fields of engineering and science. The authors in [7, 8] conducted two laboratories in field of science explaining the effect of anatomy and chemistry laboratories for academic achievement. The authors in [9, 10] conducted researches on virtual simulation and experimental operation platform to examine the difference between students who learn using virtual simulation compared to those who learn using traditional laboratory as well as to enhance their independent learning ability and critical thinking ability respectively. In [11], the author conducted a research on electronic laboratory assessments with tutoring enhancing delivery and increasing laboratory grading tasks of graduate teaching assistant. However, none of the stated studies focuses on the influence of computer laboratories on academic performance in field of software programing. Therefore, this study attempts to find the differences between online computer laboratories versus traditional computer laboratories in the field of computer science using C++ programing language in two different environments. The study also aims at generating synchronously link between program statements, output screen, and memory status.

Since there are two groups of study and each group contains two sections, the questions addressed in this work are: (1) Are there any significant differences in academic performance between the two sections in traditional computer laboratory?

fridaya@uef.fi

(2) Are there any significant differences in academic performance between the two sections in online computer laboratory?

2 Online Environment

This section discusses the importance of online laboratories, the comparison of online computer laboratories to traditional computer laboratories and technical and pedagogical limitations of an online environment.

2.1 The Importance of Online Laboratories

The students' participation in an online laboratory has various benefits for students. These benefits are reported by [3] into flexibility of doing assignments, exercises and homework without time restrictions and location constraints. In addition, it offers opportunities to various types of students due to its special characteristics. Moreover, the increase in class size, higher student-to-instructor ratio and a decrease in student-to-instructor interaction have resulted in decreasing of motivation and participation of students in the traditional laboratory. In [12], the author emphasized that laboratories are important elements in science, engineering and technical education. They allow turning the theoretical concepts into applicable ones.

The author in [2] divided computer laboratories into three types: local, virtual and online. In Local Computer Laboratory (LCL), students operate real devices, manipulate, and measure real objects while being directly co-located with the devices and objects in the same room. Students work as a team with practical help from instructor. LCLs are still the best way to get a first hands-on experience real devices.

The Virtual Computer Laboratory (VCL) contains software simulations of physical devices or real life systems, experiments and pre-recorded measurements, pictures and videos. However, VCL does not manipulate real objects. This type of laboratory is highly expensive and time-consuming if compared to local and online laboratories.

In Online Computer Laboratory (OCL), students and devices are located separately. Students work through a computer that is connected to online applications based on internet. The OCL tries to combine the characteristics of fundamentals of traditional laboratory and flexibility of virtual laboratory. Moreover, OCL develop students' skills in using an online environment.

In [2], the author found that active learning by means of an online laboratory has a positive value for distant education. Students can access laboratories assignments without attending traditional laboratories. This flexibility is important for students as it allows them to practice learning concepts and changes these concepts into applicable terms. The author in [12] stated that high cost of constructing traditional laboratory, outfitting, and maintaining laboratories have often prevented students from getting access to facilities of high importance to their education. Using an online application links has the potentialities of removing the obstacles of cost, time, technical support, and limited access to laboratories.

The author in [5] explained that the European-wide availability of an online laboratories benefit students in less developed countries where few laboratories are

fridaya@uef.fi

available. OCL can support group working over the internet from multiple locations and can remove the geographic proximity restriction that has far reaching consequences for education. In [4] and [13], the authors pointed out the advantages of using online laboratories for students that can work from distant locations 24/7 as well as share the work in online laboratories. Furthermore, OCL can increase the efficiency of an online laboratory operation. The author in [6] proclaimed that OCL reduces the cost per student teaching hours. In [12], the author indicated that OCL makes the laboratory available to a large group of students. Thereby OCL enhances the student's educational experience. Moreover, in [13] the author confirmed that OCL allows teachers in the classroom to illustrate theoretical concepts with real applicable knowledge.

2.2 Online Computer Laboratory Compared to Traditional Computer Laboratory

In [12] compared the OCL to traditional computer laboratories. The comparison revealed that OCL opened the potentiality for flexible practice and access to a large number of students from geographical distributed locations; thus, reducing the cost through sharing an online laboratory. In addition, in [5] indicated that OCL enables online students to perform practical assignments or exercises in real application at various geographical locations and facilitates the interactivity of students with the system. In [14], the author stated that there were a number of ways for accessing practical laboratory that included students access assignments or exercises over a long timeframe or times preferred by them. The OCL can be shared by universities to improve access for disabled students and facilitate greater access to assignments in distance education. In [13], the author stated that the accessibility of OCL creates an opportunity for educational institutions to make laboratory work available to students out of normal hours. Therefore, students can access the online laboratory frequently during the night hours. Hence, the application' links used by this article permits student access to the course assignments over internet as well as being used by instructor to perform demonstrations in front of a class during laboratory period.

The author in [14] declared that in North America and Europe, practical work had been carried during the last years in making the OCL accessible to disabled students in a fully participatory way. They added that OCL offers a possibility of overcoming problems of physical attendance to traditional computer laboratory. Therefore, from pedagogical point of view, the role of OCL is important in introducing students to the world of scientists and engineers in practice. Today the reality for many students is that, the OCL can work collaboratively through the computer which provides excellent context for developing online collaboration skills that will be important in their future careers.

2.3 The Use of Online Environment

The online environment is used to provide new learning method for learners to get knowledge through using basic programing language concepts. Also, how learner uses practical activities through online applications links. These links help learners to use two different methods to do their work in an online computer lab besides traditional lab.

fridaya@uef.fi

Moreover, the links develop learners' skills in participation in online classes and assist the relationships between participations. Finally, the online environment is used to help learners to do their tasks an anywhere at any time.

2.4 Limitations of Online Environment

The online environment limitations are classified into technical and pedagogical limitations.

Technical Limitations. In spite of the advantages and benefits stated for the OCL, some technical limitations are affecting the usability of OCL. These limitations are classified by [3, 6, 14] into the poor connection of internet that makes the participants quickly get frustrated and stop working with OCL. In addition, there is a lack of communications with instructor and remote tutor to solve demanding problem. Moreover, the communication with other online students and trainers is not standardized, resulting in some frustration as reported by students in the sessions. This frustration is due to lack of control or delay in feedback from remote tutors. Furthermore, the channels for communication do not have easily predictable responses due to the variability of traffic load. The actual delay of responses will depend on the network configuration, bandwidth, routing and traffic at the time.

Pedagogical Limitations. The instructor considers three pedagogical limitations when learners use OCL. The author in [12] reported the first limitation as; some online laboratories are difficult for students to access. The second, they are not integrating into a common framework. The third, some students' need more training and background for using online computer laboratory since these laboratories are used only in science and engineering education at universities.

3 Laboratories Experiments

Two methods are applied to collect the data for this study. The first method is online computer laboratory method and the second is traditional computer laboratory method. Both methods are conducted in the faculty of Science and Arts at King Khalid University, Saudi Arabia. These methods include the participation of students, the instructions and rules of the participants group, tools for data collection and the procedures used for laboratory experiment.

3.1 Participants

The participants are the undergraduate students from the Department of Information System. The participants are divided into two groups. Both groups are studying the course of Data Structure and Algorithms 2 (221 HALL-3) with total number of 35 students. Each group is divided into two sections. Section 1 consists of 12 students and Sect. 2 consists of 23 students. The justification for this distribution is to control the laboratory and to ensure that the participants follow the instructions and guidelines provided by the instructor.

fridaya@uef.fi

3.2 Instructions and Rules of Participants

The instructor creates discussion forum and gives specific guidelines for using discussion board. The online group posts their work in discussion forums. On the other hand, the traditional group does their work in computer laboratory. The instructions and rules provided by instructor to participants: online students are trained on an online environment before the actual online laboratory started. In addition, the two groups use the same course materials for their practical work and they perform seven assignments in both online and traditional environments. Moreover, one assignment per week is given for those who attend traditional computer laboratory and those who work in an online environment. The period of the study covers fourteen weeks in second semester of academic year 2013. The first seven weeks assignments are given to traditional group and the last seven weeks of the semester assignments are given to online group.

3.3 Tools for Data Collection

The first tool used for collecting the data for the online group is from an online environment using online application links. On the other hand, the second tool used for collecting data for traditional group is from traditional environment using computer laboratory.

Procedure Used for Online Environment. The procedure used for calculating time for an online environment using seven assignments starts by student login in blackboard learning system. They select the course, open discussion forum, and run the application links. Then, they perform the assignment and post the results through discussion forum or course mail messages. The tool used for determining the time is statistics tracking that turned on for each individual assignment for tracking all the online participants.

Procedure Used for Traditional Environment. The procedure used for calculating time for traditional environment using seven assignments starts by student writing a program code for assignment, compile it, correct the errors code, run the program and record the results. Each student in traditional environment before beginning to write the program code for assignment, will record the starting time and the end based on computer system. Table 1 illustrates the calculated time in online and traditional computer laboratories.

3.4 Technologies Used for Online Environment

Technology used is a web-based application developed by HTML, DHTML, Java script and explanation pages designed by PHP. RTN is integrated with e-learning system such as blackboard. The RTN can be used in discussion forums by adding each session in a separate link. Students use these links for submitting their assignments, questions and feedback using the same discussion forums or course mail messages in the blackboard learning system. Figure 1 shows the structure of an online computer laboratory.

fridaya@uef.fi

Table 1. Time calculated in two different computer laboratories

##	Traditional Laboratory	Online Laboratory
1	26.3	5.4
2	30.0	5.7
3	29.7	7.8
4	31.3	8.9
5	32.7	9.8
6	37.3	10.5
7	39.3	10.9
8	41.7	9.5
9	42.3	10.3
10	46.3	7.4
11	49.3	4.7
12	53.3	6.0
13	19.7	6.3
14	25.0	6.8
15	28.0	6.7
16	28.0	6.3
17	32.7	10.0
18	36.3	9.8
19	42.0	11.0
20	44.7	10.3
21	52.7	10.5
22	44.3	11.7
23	32.3	10.4
24	32.7	10.9
25	44.0	9.5
26	39.7	10.5
27	43.0	10.3

Table 1 demonstrates maximum time in traditional computer laboratory to perform an activity is around 53 min compared to 20 min. In an online computer laboratory, minimum time to perform an activity is around 12 min compared to 5 min.

The application connected to the server over TCP/IP protocols. The remote clients are able to interact with an online application by communicating with the server via internet connection. The web server receives responses from client. Then web server directed the received responses to the host server. The Hyper Text Transfer Protocol (HTTP) carries the replied responses back to the clients via internet connection. JDBC refers to Java Database Connectivity in which database is manipulated and managed.

fridaya@uef.fi

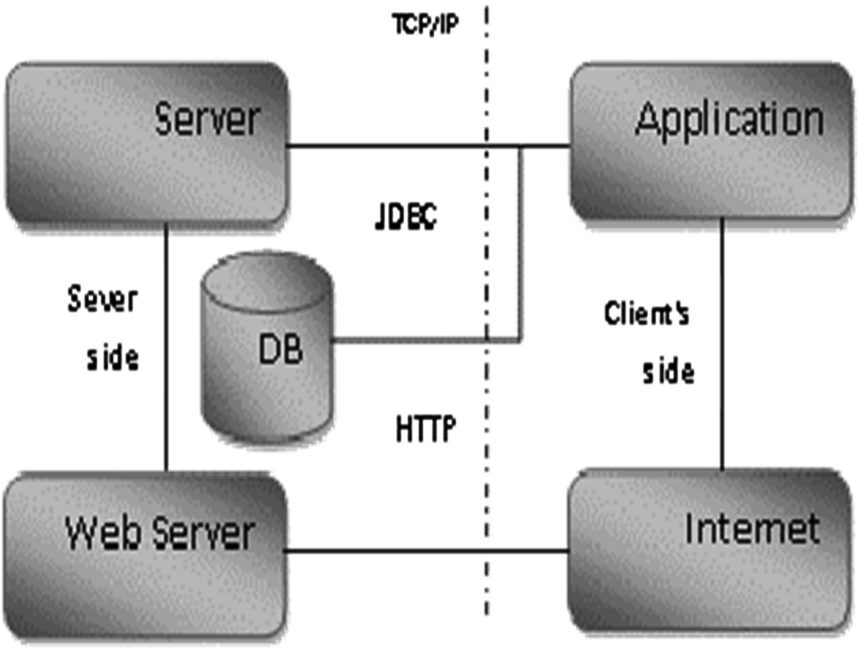


Fig. 1. Structure of an online computer laboratory.

3.5 Screenshot of Online Application Links (OAL)

The screenshot of OAL is used to demonstrate the functions of C++ programming language, which include determining rectangle area, manipulating conditional statements, calculating the sum of one-dimensional array elements and explaining the rules of priority of arithmetic operations.

Figure 2 demonstrates RTN that contains four parts of OAL. The first part C++ program source codes where learners navigate between the source code statements. The second part explains the statements inside the main function. The third part accepts the inputs from learners and displays the output. The final part shows the memory status of the input variables width and location address.

Figure 3 is similar to Fig. 2 except that it contains part five that shows the direction of flow chart path when conditional statement is true or false. Arabic and English explanations are used in Figs. 2, 3, 4 and 5.

Figure 4 demonstrates how the one-dimensional array stores its first element in location number zero and last element stores in location n−1 where n is array size and its value is greater than zero. In addition, the memory status shows the locations addressed for each element of an array.

Figure 5 demonstrates the declaration of k variable of integer type. In memory status, the undefined value of k reserved four bytes stored in specific memory address. The value of (k) varies according to priority of arithmetic operations.

fridaya@uef.fi

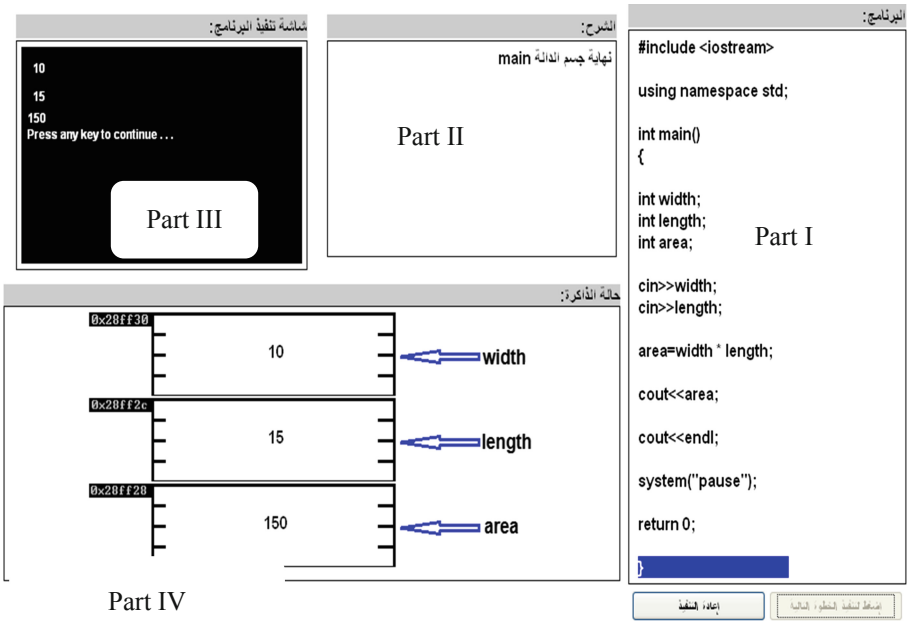


Fig. 2. RTN interface for determining the rectangle area.

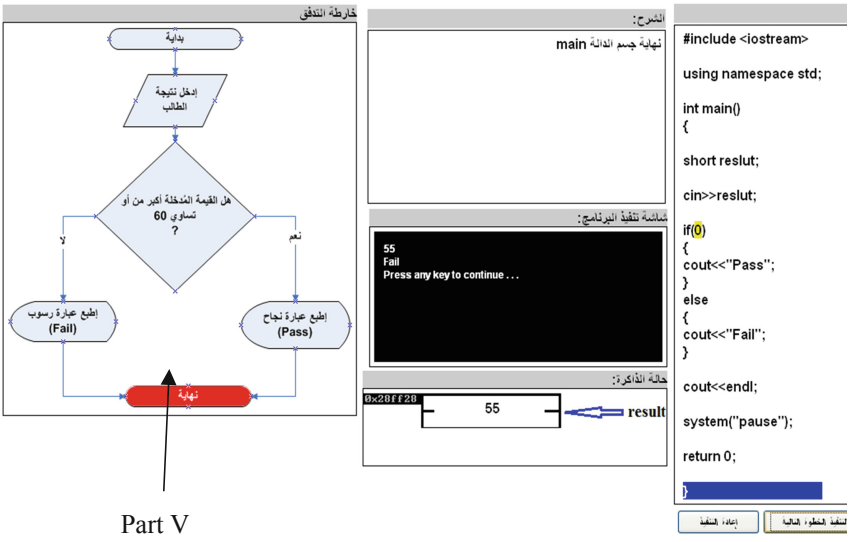


Fig. 3. RTN interface for manipulating the conditional statements.

fridaya@uef.fi

Fig. 4. RTN interface for calculating the sum of one-dimensional array elements.

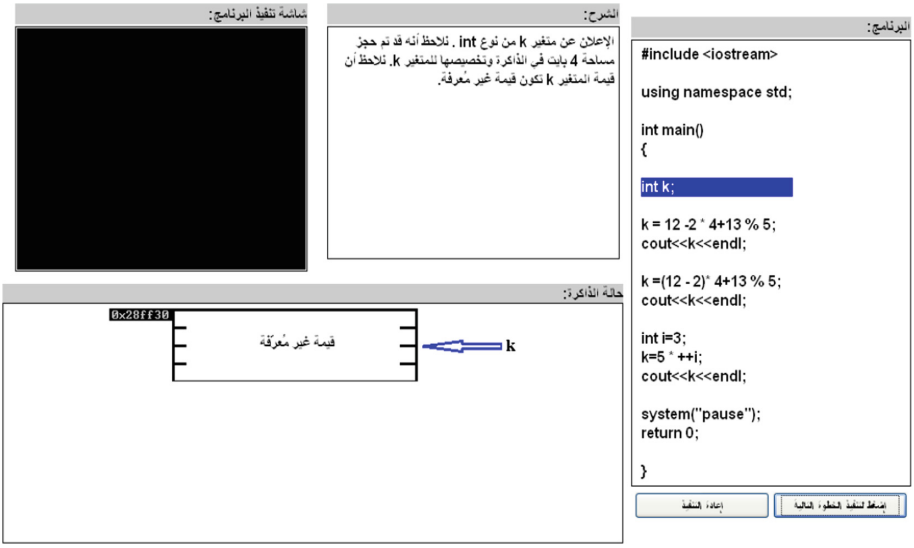


Fig. 5. RTN interface for explaining the rules of priority of arithmetic operation.

4 Data Analysis

The aim of this section is to analyze the stated questions addressed in Sect. 1. The method used in this analysis is SPSS version (22.0) to analyze the results of descriptive statistics. Three tests are used to examine the three hypotheses formulated for this article. These tests are independent t-test, Post Hoc tests for ANOVA and Paired-Test. All tests used a 0.01 and 0.05 levels of significance.

fridaya@uef.fi

The hypotheses formulated from the stated questions:

H_1: There will be statistically significant difference of academic performance in traditional environment.

H_2: There will be statistically significant difference of academic performance in online environment.

H_3: There will be statistically significant difference between the two groups in online environment compared to traditional environment.

5 Results

An independent t-test is used to examine the three hypotheses to investigate the differences of academic performance in traditional group, online group and the individual group in online and traditional laboratories. The result of the independent t-tests which is used in H_1 shows that, there is no statistically significant difference of academic performance in terms of laboratory assignment between the two sections of traditional group in traditional environment. The results obtained indicate that T.Section 1 (M = 39.25, SD = 9.845, DF = 26, T = 0.388, P = 0.701 > 0.05) and T.Section 2 (M = 40.69, SD = 9.611, DF = 26, T = 0.386, P = 0.703 > 0.05). Where M = Mean; SD: Standard deviation; T.Section = Traditional Section; DF = Degree of freedom; T-V = T value. Based on overall results of M and SD of traditional group presented in Table 2, the analysis proves that there is no significant difference between the two sections of traditional group in traditional environment. Thus, the null hypothesis of H_1 is rejected and the alternative hypothesis is accepted. The result of the independent t-tests used to test H_2 shows that there is significant difference of academic performance in terms of laboratory assignments between the two sections of online group in online environment. The results obtained indicate that O.Section 1 (M = 7.08, SD = 1.564, DF = 26, T = 4.60, P = 0.000 < 0.01) and O.Section 2 (M = 4.69, SD = 1.195, DF = 26, T = 4.424, P = 0.000 < 0.01). Where O.Section = Online Section

Table 2. The differences between the two laboratories –using t-test.

Time	T1: Time of traditional: T2: Time for online						
Group	Sections	No.	M	SD	DF	T-V	Sig. .01–.05
T1 Traditional	T.Section 1	12	39.25	9.845	26	0.388	.701 > .05
	T.Section 2	16	40.69	9.611	26	0.386	.703 > .05
Overall M & SD		28	39.97	9.728	26	0.387	.702 > .05
T2 Online	O.Section 1	12	7.08	1.564	26	4.600	.000 < .01
	O.Section 2	16	4.69	1.195	26	4.424	.000 < .01
Overall M & SD		28	5.885	1.380	26	4.512	.000 < .01

Based on the overall results of online for M & SD, the analysis concludes that there is significant difference between the two groups in an online environment compared to traditional environment. Thus, the null hypothesis of H_2 is accepted and the alternative

fridaya@uef.fi

hypothesis is rejected. The probability obtained in an online environment for the online group $p = 0.000 < 0.01$ level of significant. This means that there is significant difference between the two sections of online group in an online environment. Thus, the null hypothesis of H_3 is accepted and the alternative hypothesis is rejected. To confirm the above results another Post Hoc Test, LSD and Scheffe are used. The results in traditional environment indicate that T.Section 1 (M = 23.0, SD = 7.640) and T.Section 2 (M = 24.27, SD = 5.854) with DF = 53, F = 45.752 > 3.0 see Table 3 refers to ANOVA Test. However, the results found in an online environment indicate that O. Section 1 (M = 10.08, SD = 3.232) and O.Section 2 (M = 6.07, SD = 2.086) with DF = 53, F = 45.752 > 3.0). Based on these results, the analysis proves that there is significant difference between the two groups in an online environment compared to traditional environment. Thus, the same results confirm that the null hypothesis of H_2 is accepted.

Table 3. The differences between the two laboratories– using ANOVA test

*The mean difference is significant at 0.05 levels.

Groups	Sections	O.Section 1	O.Section 2	F-V	No.	M	SD	DF
Traditional	T.Section 1	12.917*	16.933*	45.751	12	23.0	7.640	53
	T.Section 2	14.383*	18.400*		15	24.27	5.854	53
Overall M & SD					27	23.64	6.747	53
Online	O.Section 1	–	4.017*		12	10.08	3.232	53
	O.Section 2	–	–		15	6.07	2.086	53
Overall M & SD					27	8.08	2.659	53

For more confirmation, two other tests are used to examine T1 and T2 through paired tests of two variables in the same group to see whether there is significant difference in an online environment compared to tradition environment. The results of Table 4 for online group found that O.Section 1 (M = 22.42, SD = 6.882) and O.Section 2 (M = 7.33, SD = 1.614) with (DF = 11, T = 7.758, P = 0.000 < 0.01, Correlation = 0.207). This result shows significant difference between two sections in online group in both environments and weak correlation between the two sated variables.

The results in Table 5 show that in traditional group T.Section 1 (M = 58.78, SD = 16.936) and T.Section 2 (M = 6.09, SD = 1.564) with (DF = 22, T = 16.411, Correlation = 0.984). These results show significant difference between two sections in traditional group in both environments and strong correlation between the two sated variables.

The Paired-Test for each group confirms that there is significant difference between two groups.

fridaya@uef.fi

Table 4. Comparison of online group

		T1: Time for traditional and T2: Time for Online						
	Pair 1	No.	M	SD	DF	T-V	Sign. .01	Correlation
Online	T1	12	22.42	6.882	11	7.758	.000 < .01	0.207
	T2	12	7.33	1.614	11			
	Pair 1 T1–T2	12	15.083	6.735				

Table 5. Comparison of traditional group

		T1: Time for traditional: T2: Time for Online						
	Pair 1	No.	M	SD	DF	T-V	Sign. 0.05	Correlation
Traditional	T1	23	58.78	16.936	22	16.411	0.000 < 0.01	0.984
	T2	23	6.09	1.564	22			
	Pair 1 T1–T2	23	52.696	15.399				

6 Evaluating the Two Laboratories

Table 6 demonstrates the ways of evaluating both traditional and online computer laboratories. The suitable way of evaluation for participants is based on laboratory situation [2]. The laboratory situation includes entities in both laboratories, the process to be controlled in the two labs, technology used in both laboratories and results to be obtained. The entities in the two laboratories are the same except that the computer devices in tradition laboratory are at one location, while application links in an online laboratory are at different locations. The process that is to be controlled in traditional laboratory is situated in only one location. However, the process to be controlled in an online laboratory is via internet connection. Both entities contribute in the evaluation that is intended to provide participants with controlled skills in the learning environment. Instructor manages the activities in traditional laboratory, while the software manages the activities in an online laboratory. The technologies used in traditional laboratory are Local Area Network and intranet. However, the technologies used in an online laboratory are Wide Area Network, internet and online learning tools such as email, course mail, discussion forums and links. This evaluation is also intended to provide participants with more experience in how to manage and use learning tools. The obtained results in traditional laboratory can be seen through computer screens, while the obtained results in an online laboratory can be seen via display output screens and memory status. This evaluation is intended to provide participants with the skill of how to visualize the memory status when the output is stored in computer memory. In addition, when the variable is declared, the participants will see how it is stored in memory status.

fridaya@uef.fi

Table 6. A summary for evaluating the two laboratories

	Computer laboratories evaluation	
Situation	Traditional	Online
Entities	Instructor, learners and computer devices in one location	Instructor, learners and application links in different locations
Process	Computer laboratory, Physical attendance	Via internet connection & no Physical attendance
Activities	Assignments, quizzes managed by instructor	Assignments, quizzes managed by software
Technologies	LAN, Intranet	WAN, internet and online learning tools such as email, Links and discussion forums
Results	The learners can see the results through computer screens	Learners can see the results of their action in front of them through output display screen and memory status

Figure 6 demonstrates the comparison between the time for implementing assignments and activities in online laboratories and traditional Laboratories. The results show that the means of online laboratory time varies from five to twelve minutes. On the other hand, the means of traditional laboratory time varies from 20 to 53 min. This variation of time in implementing the assignments and activities indicates that there is significant difference between the two laboratories.

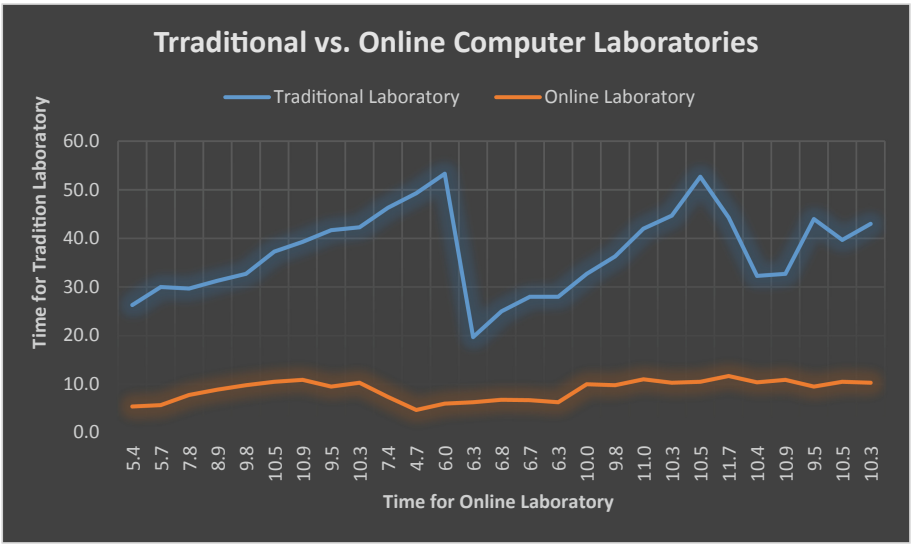


Fig. 6. Illustrates time for both tradition and online computer laboratories

fridaya@uef.fi

7 Discussion

This section discusses the results obtained from this study in relation to the findings of literature review. This work shows there is no significant difference between the two sections of traditional group in traditional environment. The obtained result is related to the finding found by [9] which confirmed that there is no statistically significant difference between the students who learn using traditional laboratory instructions. The implication of this result on practice is that traditional laboratories are considered the best for learners to get first-hand experience if the laboratories are constructed, outfitted, and maintained. Another result reveals that there is significant difference of academic performance between the two sections of online group in online environment. This result is supported by [8] which stated that virtual and online laboratories increase the level of students' achievement. The implication of online laboratories on learners' performance is that, the offered online assignments and activities are available 24/7 day long for learners. This availability of online materials in terms of assignments and short quizzes will improve students' practical skills, which will be reflected in their academic performance further. In addition, this study reveals that there is a variation in time spent by the two groups in both online and traditional laboratories. This result is advocated by [7] which explained that the time spent studying for each laboratory may explain significant differences between the traditional and online group. The implication of time variation on learners allows them to save time and spend it on other theoretical concepts.

The strengths of the discussed findings are represented in theoretical and practical perspectives of online and traditional computer concepts, approaches, techniques and practice of online applications for C++ programing language. In addition, learners overcome the restriction of time and the constrains of place to perform their work. On the other hand, the achievement of the objectives of the study are demonstrated in learners gain experience by using two different methods to perform their practical work in two different environments. Also, the learners develop their skills through participating in discussion forums and post their question via course messages. Furthermore, the learners improve their relationship by working as a team as well as exchanging comments through discussion board. However, the weaknesses of the findings are represented by technical and pedagogical limitations that discussed in Sect. 2.4, as well as the small size of sample to confirm the reality of generalization.

8 Conclusion

This study investigates the difference in the academic performance of participants in online computer laboratory compared to traditional computer laboratory. The study also explores the practical work over an online environment using online learning tools such as course mail messages and discussion forums. Based on the results obtained from laboratory experiment, this work discovers that the online laboratory is better than the traditional laboratory in terms of time restriction and location constrains. In addition, students who practice using the online laboratory gain many benefits. The impact of online laboratory on the students' academic performance is higher than traditional

fridaya@uef.fi

laboratory. As a summary, the results of this study show that there is significant difference between the online laboratory and traditional laboratory in terms of time, location and performance. Finally, this study classifies the limitations of an online environment into technical and pedagogical limitations.

References

1. Ahmeddi, A.M.: "RTN: Real Time Narrator" Lecturer in King Khalid University, Community College, Bisha, Saudi Arabia 2070-1721 (2013)
2. Tuttas, J., Wagner, B.: Distributed online laboratories. In: International Conference on Engineering Education, pp. 7–11, Oslo, Norway (2001)
3. Sagun, A., Demirkan, H., Goktepe, M.: A framework for the design studio in web-based education. J. Art Des. Educ. **20**, 332–342 (2001)
4. Hew, K.F., Cheung, W.S.: Students' use of asynchronous voice discussion in a blended-learning environment: a study of two undergraduate classes. Electron. J. E-Learn. **10**, 360–367 (2012)
5. Böhne, A., Faltin, N., Wagner, B.: Synchronous tele-tutorial support in a remote laboratory for process control. Innovations, pp. 317–329 (2004)
6. Salzmann, C., Gillet, D., Latchman, H.A., Crisalle, O.D.: On-line engineering laboratories: real-time control over the internet. Age **4**, 1 (1999)
7. Mathiowetz, V., Yu, C.H., Quake-Rapp, C.: Comparison of a gross anatomy laboratory to online anatomy software for teaching anatomy. Anat. Sci. Educ. **9**, 52–59 (2016)
8. Tüysüz, C.: The effect of the virtual laboratory on students' achievement and attitude in chemistry. Int. Online J. Educ. Sci. **2**, 37–53 (2010)
9. Hayward, M.: Comparison between computer assisted instruction as a laboratory strategy, and traditional lab instruction as applied to teaching physical science to suburban high school students (2016)
10. Niu, R., Zhang, Z.: IEBOOK-based experimental operation platform of computer basic course. Int. J. Emerg. Technol. Learn. (iJET) **12**, 51–61 (2017)
11. DeMara, R.F., Salehi, S., Khoshavi, N., Hartshorne, R., Chen, B.: Strengthening STEM laboratory assessment using student-narrative portfolios interwoven with online evaluation. In: Proceedings of American Association for Engineering Education Southeastern Conference, pp. 13–15 (2016)
12. Nejdl, W., Wolpers, M.: European e-Learning: important research issues and application scenarios. L3S Research Center and University of Hannover, pp. 35–48 (2004)
13. Gustavsson, I., Zackrisson, J., Håkansson, L., Claesson, I., Lagö, T.L.: The visir project–an open source software initiative for distributed online laboratories. In: REV 2007 (2007)
14. Cooper, M.: Remote laboratories in teaching and learning–issues impinging on widespread adoption in science and engineering education. Int. J. Online Eng. (iJOE) **1** (2005)

fridaya@uef.fi

Smart, Social, Flexible and Fun: Escaping the Flatlands of Virtual Learning Environments

Mike Brayshaw(✉), Neil A. Gordon, and Simon Grey

Department of Computer Science and Technology,
University of Hull, Hull HU6 7RX, UK
{m.brayshaw, n.a.gordon, s.j.grey}@hull.ac.uk

Abstract. This paper describes the development of intelligent, social, flexible and game-based pedagogic approaches and their applications in Virtual Learning Environment based Education. Applications of computer science technologies and techniques can enable, facilitate and change educational approaches, allowing scalable approaches that can address both individual student needs whilst managing large – sometimes-massive - cohort sizes. The benefits of these information systems include supporting the wide range of contexts met in education, in terms of individual needs and specific subject and curriculum requirements. Technologies and approaches that are considered range from the representation of knowledge and the use of intelligent systems, the use of social computing, through to the enabling opportunities of ubicomp and the practical application of game mechanics (gamification). This paper concludes with practical illustrations in the context of undergraduate computer science didactics.

Keywords: Computer science education · Computer aided instruction · Pervasive computing · Flexible pedagogy · Technology enhanced learning · Gamification based learning

1 Introduction

Computer Science provides the underpinning and inspiration for a number of educational technologies and techniques. This paper considers a number of computer science based approaches to teaching and learning that are emerging – reflecting the capacity and potential of new technologies and systems. Alongside the systems themselves, this paper also considers the potential benefits of these systems, in terms of pedagogic change, and how this may be applied in practice. In this context, learners, teachers and technologists all represent different forms of user. This paper considers the rise of smart technologies and their impact on pedagogic practice, the different approaches to learning and assessment that they afford, and then goes on to consider the impact of social media within this learning mix. It goes on to explore the personalisation and gamification that can be enabled through technology, and in Sect. 6 applies these to the teaching of programming and computer science concepts.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1047–1060, 2019.
https://doi.org/10.1007/978-3-030-22868-2_70

fridaya@uef.fi

The main context considered in this paper is higher (tertiary) education provided through universities, but the technologies, principles and issues are applicable to the broad range of education provision, from primary through to college, and on to lifelong (work placed and informal) learning scenarios.

The paper provides a discussion of the emerging approaches and systems that provide for technology-enhanced learning, where computer aided instruction is utilized to enable adaptive and flexible solutions for education. The pedagogic approaches are designed to encourage student engagement, whilst enabling the teacher to monitor and manage the student cohort. The paper does focuses in detail on virtual learning environments (VLEs) and Learning Management Systems (LMS). A VLE is a computer based information systems that support learning, teaching and assessment. A LMS is a computer-based system that typically supports learning in a structured way, providing tools to plan and assess the students. Whilst these terms can be used synonymously, a VLE is typically considered as an educational tool, whilst an LMS may support training. In many cases, VLEs consist - or in practice are used - as little more than file management systems; one naïve view is that if they include raw video of recorded lectures this will automatically constitute a positive educational experience. We know that badly authored computer delivered education material does not solve the problem from the legacy of TICCIT and PLATO [1]. In an age where tutorials spring a plenty from the internet and *how to* videos are offered from all manners of authors and vloggers, it is important and timely to think of the crucial importance of design and content from the point of view of pedagogic delivery and effectiveness for learners. Throwing multimedia at the problem is not the solution. What we propose here is a linked story to provide an effective solution.

An ongoing challenge for computer science education remains the effective teaching of programming – and this provides a fitting context through which to illustrate some of the technologies and techniques described below. In the context of England, this is particularly relevant as the nation embarks on a new computer science curriculum, where computing is developed from primary through to secondary schools, with the challenge of how to roll out programming and related skills and understanding to a majority of the population [2]. This paper provides suggestions about how to apply some of these approaches and ideas to the teaching of programming.

2 Smart Technologies and Pedagogy

2.1 Knowledge Aware Applications

Semantic Technologies are well-established [3] and can support teaching through the identification of relevant learning materials. A more recent concept – since around 1999 – is that of the Internet of Things (e.g. [4]). Originally proposed as a linkage of devices on their own IP addresses, more recently the Things have been increased in range to include Intelligent Software Agencies and Web Services. How we use and interact within these evolving environments is the context for future learning. There is a long history of applying Artificial Intelligence to Computers and Education [4] and [5]. Here we argue not only against the formal behemoth of VLEs but also Intelligent Tutoring

fridaya@uef.fi

Systems. In the world of ubicomp, we expect smart from the devices around us. The domestication of the technology, and how we use it, has led to new models of interaction [6, 7] and [8]. The influence of Semantic Technologies is going to see increasing growth of the internet of knowledge aware entities. The world is full of Smart Things. As teachers and learners, we can look to see how interaction with these smart devices is going to change our learning experiences. Knowledge Engineering considerations are thus hidden from users. They typically do not want to know the engineering algorithm behind their GPS or Virtual Assistant. However, from a pedagogical perspective we need to develop knowledge aware applications that are able to deliver this type of functionality to the user.

2.2 Computer-Supported Tools for Adaptive Teaching, Learning and Assessment

When considering traditional classroom – especially university education – one of the key challenges is how to ensure that the material is suitable for individual students. Considering the university context, the efficiency of the lecture format for demonstrating what material is expected is in tension with the actual effectiveness of the lecture format in helping students to understand the material. Such issues are not unique to university, at school level the onus is on teachers to provide individualized learning plans and support. A further issue for teaching contexts is how to manage assessment – the traditional end of material exam can assess learners' knowledge of specific areas, but is limited in terms of identifying learners' actual knowledge of a subject. Computer based systems can provide solutions to some of these issues. The first solution considered is how to provide the learner with appropriate material.

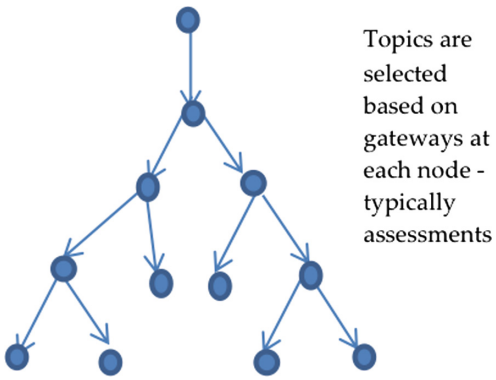


Fig. 1. Directed graph representing hierarchy of topics, concepts and skills – for modelling within the learning system to enable adaptive testing and personalized content.

Computer Science offers a number of technologies and frameworks that can address the issues above, in particular utilizing information markup and intelligent systems to enable personalized learning. As outlined in [9], an approach can be based around

fridaya@uef.fi

computer based diagnostic assessment to identify the learner's current knowledge base, followed up with the direction to the relevant material (level and content). The computer based diagnostic assessment can be made more efficient and effective through an adaptive test – something that traditional education approaches cannot provide. Here, the topics and skills are structured into a tree (see Fig. 1), with the root node question being on an initial filtering topic [10]. Further nodes branch depending on the learner's answers, providing an adaptive approach to determine a profile of the learner to be developed. The outcome of this can be an individual learning plan of resources. Further control is provided by identifying suitable gates (tests) that limit access to material until the learner can demonstrate suitable knowledge and skills. Computer systems can automate and control this approach [11], thus providing an adaptive teaching and learning experience. Alongside the diagnostic testing already mentioned, these adaptive systems can personalize the learning experience. In considering educational systems, the role of assessment becomes key.

Historically, computer based assessment has tended to be based around variants of multiple-choice questions [12]. Disciplines with established formal/technical languages – for example, mathematics, computer science and related disciplines – have the scope for symbolic evaluation, as illustrated with assessment engines built on evaluating symbolic/mathematical statements and comparing for equivalence – either algebraically or via their outputs. As natural language parsing is improving, and as large learning environments – that we consider later – are developed, another emerging trend will be on systems that can assess narrative text. Some of the current Massive Open Online Courses (MOOC) platforms already offer elements of this. Combining computer assessment with peer assessment, and moderating that with samples of teacher based assessment has the potential to alter how students are assessed.

MOOCS are a good example of such large-scale educational systems. Whilst MOOCS can support the automation of much of the teaching process, the major benefit to organizations utilizing MOOCS – alongside potential marketing and profile raising - is the potential to collate extremely large datasets on learner behavior. Linking back to the earlier discussions on personalized and adaptive systems, the difficulty in creating and managing learning content can be in identifying and maintaining suitable hierarchies of content. MOOCS – and other variants of massive learning systems - may enable machine learning to be applied here, removing the need for direct management by teachers. Such approaches are made more viable as educational web content includes more effective information on their content and use – an application of semantic web technologies [13].

2.3 Machine Learning for Education and Learning Analytics

As we live online with leave an increasing data trail. Machine Learning provides a way of looking at the history and search for patterns and hidden knowledge. From an educational perspective, what is typically of issue is predication and classification factors of such things as performance and drop out. Educational data mining is a more specialized version of the established data mining in computer science. The nature of educational environments – typically with a variety of data sources and different types of tracking, means they can be grouped under big data. Learning analytics is one aspect

fridaya@uef.fi

of this – where data about student activity (attendance, performance, access to resources, interactions with teachers, tutors and peers) can all be combined to try to profile and support students. Unlike the personalized approaches considered earlier, the onus here is on identifying students who may need intervention of one form or another. The intervention may be for those who appear to be at risk of dropping out or doing badly or it may be to look at how to challenge and develop those who are doing well. Typical educational data mining insights include

- discovering unknown associations;
- finding unknown/hidden rules;
- finding new trends and groups;
- finding new analogical cases to guide insight;
- finding new ways of solving a problem; and
- finding new ways of classifying data and predication behavior via neural networks.

If such techniques can identify students that might be in a critical zone then appropriate support, counseling, and remedial material can be made available. How can this be used to implement flexible learning? If we better understand students and are able to label and classify them and more effectively predict their outcomes, then we can improve on the flexibility of the provision provided. By knowing more about their personal performance we can customize provision in both delivery and content. An important distinction to make is in what we are data mining for. In the traditional picture, we are looking at features like marks and attendance or other longstanding benchmarks. What we are arguing for here is that other features of performance are of equal interest. Can we use data mining to find social engagement (other than just attendance)? What can we gain from social media patterns? What can we gain from game engagement patterns and what are the repercussions in terms of educational game engagement? How are they interacting with the smart and what can we mine from that?

3 Mixing Social into the Blend

3.1 Ubicomp and Social Computing

The rise of social computing means that there are now generations of users that expect to communication and socialize via computer-mediated communication. This is now a platform that we can look to exploit, delivering interactive educational media. People now instinctively look to the like of YouTube for a video of how to do something new. The rigid likes of a VLE or LMS is no longer the chosen channel of media choice. GitHub embraces both the social and the technical modes of interaction. When people now look to YouTube for a “how to” video they do not want the rigid format of a VLE or a LMS. Indeed if they are forced into such systems, they may typically quit as soon as they get what learning out of it that they originally sought. Github is a good example that embraces both this loose structure of help with a strong social/game (in terms of feedback) element.

fridaya@uef.fi

4 Flexible Technologies and Pedagogy

4.1 Pervasive and Mobile Learning Systems - Overview

The context of computers and education has changed markedly since the early days of computer based tutoring systems (e.g. [1]). Today computing technology is Pervasive and ubiquitous [7], that it to say it is located in the Smart Telephones, Smart Watches and Smart TVs, and all around our houses; this has changed how we talk and communicate [14]. Today many people carry or wear a computer and access much of their media through these devices. Laptops and Pads add to this rich picture. Thus from an educational context the era of mobile learning is within our technical grasp. How we precisely utilize this from the point of pedagogy is however another issue. Just having the delivery option is not necessarily the whole solution. Distance learning has known problems with isolation, loneliness, and self-motivation (e.g. [15]). The other problem with pervasive is a potential loss of privacy; raising questions such as who can see your mistakes and does this put potential users off?

Pervasive technologies brings with it additional possibilities brought about by games consoles and head mounted displays, which have found their way into the home via entertainment and gaming. Augmented reality [16] is one of these. Through a camera or a helmet, the user sees the world but with an overlying story placed upon it. It has the advantage of being able to use the existing world to contextualize the new information. From the point of view of the learner there are exciting possibilities to locate educational materials. Immersive environments like these allow employment of a real world scenario so that the player can learn as they go along either refining intellectual skills or developing physical skills. Indeed they can do both at the same time.

Another form of Mobile Learning is to use Cloud based systems. This is already done in some commercial Content Management Systems/Learning Environments (e.g. [17]). The Cloud ensures that the material is available anywhere and on anything that has a connection, thereby increasing the pervasiveness of the interaction.

4.2 The Transformational Impact of Social Media in Teaching and Learning

Through social media, we are able to construct a world around us, a lens by which we can shape our own unique view of the world. It is a world where one-to-one and one-to-many is possible. The pervasiveness of communication technology and every day of synchronous and asynchronous communication is one of the great changing events of our time. Email, Twitter, YouTube and Facebook are ubiquitous communication tools that form a part of everyday life. The empower users both as authors and audience for this new media traffic. They thus have the potential to transform how we deliver material to users. They also have the potential to change completely the two-way interaction in that how students interact with the teachers is changed. Students may be asked to make and share their work using a wide range of media forms. Traditional problems like loneliness and isolation that distance educators often report may be change by this new social media experience.

fridaya@uef.fi

That said there are down sides as well. Bullying and Trolls are a frequent story in the news. We must be aware that the use of social media is not necessarily all good news. How to use social media in a positive way is a big research question for designers of educational material.

4.3 Remote and Virtual Classrooms and Laboratories?

In this new world, are classrooms and laboratories dead? In [15] the whole of the Summer School was scaffolded around the notions of a tradition University Campus and the affordances that that can bring. Library materials for example where it the Library. Virtual Laboratories provide the potential for learning and collaborating just as before. However, they can now be seen here in the context of pervasive computing, augmented reality and social media. Rather than being old notions, they are now rich lands in which traditional models, but now with the constraint to physically collocated and enhance by anytime, anywhere possibilities and the social world potentially of the users' choice.

4.4 Flexible Learning and Personalization

Flexible learning [18] is concerned with the three dimensions of learner choice [19], namely the where, when and how to learn. This pedagogic viewpoint reflects the wide variety of learners across Higher Education, where assumptions that students are based in classroom or campus scenarios, with their study as their major focus are questioned. Whilst distance learning and personalized material suitable for individual learning style are not new concepts – they are difficult and impractical in large-scale teaching. What Computer Science can facilitate is the upscaling of such provision – with mobile devices allowing for choice in place, computer based teaching allowing for choice in when, and at what pace, to learn, and computer mediated learning providing flexibility in how to learn.

As with any information system solution, learning technologies do not provide an ideal answer – in mapping them on to the teaching and learning process, the typical issues of ensuring that the systems meet the range of users requirements and needs arises – something considered later in this paper.

With the adaptive systems mentioned in the previous section, this allows for the modelling of learning as a directed graph, with multiple paths depending on the choices, preference and performance of the learner. The role of the teacher with this approach moves from a focus on providing content, to a broader role where the main aim is to support the learner in managing their learning.

There is no need to include page numbers. If your paper title is too long to serve as a running head, it will be shortened. Your suggestion as to how to shorten it would be most welcome.

4.5 Authoring and Management of Learning Content

The World Wide Web and related technologies have been disruptive in allowing access to new resources and raised new issues around plagiarism, assessed work by contract,

fridaya@uef.fi

and social and peer support can veer from collaboration to collusion. The ready access to material raises the problem of intellectual property and fair use – an issue for learners and teachers alike. Reusable learning objects - followed by the notion of customizable ones – give the potential for sharing teaching materials. Approaches such as the Learning Object Metadata (LOM) [19] and Learning Materials Application Profile [20] provide a specification for the description and vocabulary for learning objects. However, but technical standards have yet to catch the attention of the majority of practitioners.

Taking inspiration from the software engineering component approach to software, a component based approach to educational material seems opportune – where interfaces are defined semantically, allowing the intelligent systems described earlier to assemble learning resources and landscapes, for the learners to traverse with their teacher guides.

The management and ownership of learning content becomes more complex with the move towards cloud based virtual learning environments, and to sharing resources. Copyright and more general Intellectual Property Right issues (IPR) issues that may be hidden within institution based lessons, are exposed where material is shared through internet sites, whether as part of an online course such as a MOOC, or part of blended learning supporting more traditional campus-based education.

5 Having Fun! Gaming Technologies and Pedagogy

5.1 Game-Based Learning, Gamification and Edutainment

For a long time play has been recognized as an important part of learning. In a study of the role of play in human culture Huizinga identifies that play has an important role in development, noting that the word ‘school’ has a foundation in play “Meaning originally ‘leisure’ it has now acquired precisely the opposite sense of systematic work and training” [21]. Celebrated game designer Chris Crawford claims that “the fundamental motivation for all game playing is to learn” [22] while Koster makes the bold claim that “Fun is just another word for learning” and those games are ultimately teachers [23]. As learners grow older it seems that the role of play is discarded in favor of more serious educational pursuits, and that this transition is accompanied by reduced engagement with learning, whilst games continue to be enjoyed by a wide variety of people. Game-based learning, gamification and serious games all attempt to use aspects of games to increase engagement with teaching subject matter.

In the case of game-based learning, subject matter is embedded within a game. A game is created with the aim of teaching concepts through playing the game. Unfortunately, the commercial genre of games label “edutainment” seems to be more associated with poorly designed, poorly executed games [24]. It is proposed that this is largely due to a combination of poor game design, poor understanding of learning & teaching, and lack of funding in comparison to mainstream games. People of a certain age may fondly remember the highly successful edutainment title “Where in the world is Carmen Sandiego”, but that title was released in 1983 when gaming was still in its infancy and budgets were small.

fridaya@uef.fi

By contrast, "serious games" are most often used by organizations with access to large amounts of funding, such as the military [18]. Serious games tend to use technology originally developed for use in games to provide serious training opportunities in immersive simulated environments, often with a trainer observing and adapting the environment to create specific scenarios and learning experiences. The field of serious games represents a clear-cut example of technology-enhanced learning, and in particular, learning enhanced specifically by technology developed primarily for use in games. Consumers of serious games often have access to significantly larger budgets and gamification has risen in popularity over the past five years. The Gartner group first plotted gamification climbing the peak of inflated expectations on its hype cycle for emerging technologies in 2011 [25]. By 2014, gamification has surpassed the peak and is currently slipping into the trough of disillusionment [26]. In both cases, gamification was expected to reach maturity in five to ten years. Zichermann and Cunningham define gamification as "The process of game-thinking and game mechanics to engage users and solve problems" [27] whilst Kapp combines elements of a variety of definitions to arrive at the definition "Gamification is using game-based mechanics, aesthetics and game thinking to engage people, motivate action, promote learning, and solve problems." [28]. Werbach and Hunter offer the more general, simplified definition of "The use of game elements and game design techniques in non-game contexts" [29] and often forms a layer of extrinsic motivators striving to complement the learning experience.

Gamification often borrows established concepts from games including feedback mechanisms (such as points and badges), element of competition (leaderboards), decreased risk (multiple lives), adaptive challenges (difficulty levels) and meaningful choice (customization) has game elements layered on top of the subject matter.

Specifically in the field of learning and teaching Dicheva et al present a review of papers reporting the application of game elements to education. 34 papers were analyzed, all of which were published between 2010 and 2014 inclusively [30]. By categorizing the content of these papers Dicheva also confirmed that "the most popular game mechanics are points, badges and leaderboards", which has become such a synonymous phrase for gamification that this trio of game elements is often shortened to the acronym PBL (Points, Badges and Leaderboards). Whitton presents a wider account of the role of digital games and learning by considering games from a series of both complimentary and contrasting perspectives [31].

As gamification has risen in popularity, work has been done to make gamification more accessible. One example of this is the Mozilla Open Badges initiative [32], which offers user a method to earn badges from multiple issuers, store them in a single combined "Mozilla Backpack" and display them across a variety of social networks. Similarly, Dicheva et al present a prototype gamification framework for skills based learning to enable such courses to be gamified [33].

As gamification becomes more popular, and more accessible as a tool for teachers to utilize as they try to engage their students, it is worth remembering that designing effective gamified experiences is not an easy task. Werbach and Hunter caution that designers of gamified experiences should not "mindlessly attach extrinsic motivators to activities that can be motivated using intrinsic regulators" [29].

fridaya@uef.fi

It is proposed that gamification is best considered as an easy to understand metaphor for a group of well-established techniques routed in psychology. In particularly, Deci and Ryan's self-determination theory of motivation [34] highlights the need for an appropriate level of challenge (mastery), meaningful choices (autonomy) and a (purpose). Csziksentmilayhi's concepts of flow [35] again demonstrates that challenges are most engaging when they are at a level appropriate to an individual's skill (customization). Finally, Dwecks theories on motivation applied to learning [36] promote a mindset that accepts failures as opportunities to learn, and highlights the need for activities that are low risk and provide fun failure.

Similar themes arise in the field of games design. Games designers try to engage their players in the same way as teachers try to engage their students. Schell proposes that a game can be considered a designed experience [37]. It is proposed that learning and teaching is different only in that designed learning experiences have an ultimate goal of achieving a specific learning outcome. Whilst being mindful of that difference it is proposed that some games design methodologies can be extended to any designed experience.

In particular, the Mechanics-Dynamics-Aesthetics model for game design [38] recognizes the different approaches taken by game designer, building a game's mechanics, and the game player, partaking in the game's aesthetic experience. Mechanics are the rules of the game as created by the game designer. Dynamics are behaviors, both expected and unexpected, that arise because of those rules. Aesthetics are the experiences of individual players. The same concept can be applied to anyone designing any experience for another person. In the case of learning and teaching, the teacher is responsible for the mechanics. They create teaching material to interact with, rules to adhere to, such as compulsory monitored learning activities and goals to strive towards, such as assessment. The dynamics that define the learning experience (aesthetics) might be that students focus mostly on activities that they perceive will move them closer to the assessment, or activities that will avoid incurring a penalty such as attending monitored activities. Similarly, when creating a gamified activity it is important to consider the underlying motivations of learners, and the likely effect of any additional extrinsic motivation may have on the overall learning experience.

6 Flexible Approaches to Introductory Programming, Algorithms, and Computer Science

We have here argued for the notions of smart, social, flexible and fun to be key motivation devices in design of future education delivery. In line with the above work we have applied these ideas in the following Computer Science Undergraduate contexts

- A Flexible Approach to Teach Introductory Programming

Providing options of pace and place, programming learning is supported with mini-workshop activities (akin to mini-lectures) to break the concepts down into manageable elements for students. By scaffolding their programming within a constrained environment, they focus on the specific programming construct rather than being overwhelmed or confused by needing to fit within a full programming IDE. However, the

fridaya@uef.fi

focus is on text-based programming, rather than the block based graphical approach, which can make the programming learning less effective.

- Enquiry Based Learning for Exploring Computer Science

Enquiry based learning lets the learner be the driver of their learning experience and follow their own user-derived path. Thus the flexibility is in the users own volition and task agenda.

- Guided Discovery Tutoring of Introductory Programming

Building on discovery learning (the idea the to construct new ideas users are best served by discovering new knowledge in terms of their already existing mental models [39] and [40]) the approach here was after [41] in terms of guiding users to their discovery in the context of first year computer programming.

- Gamification of Programming Exercises

Going beyond the normal quizzes, we have sought to make these games in their own right so that the underlying story is a game with programming just being the location. The idea of the interaction is fun, with programming providing a location for that fun.

- The use of social interaction techniques and discussions as a location and forum talking about programming and algorithmic issues as illustrated in Fig. 2

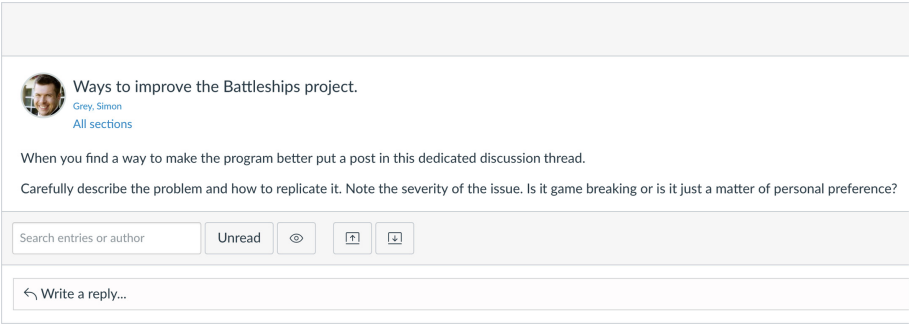


Fig. 2. An example of a social media discussion on the topic of algorithms. Here discussing ways to improve a game algorithm.

- Flexible and highly interactive methods to socially engage and provide fun for students

Whilst lectures are sometimes decried for their lack of engagement, modern technology such as in-class voting and interaction – means that students can provide feedback and actively participate in a lecture-based activity. This can range from voting on an answer, to free text (though the latter can lead to inappropriate content, so needs judicious use). This technology can enable students to see the process of research – both in gathering data, as well as the pit-falls of leading questions.

fridaya@uef.fi

- Knowledge Aware Apps for personalising services – based upon existing profiles smart knowledge ware technologies personalise the delivery of services. This can include the combination of separate needs knitted together into a single services

These therefore are starting locations for the work outlined here that we aim to move forward with in the future.

7 Conclusions

Computing's influence on educational has been significant to date, but the next generation of systems – with ubicomp providing anytime, anyplace with intelligent systems providing bespoke material tailored to the individual learner's needs – offers a step change in how people learn. The move towards cloud based learning environments, and support for augmented reality – will further soften the distinction between formal and informal learning. For traditional education providers, the challenge is in ensuring that the blend of online and ubicomp learning complements and enhances other learning. For many new learners, such blends of traditional and computer mediated learning will not be relevant, and intelligent pervasive educational systems will be providing for the lifelong learning that is so difficult in many economies and contexts.

Flat content management systems for education do not reflect what learners require and expect. They are frequently very rigid with a one size fits all mentality. The education designer is frustrated having to fit the goals of their learning outcomes into the limited tools available (e.g., multiple choice quizzes). They do not compete against the rest of the dynamic and fun elements of today's online world. If we are to generate new learning management systems, they must reflect what is around. This is not only just in terms of content, but also of interactive experience. When people go online today their experience is varied and to their own calling. In other words, it is highly flexible. We live in world where smart is all around. We live social lives online. We like having fun. The challenge for computer science is to ensure that that the next generation educational technologies are truly enhancing learning, reflect the above concerns, and so developing better ways to capture the requirements and to deliver future neophytes.

References

1. Self, J., O'Shea, T.M.: Learning and Teaching with Computers. Harvester Press, Brighton (1983)
2. Departement for Education: Schools Minister Announces Boost to Computer Science Teaching (2018). https://www.gov.uk/government/news/schools-minister-announces-boost-to-computer-science-teaching. Accessed 14 Feb 2019
3. Berners-Lee, T., Hendler, T., Lassila, O.: The Semantic Web: A New Form of Web Content That Is Meaningful to Computers Will Unleash a Revolution of New Possibilities, Scientific America (2001). https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf. Accessed 12 Oct 2018

fridaya@uef.fi

4. Höller, J., Tsiatsis, V., Mulligan, C., Karnouskos, S., Avesand, S., Boyle, D.: From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence. Elsevier, Amsterdam (2014)
5. Sleeman, D., Brown, J.S.: Intelligent Tutoring Systems. Academic Press, London (1982)
6. Madianou, M., Miller, D.: Polymedia: towards a new theory of digital media in interpersonal communication. Int. J. Cult. Stud. **15**(1), 169–187 (2012)
7. Madianou, M., Miller, D.: Migration and New Media: Transnational Families and Polymedia. Routledge, London (2011)
8. Karatzogianni, A.: Firebrand Waves of Digital Activism 1994-2014: The Rise and Spread of Hacktivism and Cyberconflict. Palgrave Macmillan, Basingstoke (2015)
9. Gordon, N.: Enabling personalised learning through formative and summative assessment. In: O'Donoghue, J. (ed.) Technology-Supported Environments for Personalized Learning: Methods and Case Studies. Information Science Publishing, Hershey (2009)
10. Appleby, J., Samuels, P., Treasure-Jones, T.: Diagnosys—a knowledge-based diagnostic test of basic mathematical skills. Comput. Educ. **28**(2), 113–131 (1997)
11. Wen, L., Brayshaw, M., Gordon, N.: Personalized content provision for virtual learning environments via the semantic web. Innov. Teach. Learn. Inf. Comput. Sci. **11**(1), 14–26 (2012)
12. Thelwall, M.: Computer-based assessment: a versatile educational tool. Comput. Educ. **34**(1), 37–49 (2000)
13. Brayshaw, M., Gordon, N.: Managing learning resources–technical issues with learning objects, repositories and searching the semantic web. In: Proceedings of the 2009 eTeaching and Learning Workshop, pp. 28–32 (2009)
14. IEEE Pervasive Computing: http://www.computer.org/csdl/mags/pc/2015/01/index.html (2015). Accessed 12 Oct 2018
15. Eisenstadt, M., Brayshaw, M., Hasemer, T., Issroff, K.: Teaching, learning and collaborating at an open university virtual summer school. In: Dix, A., Beale, R. (eds.) Remote Cooperation: CSCW Issues for Mobile and Teleworkers. Springer, London (1996)
16. Shumaker, R., Lackey, S., (eds.): Virtual, Augmented and Mixed Reality (Part 1 and 2). LNCS 8525 and 8526, Springer (2014)
17. Desire2Learn: http://www.d2l.com/ (2015). Accessed 12 Oct 2018
18. Higher Education Academy: Flexible Learning (2013). http://www.heacademy.ac.uk/flexible-learning. Accessed 12 Oct 2018
19. Zope, The UK LOM Core Home Page: http://zope.cetis.ac.uk/profiles/uklomcore (2008). Accessed 12 Oct 2018
20. LMAP Learning Materials Application Profile Scoping Study: http://www.icbl.hw.ac.uk/lmap/ (2008). Accessed 12 Oct 2018
21. Huizinga, J.: Homo Ludens: A Study of the Play-Element in Culture, p. 148. Beacon Press, Boston (1955)
22. Crawford, C.: The Art of Computer Game Design. McGraw-Hill Inc., New York (2010)
23. Koster, R.: Theory of Fun for Game Design. O'Reilly Media, Inc., Beijing (2013)
24. Goldstein, K.: Why Did Edutainment Become a Bad Word? Computers in Entertainment (2014)
25. Fenn, J., Hung, L.: Hype Cycle for Emerging Technologies. Gartner (2011)
26. Gartner: Hype Cycle for Emerging Technologies (2018). https://www.gartner.com/doc/3885468/hype-cycle-emerging-technologies-. Accessed 12 Oct 2018
27. Zichermann, C., Cunningham, C.: Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps, p. xiv. O'Reilly Media Inc., Beijing (2011)
28. Kapp, K.: The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education, p. 17. Wiley, San Francisco (2012)

fridaya@uef.fi

29. Werbach, K., Hunter, D.: For the Win: How Game Thinking Can Revolutionize Your Business, p. 26. Wharton Digital Press, Philadelphia (2012)
30. Dicheva, D., Dichev, C., Agre, G., Angelova, G.: Gamification in education: a systematic mapping study. Educ. Technol. Soc. **18**, 75–88 (2015)
31. Whitton, N.: Digital Games and Learning: Research and Theory. Taylor & Francis, Abingdon (2014)
32. Open Badges: http://openbadges.org/ (2014)
33. Dicheva, D., Irwin, K., Dichev, C., Talasila, S.: A course gamification platform supporting student motivation and engagement. In: Web and Open Access to Learning (ICWOAL) International Conference (2014)
34. Deci, E.L., Ryan. R.M.: Self-determination theory. In: Handbook of Theories of Social Psychology, vol. 1, pp. 416–433. Sage Publications Ltd, Ca (2011)
35. Csikszentmihalyi, M.: Flow. Springer, The Netherlands (2014)
36. American Psychologist, Special Issue: Psychol. Sci. Educ. **41**(10), 1040–1048 (1986)
37. Schell, J.: The Art of Game Design: A Book of Lenses. CRC Press, London (2014)
38. Hunicke, R., LeBlanc, M., Zubek, R.: MDA: a formal approach to game design and game research. In: Proceedings of the AAAI Workshop on Challenges in Game AI, vol. 4 (2004)
39. Bruner, J.S.: The act of discovery. Harvard Educ. Rev. **31**(1), 21–23 (1961)
40. Papert, S.: Mindstorms: Children, Computers and Powerful Ideas. Harvester Press, Brighton (1980)
41. Elsom-Cook, M.: Guided discovery tutoring. In: Elsom-Cook, M. (ed.) Guided Discovery Tutoring: A Framework for ICAI Research. Paul Chapman, London (1990). ISBN 0442308353

fridaya@uef.fi

Smart Mobile Learning Environment for Programming Education in Nigeria: Adaptivity and Context-Aware Features

Friday Joseph Agbo(✉) and Solomon Sunday Oyelere

School of Computing, University of Eastern Finland, Joensuu, Finland
{fridaya, solomon.oyelere}@uef.fi

Abstract. The proliferation and diffusion of mobile devices and the growing number of its users presents the opportunity for developers of educational technology to provide tools that will support both the learners and the instructors. Adaptive and context-aware features of smart learning systems are built to render support to learners in such a way that learning is possible anywhere, anytime and at the learner's convenience. Nowadays, mobile devices including smartphones, personal digital assistant, and wearable devices possess advanced technological features that can support the development of the smart learning environment (SLE). This study investigated the recent research contributions on adaptivity and context-aware features of SLE for programming education. The study also surveyed the perception of students and teachers in Nigeria higher education institutions regarding adaptive and context-aware features of SLE. To provide answers to the research questions raised in the study, we applied a mixed method approach consisting the literature review and quantitative methods. The study presents some existing solutions for programming education with adaptive or context-aware features of SLE. Finally, we proposed a model for designing an SLE with adaptivity and context-awareness features that take into cognizance the learner's preferences and needs.

Keywords: Mobile technology · Programming education · Context-aware · Adaptivity · Smart learning environment · Nigeria

1 Introduction

The dream of every learner is to acquire knowledge about subjects of interest in a most convenient and straightforward manner. Instructors also seek for ways to present knowledge to learners to enhance their learning experience. Computing education requires a pragmatic approach to make learning and teaching motivating and enjoyable. The idea of computing, where the computers can adapt to the human environment and learning condition, rather than engaging humanity to learn in certain way, would make computing activities simpler and attractive to many users [1].

Just like a smart government, smart health, smart communication and so on, a smart education is also one of the components of smart technological development. However, adaptability and adaptivity are fundamental features of the smart mobile learning system, and it is widely reported in the recent past [1, 2]. Learning technologies

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1061–1077, 2019.
https://doi.org/10.1007/978-3-030-22868-2_71

fridaya@uef.fi

employed for formal and non-formal educational settings adapt to learner's changing environment, context, and devices [3]. Adaptive learning according to [4] is based on adapting the learning methods to students' preference and learning styles. This concept deals with an individualized way of teaching/learning that will help students to learn at a fast pace, more effectively, and with better experience [4].

Smart mobile learning environment (SMLE) as used in this paper is a learning technology accessed through mobile devices such as smartphones and wearables, with features for adapting to learner's needs and preferences. Though the primary aim of mobile learning is to provide users with a learning environment that is not restricted to a single and physical location [5], the concept of smart learning allow learners to personalize learning experience and be totally immersed in the learning itself based on specific technological features—context-awareness, adaptivity, ubiquity, and social-awareness. In computer programming, the learning process for novices is expected to be simple and flexible to encourage active participation and to create a fast learning curve. Hence, it is essential to leverage the advanced technology that is capable of making the learning of computer programming interesting and taking into cognizance the learners' needs and preferences.

According to [5], context-aware, socially-aware, ubiquitous, and pervasive learning are the latest trends in harnessing learning technologies in programming education. In a smart learning system, physical environment, which forms the context of the learner, provides vital information and resources that support the high engagement of the learner. However, according to [6], many electronic learning solutions are sufficiently unaware of the context of the learner, which include individual traits, organization, work processes, and tasks. The concept of context-awareness has been defined as "any information that can be used to characterize the situation of an entity" [1]. Many pieces of literature referred to context as location, the identity of users, activity and time [7, 8]. These attributes influence the learning process, particularly in programming education. The location of learners is a critical factor that determines the level of their understanding. Identifying the type of learner and preferred ways of learning also play a vital role in the learning outcome. Also, the period within which, one engages in learning may make learning convenient or not. For instance, sorting list integers may not be convenient to practice in a rowdy environment such as a marketplace. The context-aware smart system can handle this kind of scenario by under-studying the learner to present a better learning approach and activity.

Programming education, being a core aspect of computer science is relatively considered difficult by novices. Learning or building the skill of programming takes time and need motivation [9]. Efforts have been made to ease the task of learning how to write computer programs by applying different approaches such as visualization, games and puzzle-based learning to motivate students and increase interactions among learners. [10–14]. However, not so much work has been done to consider the learner's style and characteristics, so that learning tools are adaptive to the needs, preferences, and context of the learner. Furthermore, from observation and experience with computer science students of higher education institutions in Nigeria, the majority of them find it uneasy to develop the skill of programming, but rather attend the lecture, read the theoretical materials to pass the exams and graduate to the next level. Therefore, an opportunity is created to investigate the two features of the SLE—*adaptivity*, and

fridaya@uef.fi

context-awareness, which are presented in Sect. 2. This study is important because it aims to provide insight into the design and implementation of SLE that would help students, learners, and instructors of programming education.

This research is part of the first author's doctoral studies with the aim to design and develop an adaptive context-aware SMLE for programming education in Nigeria context. Hence, the study attempts to investigate research articles that focused on context-awareness and adaptivity of learning environments for programming education in the recent past. We also seek to know the stakeholder's perception regarding these features of the smart learning environment [15] by asking the following research questions.

1. What are the recent research contributions regarding the adaptive and context-aware features of SMLE for programming education?
2. What are the students' and teachers' perception of the adaptivity and context-aware features of SMLE in Kogi State, Nigerian?

The outline of this paper is as follows: First, we review the literature that focuses on adaptivity and context-awareness of SMLE for programming education. Next, we analyze the data gathered from a survey regarding the adaptivity and context-aware features of SMLE to elicit the perception of the students. All these activities are to provide answers to the research questions. In the discussion section, we propose a conceptual model for the design of an adaptive smart mobile learning environment. The last section contains the conclusion and future work.

2 Literature Review

2.1 Mapping Nigeria: The Use of Smart Mobile Phones and Programming Education

The diffusion of smartphones in Nigeria has created the opportunity for learners and instructors to have access to solutions that enhance the learning experience and makes pedagogy flexible and exciting. Adepetun [16], reported in one of the Nigerian print media, that the number of mobile phones that penetrates Nigeria shows about 30%, and those with features of smartphones are 70%. Similarly, Statista, a statistical web page [17], estimated that at 2016 alone, Nigeria was expected to have about 15.5 million new users of smartphones [17]. The reason for the diffusion of smart mobile phones in Nigeria could be due to its affordability, portability, and its ability to enhance social interaction [18], hence many average Nigerians possess a smart mobile phone. However, some challenges have been identified to mitigate the adoption and use of these devices in Nigeria for teaching and learning [18]. They include lack of adequate electricity, cost of internet subscription, and low device storage memory.

The revelations from the survey conducted by Abel et al. [19], shows that "a large percentage of first-year students at the university in Nigeria use Android smartphones (62%). Likewise, 10% of the students own Laptop PC". This report clearly shows that the majority of students in Nigeria Universities possess smartphones compared to Laptop. This increasing number of smartphone users in Nigeria can again be linked to

fridaya@uef.fi

its durability, ease of use, and the embedded social features. The social activities—chatting, messaging, and sharing of photos, videos, and other social materials are commonly practiced among a network of friends from different locations in the world. A study recently conducted by Kareem [20], also showed that the use of tablets for secondary school's students in Nigeria to support learning had improved their performance as they can easily adapt to the technology. Based on this finding, the author recommended that mobile devices such as tablets should be made available to the students at the secondary schools by the State and Federal governments. Similarly, studies [11, 15, 21] shows that a good number of Nigerian teachers and students possess smartphones and are excited about making use of them to aid learning and other personal activities. The concept of smart learning and smart education seemed to be introduced considering the features of smartphones that could be useful to immensely transform the traditional learning process into an interactive, engaging, and active learning that entirely take into cognizance learner characteristics.

2.2 The Concept of Smart Learning

The use of information technology (IT) in the educational environment has witnessed transitions from one phase to another in the last decade. For instance, according to [22], a major paradigm shift from e-Learning—driven mainly computer-based learning, to m-Learning—powered by the mobile devices and wireless technologies, has occurred over the years [23, 24]. It is important to note that m-Learning is very similar to e-learning, but the only significant difference is the use of mobile devices and wireless technology in the later and wired technology in the former. Similarly, the advancement of mobile technology further brought about the transition of the educational environment from m-learning to ubiquitous learning (u-learning). Ubiquitous learning is defined as that environment that allows students to study anytime and anywhere through various terminals, without limitations of time, place, or environment [22, 25]. Nowadays, the field of education is becoming interesting as educational environment leverage on Social Learning (SS). Social learning makes use of smart devices and Social Networking Services (SNS) as the educational instrument, which brings about the shift from u-learning to smart learning (i.e., s-learning).

So far, there is no consensus, clear and unified definition of smart learning [22, 26]. Since the concept is still new, some scholars argued that smart learning should not be limited to just utilizing smart devices. According to Gwak [27], the concept of smart learning is focused on learners and content more than on devices; and is effective, intelligent, tailored learning based on advanced IT infrastructure. Gwak argued that although technology plays a vital role in supporting smart learning, it should not be the focus. Thus, the government, academia, and the educational industry have been working on defining and categorizing smart learning [28]. Noh [29], revealed that a concept of smart learning was proposed at the Smart Learning Korea Forum 2010 and that Korean Ministry of Education Science and Technology (MEST) defined Smart Learning as *"Self-directed, Motivated, Adaptive, Resource-enriched, and Technology-embedded (S.M.A.R.T)"* [30] cited in [22].

fridaya@uef.fi

2.3 Design Frameworks of SLEs

In this section, we present frameworks of SLEs proposed by few authors such as [31–33]. Spector et al. [32], presented a framework that focused on the fundamental characteristics of SLE. It is being considered as a preliminary and partial set of characteristics for SLEs. The framework is classified into; necessary, highly desirable, and likely. Spector's framework characteristics, from the foundational perspectives, can be mined and used as a preliminary indicator of the smartness of a learning environment [32].

In the proposed framework by [31], SLE is seen to play the role of a friend who seeks for ways to advise and encourage learners in their daily lives by taking their needs and preferences into cognition. The criteria for designing SLE formed the major components of the four-layered framework, which consist of modules: (i) for detecting learner's status; (ii) for detecting learner's performance and evaluation; (iii) for adaptive learning task; (iv) for adaptive learning content; (v) to support learning; (vi) data stores; (vii) as an inference engine

Rawal and David [33], proposed SLE framework aimed at establishing a useful link between the technologically enhanced learning domain and smart cities domain. The four-layered framework can work in the context of a smart city, which comprises a learning manager, data manager, organization manager, and technology manager.

Other models for SLE proposed by authors such as [6, 23, 31–34] has shown that the design approach is wide-ranging considering the different domain of learning. However, this current study proposed a model for the design of SMLE specifically concerned with programming education that focuses on adaptivity and context-awareness. The approach though followed the precedence of theories put forward by earlier researchers but with the contribution of context classification.

2.4 Adaptivity and Context-Aware Features of the Smart Learning Environment

In this section, we present the related discussion by researchers regarding adaptivity and context-awareness of SMLE as the features in focus. The technological features of mobile devices have provided the opportunity for one to consider SMLE that adapt to learners need and preferences taking into account the contextual information of the learners. Adaptive systems, intelligent tutoring, ubiquitous and context-aware learning systems are aimed at providing learning supports in a personalized manner based on students' learning styles, preferences, status, and other factors [31]. Zhao and Okamoto [35], proposed a method in order to ascertain mobile device capabilities and learner's preference while the personal social characteristics are used to detect the learner's behaviour. As part of the effort to model personalized adaptive content for mobile learning system, [36], presented an adaptation content architecture, which is divided into four engines: *(i) Learner Context (ii) Detector (iii) Adaptive Content Delivery (iv) Transcoding*.

According to Jones et al. [4], the concept of adaptive learning is based on the idea of adapting learning methodologies to students' learning styles. Jones et al. [4], believed that an individualized method of teaching would help students learn

fridaya@uef.fi

effectively, at a fast pace, and with more understanding. Divéky et al. [37], developed the Smart multipurpose interactive learning environment (SMILE) that encapsulates and combined the advantages of two distinct aspects, interactive educational materials, and popular computer games. SMILE concept of educational games was designed with both adaptability and adaptivity in mind. By adaptivity, it means the system attempts to be different for different learners and groups of learners by taking into account information accumulated in learner's or group's profiles; and by adaptability, it means that the system can be adapted by every user according to his needs and abilities.

According to Gwo-Jen [31], the context-aware ubiquitous learning system is capable of detecting the real-world scenario of learners by utilizing sensing technologies that interact with the system through wireless means of communication, present learning support, and offer supplementary materials or feedback through mobile devices. Krishnaswamy et al. [38], in their work, presented a framework called intelligent context-aware learning and teaching (iCALT) for obtaining real-time feedback from students and enabling the teachers to use such feedback to understand the situation of the students in order to provide the right support. As described by the author, "*iCALT is a distributed system comprised of different interfaces to support context-aware learning and teaching for the student and teacher.* " The characteristics of iCALT included (a) the ability to obtain real-time feedback from students in a lecture setting; (b) the capability to present such feedback to the teacher in a simple and easy to use format; (c) the capacity to provide support to the students.

3 Research Design and Methods

In this research, we applied a mixed method approach [39] in order to provide answers to the research questions earlier raised. According to Mckim [40], a mixed methods approach was perceived to add more value and provide the reader with a deep understanding of the study. In this study, the mixed method consists of a literature review and quantitative approach. To start with, the authors considered the first research question by conducting a literature review of recent articles relevant to adaptivity and context-aware features of SMLE. The literature review helps to map out the existing related knowledge in order to give insight and reveal the gaps that may exist in this area of research. The next methods (i.e., quantitative approach) was conducted to provide insight into the second research question regarding the perception of stakeholders regarding the adaptivity and context-aware features of the SMLE.

3.1 Data Collection

The data gathering for the literature review followed the approached described by Anohah et al. [41], which is related to this current study in terms of context. Conducting a review of literature relevant to the features of the SMLE for programming education required that we first search through the databases of publishers of scientific journals. The lists of databases that we have used consist of the ACM, Science Direct, IEEE Xplorer, and ProQuest. We also conducted manual search on Google Scholar and Computer & Education Journal databases in order to identify other related articles.

fridaya@uef.fi

In each of the databases, we used the *advanced search* option and a combined list of keywords or phrases. For example, our advanced search keywords with the Boolean Logic AND operator took the form *(Adaptivity and Context-awareness) AND (Mobile Learning Environment) AND (Learning Environment) AND (Smart Learning).* We observed that vast numbers of results were returned from Science Direct database compared to other databases. For example, Science direct returned 2,415 results from the search. Downloading all the articles might be difficult, and hence we applied criteria for download list, which include that each article title must have at least one of the keywords or phrases. A total of 135 articles were gathered and analyzed to provide an answer to the first research question. Regarding the perception of the students and teachers on the features of SMLE for programming education, a survey was conducted using questionnaire instruments where the participants gave their answers as it relates to their learning and teaching experience.

3.2 Research Context and Data Validation

To provide an answer to the second research question, data were collected from the students and teachers of the Federal University Lokoja, Kogi State Nigeria. Before embarking on the survey, the consent of the participants was sought to voluntarily partake or decline to participate in the survey. Also, a pilot survey was conducted to validate the instrument which was generated to fit the research purpose. Similarly, an expert in the field of computing was consulted to screen the instruments whether they meet the requirements and standard to be administered. The questionnaires were thereafter given out to a random sampling of students cutting across levels two to four of students majoring in computer science. The study aims to gather data regarding their perception about the adaptivity and context-aware features of SMLE. Our findings are discussed in the result section.

4 Results and Discussion

This section presents results based on the research questions and the discussion of the findings. We start by analyzing some selected literature and proceed to present our quantitative result.

4.1 Adoption of the Smart Mobile Learning Environment for Programming Education

Researchers over the years have developed systems with smart features such as adaptivity, ubiquity, context-awareness, and location-awareness, which support learning in different disciplines [42]. Similarly, models, artifacts and designs on smart learning environment have been discussed in the past. For example, a Japanese language teaching system JAPELAS [43] was developed using the concept of ubiquitous learning and deployed on PDAs with smart features. However, efforts towards developing smart systems that enhance learning of programming have not received much attention from researchers. Also, little efforts have been made towards designing

fridaya@uef.fi

models that aid teaching and learning computer programming education in particular. This section presents in brief, mobile learning environment for programming education with smart features regarding the pedagogy around teaching and learning activities.

Martin et al. [44], developed **iPro**, meaning (I can Program). An application designed to help students learn basic programming at high school and college level. iPro users program virtual robots to play soccer. The learning environment is developed for ioS and deployed on an iPod Touch mobile device. The three main design features of the tool and associated learning environments created were: (1) to enabled mobility; (2) to encourage teamwork; (3) included a full-size physical replica of the agent environment in the classroom.

Santhosh et al. [45], proposed the implementation of a solution for teaching Java programming on smartphones. Their idea is to have an innovative pedagogy that is a pointer for a paradigm shift in the academic domain where learning and teaching computer programming can all happen on a smartphone. The approach aimed at teaching programming by developing application rather than teaching programming for creating an application. The main premise of the programming education was Java. A pilot evaluation was conducted among a group of students and the result showed that students could easily understand programming concepts, interface development, programming architecture and this, therefore, improve their learning curve.

The future of teaching programming courses on mobile devices has been predicted by Tillmann et al. [46]. These researchers investigated the opportunities in teaching programming on mobile devices using a mobile programming environment- *TouchDevelop*. Their study was based on these two hypotheses: (i) teachers, and students are soon more likely to own a smartphone than a traditional PC or laptop; (ii) students will carry the smartphones with them at all times. The authors went further to conduct a 9 weeks study of developing games and applications with *TouchDevelop* among high school students. Within the 9 weeks, it was reported that the students became familiar with the environment and they can develop games and applications. Arising from the experiences gathered from the study conducted among students, the authors reported that a programming environment that runs directly on mobile devices has the potential to dramatically reduce the technical learning overhead. This is achieved by leveraging the native features available on the devices such as sensors and data instead of having to learn how to develop application, cross-compile and deploy on another device.

Similarly, Tillmann et al. [47] developed Pex4Fun, interactive game-based teaching and learning platform that can be used for teaching and learning computer programming and software engineering courses. The authors applied a gaming concept called *coding duels*. According to the description of a learning scenario by the authors, a student playing the game component of Pex4Fun learns various concepts and skills in programming or software engineering. As at the time the study was published, the authors reported that only informal interaction with the students was conducted. It was however discovered during the interaction that the students were excited about the guidance provided to them by Pex4Fun in learning programming courses.

Oyelere [48] developed a mobile learning tool for computer science education known as *MobileEdu*. According to this author "The idea of the *MobileEdu* app is not only for use in programming education but also other computer science courses, such as system analysis, design, algorithms, etc.". This mobile solution was aimed to

fridaya@uef.fi

provide support towards learning computer science courses on mobile devices in Nigeria [48]. Moreover, version 2 of *MobileEdu* has a game-based feature that is aimed at motivating learners during programming courses. The author conducted the evaluation of the tool with third-year students of a Nigerian University where the results showed that students had a better learning experience after they have undergone computer science courses using *MobileEdu*.

Gerhard [49], introduced ELM-ART as an adaptive web-based system for learning LISP programming on the world wide web (WWW). According to Gerhard, ELM-ART is capable of allowing interactivity and adaptability. The system can allow the users to specify a need and influence it to adapt (cooperative learner model). Essentially, ELM-ART provides an interface for students to communicate the learning stage or their starting knowledge. Table 1 presents an overview of some platforms and solutions developed for computer programming education.

Table 1. Some smart systems for computing education

Tools/sources	Focus/description	Learning domain	Supported technology	SMLE aspect
John and Rani [45]	Teaching programming language by developing an application (peer learning)	Java programming language	Android	Feedback mechanism
TouchDevelop [50]	Task-based learning with a little guide	Java or C++	Android	Mobility & accessibility
Pex4Fun [47]	Game-based, fun-filled learning, and teaching of programming	MOOC for computer science programming and software engineering	Web-based	Feedback and motivation
MobileEdu version 2 [48]	Puzzle-based learning environment	Introductory programming using C related code	Android	Motivation, feedback mechanism, and social-aware features
iPro [44]	Game-based tool to learn basic programming	Basic programming	iOS	Mobility & collaborative work

(*continued*)

fridaya@uef.fi

Table 1. (*continued*)

Tools/sources	Focus/description	Learning domain	Supported technology	SMLE aspect
Minerva [51]	A programming education game that adapts learning content and gameplay	General programming concepts, e.g. inputs outputs, repetition and decision	Web-based	Adaptation to learning style and motivation
Lightbot [52] cited in [53]	Puzzles-based learning platform	Learning of fundamental programming concepts (JavaScript and Python)	Web browser, Android, and iOS	Motivation and feedback
PyKinetic [54]	Python tutor	A mobile tool to learn	Android	Adaptive and personalized learning
ELM-ART [49]	LISP programming	Programming on WWW	Web	Adaptive

4.2 Students and Teachers' Perception Regarding Adaptivity and Context-Aware Features of SMLE

Tables 2 and 3 contain the descriptive statistics of students' and teachers' responses to the features of SMLE.

Table 2. Descriptive statistics of student's responses on the features of SMLE (n = 134)

Use of smart learning environment	Mean	Std. deviation
Q1. The smart learning solution I have used to learn computer programming has intelligent tutoring feature	2.56	0.95
Q2. The smart learning solution I have used to learn computer programming has a context-aware feature	2.51	0.87
Q3. The smart learning solution I have used to learn computer programming can adapt to learner's style	2.58	0.86

Comparing the statistics in Tables 2 and 3, it is observed that both categories of respondents—students and teachers, somewhat disagreed with the list of items in the questionnaire. For example, the question *"the smart learning solution I have used to learn/teach computer programming has context-aware feature"* had a mean value of 2.51 for the students and 2.33 for the teachers, while their standard deviation is 0.87 and 1.29 respectively. These results show that the degree of their responses varies and

fridaya@uef.fi

Table 3. Descriptive statistics of teacher's responses on the features of SMLE (n = 15)

Use of smart learning environment	Mean	Std. deviation
Q1. The smart learning solution I have used to teach computer programming has intelligent tutoring feature	2.67	0.98
Q2. The smart learning solution I have used to teach computer programming has a context-aware feature	2.33	1.29
Q3. The smart learning solution I have used to teach computer programming can adapt to learner's style	2.33	1.29

Table 4. Summary of descriptive statistics for student's and teacher's responses

Items	Student mean/teacher mean	Student SD/teacher SD
Question on intelligent tutoring feature	2.56/2.67	0.95/0.98
Question on context-aware feature	2.51/2.33	0.87/1.29
Question on adaptivity feature	2.58/2.33	0.86/1.29

not tilt towards positive. For clarity, a comparative summary of the findings from the analysis of the students and teachers' responses is given in Table 4.

From the statistics in Table 4, it is possible to draw an inference that the respondents have not either used SMLE or the one they use do not have the features for intelligent tutoring, context-awareness, and adaptivity.

4.3 Proposed Model for the Smart Mobile Learning Environment

The overall aim of smart learning environment is to aid pedagogy and make learning simpler and interesting to learners. One of the features of SLE that seeks to achieve this aim is the adaptivity of such system. Learning process of computer programming could be difficult especially for novices if the approach does not consider learner's individualized need, preference, and capability [18]. The conceptual adaptive model of SMLE for programming education (see Fig. 1) is the author's contribution to the existing knowledge on adaptive and context-aware learning system focused on making learning a better experience.

Adaptivity and context-awareness are vital characteristics of the SMLE. This importance is evident in the characteristics and comparative analysis of smart learning environment conducted by Darsham and Hassan [55]. A smart learning environment is expected to be adaptive to the learner's style, learning task, and learning contents to support learning. Hence, the need to pay attention to these features and to ensure that the design of SMLE for computer programming reflects learner's characteristics, environment, needs, and situations. For example, the learning styles, motivation, learning contents, cognitive capabilities, affective domain, are pointers toward providing a smart, adaptive, and intelligent system.

fridaya@uef.fi

The proposed SMLE model is a four-tier abstraction with embedded intelligent and smart features to detect the learner's status and performance. In the model, sensors and actuators are used for capturing data about the learner's context and location [56]. The next stage of the model is the context pre-processing layer where the acquired data are categorized into social trait, user's profile, learning style, cognitive ability, emotions, and user's interest or preference. Based on this categorization, the content adaptive layer communicates with the appropriate learning module to be displayed on the learner's screen.

For instance, some learners of programming may want to begin learning of programming by studying all the syntactic constructs and structures, and the keywords before delving into practicing. Others may prefer the 'learn to practice' approach. It is also possible to have learners that are motivated when an explanation of a concept is presented first then followed by a task to allow them to apply the concept and to figure out the solution (theory of constructivism [57]). In any case, all these scenarios explicate different learning situations of computer programming education.

4.3.1 Explanation of the Proposed Adaptive Model for the Design of SMLE

i. *Learning ecosystem layer:* The learners interact with their environment using the smartphone devices—having technological sensors such as camera, GPS and other actuators—relevant to capture contextual data from the environment. For example, it is possible for a learner to be at the garden, supermarket, party, on transit, or at anywhere, anytime and wanting to learn.

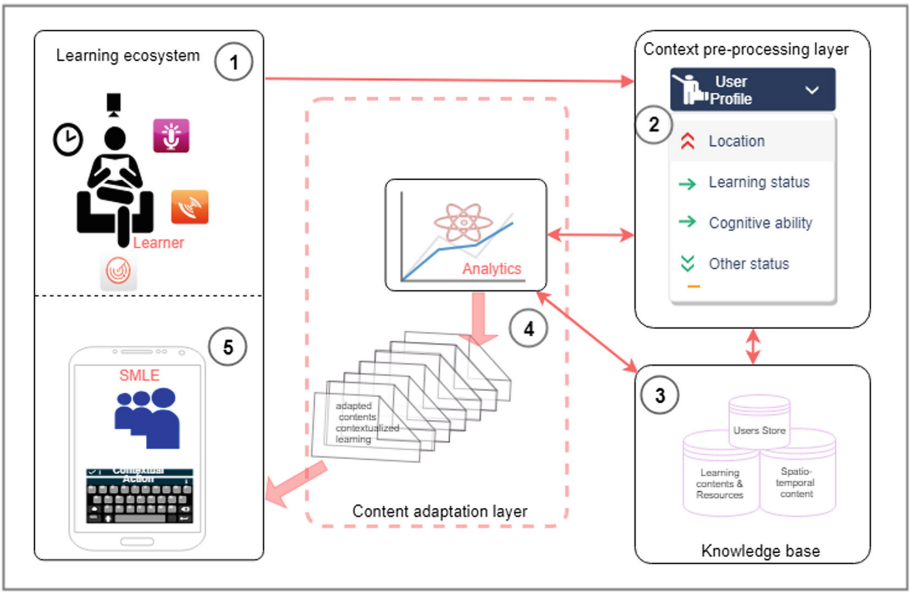


Fig. 1. A proposed conceptual model of SMLE

fridaya@uef.fi

ii. *Context pre-processing layer:* All contextual data captured from the ecosystem are categorized into learning style, social/emotional context, cognitive ability, or environmental context. Categorization is one of the new concepts introduced in this model to determine the learning contents, resources, and module to present to the learner at any instance.
iii. *The knowledge-based layer* is the repository that stores the user's profile (meta-data) and learning progress log. It is also the knowledge base for all categories of context gathered from the users' ecosystem. Besides, this layer warehouses the rules defined for analytic and intelligent decision making; communicate with the study modules to ensure that users are presented with the learning contents based on the individual learning characteristics and peculiarities.
iv. *The content adaptive layer* contains the learning module and the adaptive interface that directly feed the user interface with learning contents. When the context and status of the learners are determined, the system communicates with the database and the learning modules to feed the learning interface with the appropriate learning contents and learning resources.

The entire process of goes from step 1 to 5 as depicted in Fig. 1. It is worthy to state that SMLE is expected to be simple and flexible for users so that the aim will not be defeated. Also, when the system detects a new user without prior knowledge of the domain, i.e., a novice, the system begins with such a learner from the simplistic module of the learning resources to the advanced aspect. The proposed SMLE model is an extension from the past researchers' contribution by introducing the concept of context categorization and streamlining of processes into four compact layers. A unique feature introduced into the model is the learning analytics that evaluates the learner's performance and learning progress to have the right feedback that determines the next stage of learning or to rather store the result of the analysis on the database for future reference. The model aimed to help designers of smart learning environments for programming education where teaching and learning of programming education are modularized for an enhanced learning experience.

5 Conclusion

This research investigates the perception of the students and the teachers regarding the adaptive and context-aware features of SMLE. It also reviewed several pieces of literature and presented an analysis of recent studies regarding the use of SMLE features for teaching and learning. The study has shown that, to the extent of the reviewed literature, there is no precise description of the model for designing a smart learning environment for programming environment reported by authors. In addition, some existing frameworks in the literature regarding smart learning were presented. Finally, we proposed a model that focused on the use of SMLE for programming education, which can guide educational technology developers to achieve an adaptive, context-aware, and user-centered SMLE.

Based on the outcome of this study, we strongly recommend the adoption of the proposed model while designing SLEs to enhance learners' experience. The educators

fridaya@uef.fi

and other relevant stakeholders should also consider the integration of SMLE into the mainstream of computer science education in higher education institutions. One of the limitations of this work is the fact that the proposed adaptive model for designing SMLE for programming education has not been evaluated. As part of the author's research progress, our next work is to apply the proposed adaptive model in the design of smart mobile learning environment for computer programming in Nigeria context and investigate its efficacy. This evaluation can also be done by other researchers working in the same area of interest.

References

1. Soylu, A., Causmaecker, P.D., Desmet, P.: Context and adaptivity in context-aware pervasive computing environments. In: Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing, IEEE, pp. 94–101 (2009)
2. Gary, M.O., Olson, S.J., Bill, C.: Adaptive user support. In: Oppermann, R. (ed.). L. Erlbaum Associates Inc., Hillsdale (1994)
3. Richard, A.W.T.: Framework for Context-Aware Learning Systems, 298th edn. Publications of the University of Eastern Finland, Joensuu (2017)
4. Jones, V., Jo, J.H.: Ubiquitous learning environment. An adaptive teaching system using ubiquitous technology. In: Conference Proceedings of Australasian Society of Computers in Learning in Tertiary Education (2004)
5. Teemu, H.L., Mike, J.: Survey on context-aware pervasive learning environments. Int. J. Interact. Mobile Technol. **3**(1), 70–76 (2009)
6. Lau, S.B.-Y., Lee, C.-S.: Context aware reference model: architecture and implications for adaptation of learning activities. In: Proceeding of 4th International Conference on Mobile Technology, Applications and Systems (Mobility 2007) ACM, Singapore, September 2007
7. Ryan, N., Pascoe, J., Morse, D.: Enhanced reality fieldwork: the context-aware archaeological assistant. Computer Applications in Archaeology (1997)
8. Dey, A.K., Abowd, G.D.: Towards a better understanding of context and context-awareness. Georgia (1999)
9. Maleko, M., Hamilton, M., D'Souza, D.: Novices' perceptions and experiences of a mobile social learning environment for learning of programming. In: Proceeding in 12th International Conference on Innovation and Technology in Computer Science Education (ITiCSE), Haifa, Israel, July 2012
10. Kazimoglu, C., Kiernan, M., Bacon, L., Mackinnon, L.: A serious game for developing computational thinking and learning introductory computer programming. Soc. Behav. Sci., ScienceDirect-Procedia **47**, 1991–1999 (2012)
11. Oyelere, S.S., Suhonen, J., Wajiga, G.M., Sutinen, E.: Design, development, and evaluation of a mobile learning application for computing education. Educ. Inf. Technol. **23**, 467–495 (2017)
12. D'Souza, D., Hamilton, M., Harland, J., Muir, P., Thevathayan, C.: Transforming learning of programming: a mentoring project. In: Proceedings of the 10th Australasian Computing Education Conference, Wollongong, Australia (2008)
13. Stamouli, I., Doyle E., Huggard, M.: Establishing structured support for programming students. In: Proceedings of the 34th American Society of Engineering Education ASEE/IEEE Frontiers in Education Conference, Savannah (2004)
14. Williams, L., Wiebe, E., Yang, K., Ferzli, M., Miller, C.: In support of pair programming in the introductory computer science course. J. Comput. Sci. Educ. **12**, 197–212 (2002)

fridaya@uef.fi

15. Agbo, F.J., Oyelere, S.S., Suhonen, J., Tukiainen, M.: Identifying potential design features of a smart learning environment for programming education in Nigeria. Int. J. Learn. Technol. (2019, in press)
16. Adepetun, A.: *The Guardian*. [Online] July 2016. https://guardian.ng/business-services/smartphone-penetration-hits-30-in-nigeria/
17. Statista: The Statistics Portal. [Online] April 2015. https://www.statista.com/statistics/467187/forecast-of-smartphone-users-in-nigeria/
18. Oyelere, S.S., Jarkko, S., Shaibu, A.S., Mike S.J.: Discovering students mobile learning experiences in higher education in Nigeria. In: IEEE (2016)
19. Abel, O.A., Abah, A.J., Ogbole, B.A.: Perceptions of tech-augmented learning in basic mathematics among university students: a case of matrix algebra tools. Int. Refereed J. Arts Sci. Commer. **6**(1), 121–131 (2018)
20. Kareem, A.A.: Comparative study of the effects of computer assisted instruction on students' academic achievement in science subjects in high schools in Osun State, Nigeria. Int. J. Innovative Technol. Integr. Educ. **1**(1), 15–22 (2018)
21. Osang, F.B., Ngole, J., Tsuma, C.: Prospects and challenges of m-learning implementation in Nigeria: case study National Open University of Nigeria (NOUN). In: International Conference on ICT for Africa, pp. 1–13 (2013)
22. Kim, T., Ji., Y.C., Gyou, L.B.: Evolution to smart learning in public education: a case study of Korean public education. In: IFIP International Federation for Information Processing, vol. 295, pp. 170–178 (2013)
23. Kim, S.J.: Developing a framework of u-Learning for research and practice. Doctoral dissertation, Yonsei University, South Korea (2010)
24. Lee, B.G., Kim, S.J., Park, K.C., Kim, S.J., Jeong, E.S.: Empirical analysis of learning effectiveness in u-learning environment with digital textbook. KSII Trans. Internet Inf. Syst. **6**(3), 869–885 (2012)
25. Chen, Y.S., Kao, T.C., Sheu, J.P., Chiang, C.Y.: Mobile scaffolding aid-based bird watching learning system. In: Proceeding of IEEE International Workshop on Wireless and Mobile Technologies in Education, pp. 15–22 (WMTE 2002) (2002)
26. Zhu, Z.-T., Yu, M.-H., Riezebos, P.: A research framework of smart education. In: Smart Learning Environments, pp. 1–17. Springer, Heidelberg (2016). https://doi.org/10.1186/s40561-016-0026-2
27. Gwak, D.: The meaning and predict of Smart Learning. In: Smart Learning Korea Proceeding, Korean e-Learning Industry (2010)
28. Kim, T., Cho, J.Y., Lee, B.G.: Evolution to smart learning in public education: a case study of Korean public education. In: International Federation for Information Processing (IFIP), pp. 170–178 (2012)
29. Noh, K.S.: An exploratory study on concept and realization conditions of Smart Learning. J. Digital Policy Manag. **9**(2), 79–88 (2011)
30. Science and Technology MEST: Ministry of Education: smart education promotion strategy. President's Council on National ICT Strategies, Republic of Korea (2011)
31. Gwo-Jen, H.: Definition, framework and research issues of smart learning environments - a context-aware ubiquitous learning perspective. Springer Open J. **1**, 4 (2014)
32. Spector, J.M.: Conceptualizing the emerging field of smart learning environments. Smart Learn. Environ. Springer Open J. **1**, 2 (2014)
33. Rawad, H., David, L.: Towards a smart learning environment for smart city governance. In: International Conference on Utility and Cloud Computing, pp. 185–190 (2016)
34. Ronghuai, H., et al.: A Conceptual Framework for a Smart Learning Engine. Springer Science Business Media, Singapore (2017)

fridaya@uef.fi

35. Zhao, X., Okamoto, T.: A device independent system architecture for adaptive mobile learning. In: The 8th IEEE International Conference on Advanced Learning Technologies (ICALT'08) (2008)
36. Zhao, X., Fumihiko, A., Ninomiya, T., Okamoto, T.: Personalized adaptive content system for context-aware mobile learning. Int. J. Comput. Sci. Netw. Secur. IJCSNS **8**, 153–161 (2008)
37. Divéky, M., Jurnečka, P., Rudolf, K., Ľuboš, O., Bieliková, M.: Adaptive educational gameplay within smart multipurpose interactive learning environment. In: Second International Workshop on Semantic Media Adaptation and Personalization (2007)
38. Krishnaswamy, S., et al.: iCALT: intelligent context-aware learning and teaching environment. In: Proceeding of 10th IEEE International Conference on Advanced Learning Technologies (2010)
39. Oyelere, S.S., Paliktzoglou, V., Suhonen, J.: M-learning in Nigerian higher education: an experimental study with Edmodo. Int. J. Soc. Media Interact. Learn. Environ. **4**, 43–62 (2016)
40. McKim, C.A.: The value of mixed methods research: a mixed methods study. J. Mixed Methods Res. **11**(2), 202–222 (2017)
41. Anohah, E., Oyelere, S.S., Suhonen, J.: Trends of mobile learning in computing education from 2006 to 2014: a systematic review of research publications. Int. J. Mob. Blended Learn. **9**(1), 16–33 (2017)
42. Lay, D.: Ubiquitous Computing, vol. 2, pp. 66–79. British Educational Communication and Technology Agency (2007)
43. Yen, C., Ogata, H., Yano, Y.: JAPELAS: supporting Japanese polite expressions learning using pdas towards ubiquitous learning. J. Inform. Syst. Educ. **3**(1), 33–39 (2004)
44. Martin, T., Berland, M., Benton, T., Smith, C.P.: Learning programming with IPRO: the effects of a mobile, social programming environment. Int. J. Interact. Learn. Res. **24**, 301–328 (2013)
45. Santhosh, J., Rani, M.S.: Teaching java programming on smartphone-pedagogy and innovation; proposal of its ontology-oriented implementation. Procedia Soc. Behav. Sci. Published by Elsevier **176**, 787–794 (2015)
46. Tillmann, N., Xie, T.: The future of teaching programming is on mobile devices. In: Proceedings of Annual Conference on Innovation and Technology in Computer Science Education, Haifa-Israel (2012)
47. Tillmann, N., Halleux, J., Tao, X., Gulwani, S., Judith, B.: Teaching and learning programming and software engineering via interactive gaming. In: Proceeding of the 2013 International Conference of Software Engineering ICSE'13, San Francisco, pp. 1117–1126 (2013)
48. Oyelere, S.S.: Design and development of a mobile learning system for computer science education in Nigerian higher education context. Joensuu: Ph.D. thesis, University of Eastern Finland (2017)
49. Gerhard, W.: Adaptive learning systems in the world wide web. In: UM99 User Modeling, pp. 371–377. Springer Science+Business Media, New York (1999)
50. Nikolai, T., Halleux, J.: Pex4Fun: Teaching and Learning Computer Science via Social Gaming. In: IEEE (2011)
51. Renny, S.N., Hasanov, L.A., Laine, T.H.: Improving play and learning style adaptation in a programming education game. In: Proceedings of the 9th International Conference on Computer Supported Education (CSEDU) (2017)
52. Lightbot. [Online]. https://lightbot.com/flash.html. July 2017
53. Figueiredo, J., García-Peñalvo, F.J.: Improving computational thinking using follow and give instructions. Assoc. Comput. Mach., ACM **978**(1), 4503–5386 (2017)

fridaya@uef.fi

54. Fabic, G.V., Mitrovic, A., Neshatian, K.: Adaptive problem selection in a mobile Python tutor. In: Proceedings of ACM User Modeling, Adaptation and Personalization Conference, Singapore (2018)
55. Darsham, S.A., Hassan, M.: Pursuit of smart learning environments for the 21st century. In-Progress Reflection No. 12 on Current and Critical Issues in Curriculum, Learning, and Assessment (2017)
56. Aziz, H., Laine, H.T.: A survey of context-awareness in learning environments in 2010–2016. In: Proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017), vol. 1, pp. 234–241 (2017)
57. Mordechai, B.: Constructivism in Computer Science Education. ACM, New York (1998)

fridaya@uef.fi

An Integrated Approach for Educating the Marine Forces Reserve Inspector-Instructor Cadre

Alejandro S. Hernandez(✉) and Lisa R. Spence

Naval Postgraduate School, Monterey, CA 93943, USA
ahernand@nps.edu

Abstract. This paper describes the application of a systems engineering approach to combine live instruction, web-based learning, and scenario methodologies into a unique curriculum design. This innovative instructional design serves a large, greatly distributed organization with a high turnover rate, which demands an adaptive program. By following a systems engineering model, the authors outline the evolution of an integrated system to educate the Marine Forces Reserve Inspector-Instructor cadre who minister the Marine credo and expertise to reserve units. These active duty professionals often have no, or limited experience with Marine Reservists or policies. An annual orientation conference was the lone organizational education event that prepared new cadre before reporting to their assigned units. The Marine Forces Reserve leadership judged the preparation of I-I cadre as insufficient for the challenges that can emerge in a reserve assignment. For the past four years, the Naval Postgraduate School and the Marine Reserves have seamlessly fused online education, formal lectures, and scenario techniques as a sustained transfer of information and knowledge that cadre can reference throughout their tenure with reserve units. The systems engineering process that brought these different techniques together presents a formula for creating new instructional designs for similar audiences and conditions.

1 Overview

The United States Marine Corps (USMC) is one of the most formidable expeditionary forces in history. An ability to project power rapidly anywhere in the world is a significant comparative advantage for the US. The success of the Marines, especially in the current global environment, is dependent on the reserve forces that augment the ranks of active Marine components. Preparing reservist to support active duty units is the primary mission of the Marine Forces Reserve (MARFORRES). To achieve their objectives, the Marine Corps assigns active duty Marines to serve as Inspector-Instructors (I-I) at Marine Corps reserve units. As such, a key component to MARFORRES success is the knowledge of the I-I cadre [1]. However, most I-I lack prior experience working with the unique challenges that are in reserve units. The I-I knowledge about their job is mainly through somewhat ambiguous on-the-job-training (OJT) and a deluge of slide shows. This paper describes the application of the systems engineering model to create a novel education design that the study team implemented over a four-year period.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1078–1096, 2019.
https://doi.org/10.1007/978-3-030-22868-2_72

fridaya@uef.fi

2 Background: The Marine Forces Reserve

Marine Corps leaders are acutely aware that the strategic and operational environment will remain multi-polar, dynamic, and complex. The US, China, and Russia fight to remain regional and global powers. Globalism and the emergence of regional powers in Central Asia, South and Central America, and Africa threaten actions that can have large-scale consequences. For US forces, simple military actions are seldom a solution. Economic, ethnic, cultural and/or religious factors challenge many commanders and their staff.

The USMC recognizes that its reserve forces are crucial to its success. Figure 1 shows that reservists make up 37% of the total Marine force [2]. Reservists must be ready to conduct joint operations and work in an interagency environment. Standards and achievements for training and education must be at the highest levels. The MARFORRES challenge of reinforcing active duty Marine units with trained and ready personnel comes at a time of budgetary concerns and increasingly multidimensional environments. MARFORRES may be called upon in time of war, national emergency or contingency operations. They are essential in relieving active component Marines from an intense operational tempo. Meantime, peacetime activities for MARFORRES units are a gamut of (sometimes unique) services to Marine Corps headquarters, the Department of Defense (DOD), and local communities.

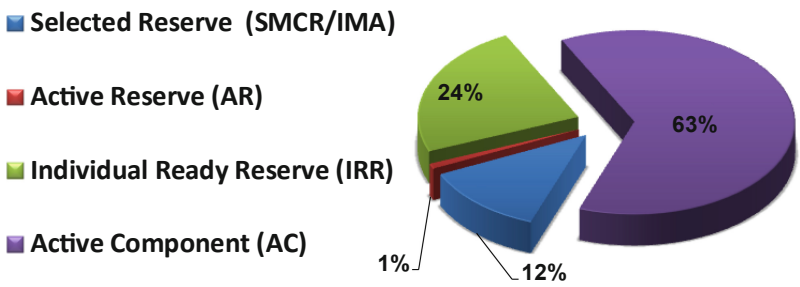


Fig. 1. Proportion of reserves in total Marine Corps [2].

Commander, MARFORRES, has Title 10 responsibility for ensuring that reserve units are equipped and trained to the same rigorous standards as active Marine forces [3]. Therefore, it is incumbent on MARFORRES to rapidly respond with trained and ready Marines to support active components. In meeting this responsibility, MARFORRES has developed a unique reserve structure. It is the only one of the US military Services to place active duty Marines as I-I to be cadre in the reserve units. This active

fridaya@uef.fi

duty staff is accountable for maintaining the daily operations of the unit until reservists report for their periodic training. During training events, the I-I hand over the control of the unit to the reserve commander and staff to conduct operations. At this point, the I-I cadre serve as advisors and subject matter experts (SME).

Inspector-Instructors have a marked disadvantage when reporting to a reserve organization. Few Marines have ever served in a reserve unit. In a recent survey of 111 newly assigned I-I, less than 15% have any previous duties with reservists [4]. This lack of experience creates a challenge for understanding the unique needs of a reserve organization. For instance, to qualify on their assigned weapons, an active duty unit would have available ranges, transportation, and ammunition already in place to conduct the event. If circumstances warrant, the planning cycle for the training consists of days, or even hours. On the other hand, a reserve unit plans and coordinates such an event *at least six months* from the actual date. There are many reasons for this vast difference in planning time. Reserve units are located outside of military installations, and therefore have no infrastructure for the event. Consequently, the cadre must request to share rifle ranges with active duty units. Transportation must be arranged to bring the reservists from their compound to the range. Additionally, separate transportation is required for the weapons, along with a security detail. These are just a few considerations that I-I must take into account, which active duty components do not. MARFORRES leaders acknowledged that the education program to prepare the cadre to meet the distinct needs of reserve units was inadequate, thus prompting requests to address the problem.

3 Challenges to Educating the Inspector-Instructor Cadre

Reservists are assigned to more than 230 selected Marine Corps reserve (SMCR) units that are spread in 47 states, including the District of Columbia and Puerto Rico (Fig. 2) [2].

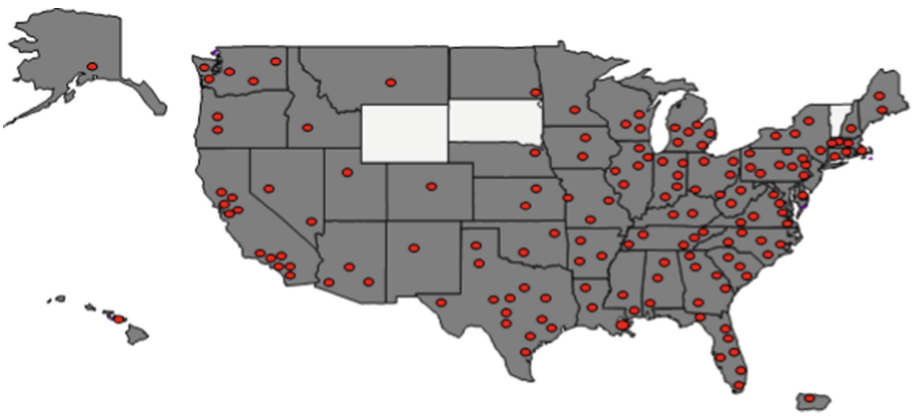


Fig. 2. Distribution of selected Marine Corps reserve units [2].

fridaya@uef.fi

Limited peer interactions, remoteness from other SMCR units and MARFORRES Headquarters, constraints to budgets and training opportunities, and relatively high turnover rates add to the difficulties in these assignments [1]. Cadre adapt relatively slowly to their reserve units. Powley and Borrego [5] find that approximately 18 months pass before cadre have a full understanding of the reserve culture (Fig. 3). Accelerating how cadre adjust to reserve surrounding can be addressed, in part, through the design of a robust I-I education program.

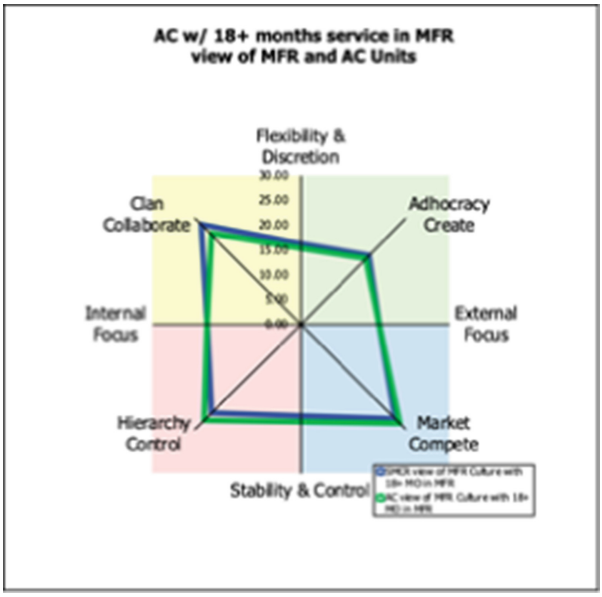


Fig. 3. After 18 months on duty, the I-I cadre's perceptions about reservists and active duty Marines begin to match [5].

To address the educational needs of the I-I cadre, MARFORRES hosts an Annual I-I Orientation Conference to educate cadre prior or early in their reserve assignment. This approach consisted of four days of PowerPoint presentations from a series of headquarter staff. Each presentation is between 30 and 120 min, and can involve over 100 slides. Overall, reviews from I-I conference participants rated the approach as ineffective. The MARFORRES operations and training (G-3/7) staff was also dissatisfied with the efficacy of a PowerPoint lecture format. The staff sought a systematic review of the program and a traceable approach for building a new one.

fridaya@uef.fi

4 A Systems Engineering Approach to Education Design

Education is a major domain in which systems engineering (SE) approaches and methods are directly applicable [6]. This is not a new idea; Gaff [7] discussed education design using a systems approach, which originates from the scientific method. Reigeluth [8] begins a description of designing and implementing new models for schooling with a discussion on system design. More recently, Reigeluth [9] explains engineering concepts and systems theory as the basis for educational systems design, which is commonly referred to as the ADDIE (analyze, design, development, implementation, and evaluation) model for instructional design [10]. Sehanovic and Krelja-Kurelovic [11] explain the development of interactive web-based education that conform to educational objectives. The principles of the approach adhere to form-follows-function, a tenet in system design and development, as well as scenario methodologies. The exponential growth of online education prompted Lister [12] to conduct research on the components of e-learning that influence course design. The approach for implementing findings from the study mirrored a functional decomposition and analysis in SE. The general principles from these examples and other education design discussions follow a SE methodology, including the development of objectives based on sponsor needs and the desired knowledge that the students should draw from the experience.

In 2014, MARFORRES established a partnership with the Naval Postgraduate School (NPS) to examine the existing I-I educational structure and to devise means to address shortcomings in the program. The NPS principal investigator (PI) used a SE approach to design the I-I education program, treating the program as a case study. The overall project required a teaming effort that continues today and has fostered the evolution and present state of the I-I education program described in the latter part of this paper.

4.1 Stakeholder Analysis and Brainstorming

The SE process necessarily begins with problem definition and needs analysis [6]. In a one-week workshop, the NPS team engaged key SME from the MARFORRES Headquarters to accelerate the knowledge flows required for effective I-I performance. These discussions highlighted the knowledgebase of experienced professionals at MARFORRES Headquarters that should make up the I-I education. The dialogue identified not only essential I-I knowledge, but important constraints and limitations that impede current I-I knowledge flows [1]. An examination of the existing I-I education program with the SME established a structure for the fundamental knowledge that an I-I requires. Figure 4 illustrates the major areas of the educational material.

This thorough and candid session with MARFORRES is an essential phase in the SE process. Stakeholder analysis describes the needs of the sponsor for the given problem [6]. The results of the workshop set the stage for defining requirements for the system of interest—the I-I education program.

fridaya@uef.fi

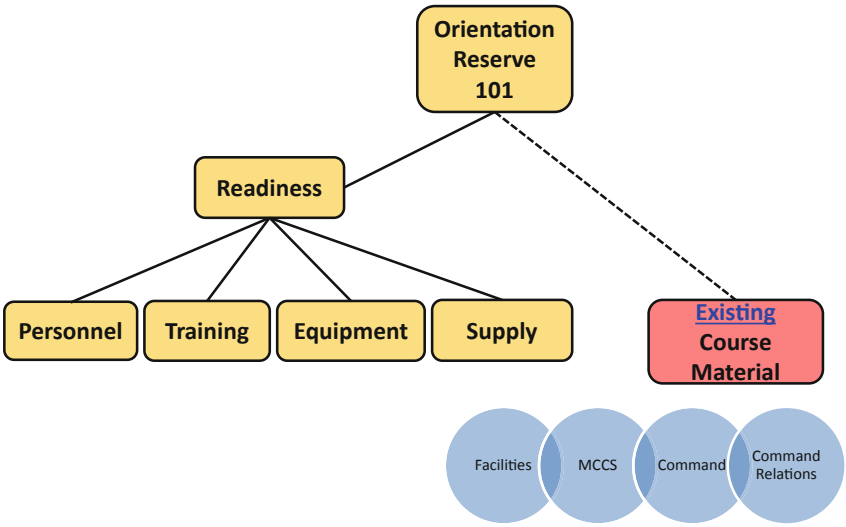


Fig. 4. Knowledge areas that are essential to I-I success in SMCR units [1].

4.2 Functional Decomposition and Requirements Development

The overarching system level requirement for the education program is to educate all I-I about the information in each of the major areas of knowledge (Fig. 4). Additionally, requirements reflect means to address obstacles that have impeded I-I education. Marine Corps regulations and policies have already established the content. The true challenge in this project is in the design and construction of an education framework that conveys the material in a manner that meet MARFORRES needs and objectives.

A functional analysis and objective hierarchy are SE techniques that analyze the engineering problem and guide the design and development of a useful education system for the I-I cadre. Functional analysis is a way to describe **what** the new education format should do. An objective hierarchy constructs MARFORRES's desired outcomes from a new education program. Using these tools, the team of NPS faculty, instructional design engineers, and MARFORRES SME can create a set of workable options.

Functional analysis is a decomposition of the system's primary purpose. Each successive breakdown of a function provides greater detail about the activities that must occur for the overall system to achieve its main purpose. The MARFORRES objectives, as described by the SME, reflect the functions in terms of a stakeholder's specified criteria. Figure 5 presents a functional tree for the I-I education system that encompasses the overarching requirement for the systems and MARFORRES stakeholder objectives.

Compiling functions and sub-functions into categories aids in defining system requirements and shaping the modality of the education program [6]. For instance, the system may have to provide near-instantaneous access to reference material and expert knowledge. This requirement addresses the aggregate of the four highlighted or

fridaya@uef.fi

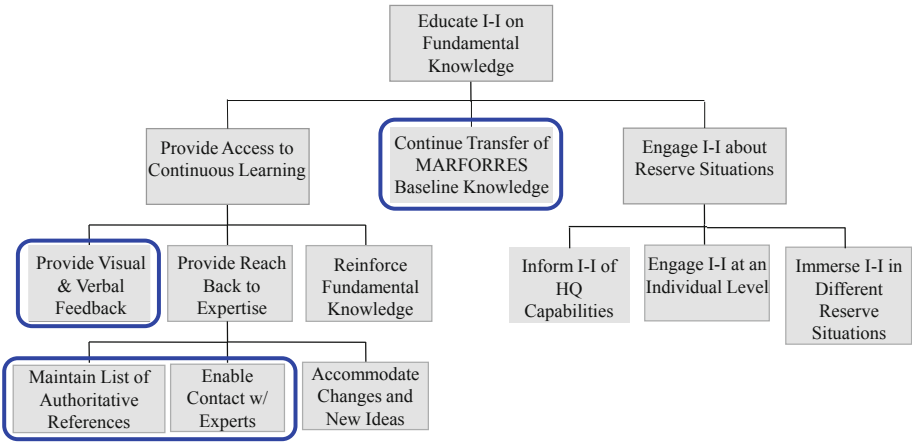


Fig. 5. Functional decomposition of the I-I education program.

outlined functions in Fig. 5: Maintain a List of Authoritative References, Provide Visual and Verbal Feedback, Enable Contact with Experts, and Continue Transfer of MARFORRES Baseline Knowledge. Continuing in a similar manner, the team developed a set of requirements based on **what the system must do**. These requirements drive the design process. During preliminary and conceptual design, the functional tree plays a major role in creating alternatives.

4.3 Component to Function System Design

As previously discussed, the SE design phase follows the principle of form-follows-function. This point is manifest in the process to develop a generic design of a system that will address the set of requirements. For each function, there should a component or group of components that achieves at least one function. A primary tool for developing system design options is a morphological box [13]. This technique is a combinatorial method for conceptually arranging the components of a system to create a whole that achieves the system's primary purpose.

There are four main steps for creating a morphological box or matrix. The process begins with identifying system attributes, elements, or functions that are necessary to realize the system's purpose. Most often, the functions of the system are the main drivers of system design. These functions comprise the columns of the matrix. Under each column, the designer generates as many feasible alternatives that will achieve the function. Selecting one alternative from each column produces one instantiation of a candidate design. As an example, the combination of unique designs for a system with five main functions, each with four alternatives is 4^5 or 1,024 possible design solutions. In the case of the I-I education program, the top three functions from Fig. 5 are used to build the simple morphological box that is shown in Table 1.

For the example in Table 1, there are 27 different combinations that can address the primary system purpose. As the number of alternatives for each function grows the set of candidate solutions increases exponentially. The shaded boxes in Table 1 was the

fridaya@uef.fi

Table 1. This morphological box demonstrates the technique with only the top-level functions of the I-I education program.

Provide Access to Continuous Learning	Transfer MARFORRES Baseline Knowledge	Engage I-I about Reserve Situations
Share Drive	Library: Book of Knowledge	Practical Exercises
Website Portal	Learning Environment	Seminars\ Testimonials
Interactive Site	Formal Lectures	Simulator

baseline approach that the MARFORRES used to educate the I-I cadre. Several workshops and continued dialogue with MARFORRES staff and SME generated numerous options beyond the baseline. Through iterative evaluation of options and objective value generation, which are not part of the discussion in this paper, the NPS team and MARFORRES leadership agreed with an education design that met MARFORRES needs and constrains. Contained within the final design decision were fixed constraints. In one case, the MARFORRES Commanding General ordered that formal presentations about MARFORRES processes and history be part of the final design solution. Such objectives are nonfunctional requirements, but are necessary to meet stakeholder needs.

5 A Framework for the I-I Education Program

The design processes described in the previous sections resulted in a system design that satisfied stakeholder needs and system requirements. There are three major components of the I-I education program: a formal orientation event, an online course and reach back portal, and practical exercises. These components map onto specific education techniques. The following sections describe the development and integration of these education components, including a discussion of the theoretical basis for using the specific technique.

5.1 Maintain Baseline Formal Lectures

Commander, MARFORRES directed that the annual orientation conference continue. The conference is not only an education channel, it serves to initiate and develop relationships among the I-I cadre, as well as the MARFORRES Headquarters staff. Therefore, the conference is also a natural platform to introduce basic knowledge about the reserves, and particular internal processes that are not present in active duty units.

Instructional development or improvement is a confluence of the organization and the instructors (faculty). Powley and Borrego's [5] study on the reserve culture shows initial results that classify reserve units as clannish and hierarchical. Consequently, a formal lecture style of instruction is understandably a component that the MARFORRES leadership would wish to keep intact.

fridaya@uef.fi

The formal lectures that remain from the previous education format consists of information from the MARFORRES Headquarters primary staff. These major areas include instruction on personnel administration, operations and training, and supply, logistics, and support processes. Other presentations to establish a knowledgebase are scheduled throughout the conference, alternating with other events such as introductions to the I-I's major subordinate command (MSC), the higher headquarters to which a SMCR is assigned. While it is not technically an educational element, these meetings and command relationships are essential to the I-I cadre's success. Therefore, a systems engineering solution must consider the high value with which stakeholders regard the event.

5.2 An Online Portal for Education Modules

Another element of the I-I education program is an interactive, web-based learning (WBL) environment. This component of the education system leverages the advantages that WBL presents [14]. These benefits have grown considerably in the past decade. Using the initial stakeholder workshop, the NPS team designed and developed an online portal that guides incoming I-I through a series of education modules, which introduce MARFORRES processes and policies. This online learning experience prepares the new I-I to attend the formal Annual I-I Orientation Conference. This portal has evolved over the past four years since the initial launch of the online portal and is now an integral part of the I-I education program.

Early study of the MARFORRES in 2014 identified the organization as one that needs rapid knowledge movement [1]. Nissen [15] provides detailed explanation of the five dimensions of knowledge flow depicted in Fig. 6. The initial project describes the relationship between the primary tools (slideshows and OJT) that MARFORRES uses for educating the I-I and the five dimensional axis of knowledge flow [1]. The placement of these tools in this five-dimensional space helps direct remedies for shortfalls in the existing education program.

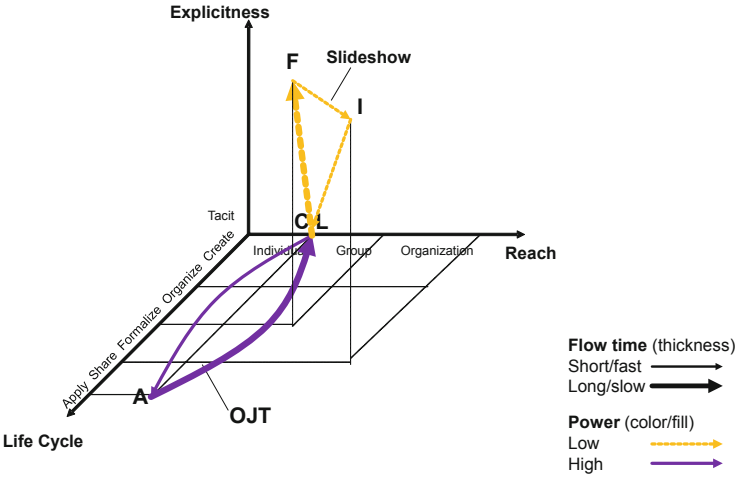


Fig. 6. Principal I-I knowledge flows [1, 15].

fridaya@uef.fi

To orient the reader, Nissen et al. [1] offers a brief description of each axis in Fig. 6. The vertical axis (Explicitness) is the degree that knowledge has been delivered in concrete terms or form. Near the origin is the amount of tacit knowledge (e.g., experience, know-how, understanding) that the audience possesses. The range from tacit to explicit endpoints represents the level of knowledge. The horizontal axis (Reach) is the extent to which knowledge has spread through an organization (e.g., flowing from individuals, through groups, to organization). The axis for "Life Cycle" is the activity associated with knowledge and includes five classes (i.e., create, organize, formalize, share, apply). These three axis are orthogonal and behave as independent variables. "Flow Time" is a fourth dimension that represents the speed that knowledge flows. The thickness of a line signifies speed. Lastly, "Power" is the degree that knowledge is actionable. High power is represented with a solid purple line and low power is seen as orange dotted lines. Nissen [15] provides more information about the five dimensions of knowledge flow.

As a result of these studies, the team moved to apply advances in information technology. The graph shows that slideshows, which was the primary means to impart knowledge to the I-I, has very low power, while OJT or individual hands on experiences have high power. These results align with the educational theory of constructivism, which asserts that people learn best through experiences [10]. Methods to engage I-I individually became an imperative for creating a foundation for building their knowledge about the reserves. Dalton [16] points out that technology has changed how individuals retain information and facilitate blended learning. In the relatively recent past, education design has leaned heavily on internet capabilities [17]. The proliferation of online education throughout the globe is just one signal for the evolving nature of education design used for interactive web-based learning [11].

Employing a class of systems that are mainly used for distributed and remote learning, the team created a prototype distance or distributed learning (DL) site within the NPS learning management system (Sakai, https://sakaiproject.org/). Through this initial education portal, the NPS team introduced preliminary training modules to I-I attendees of the 2014 I-I Orientation Conference. Post-conference survey results indicated that continued development of the site could pay great dividends for improving the I-I cadre knowledgebase. Figure 7 presents the prototype of the site on the top and the latest version of the site on the bottom. The continued refinement of the DL site is a function of the configuration control board (CCB) that is discussed later in this paper.

5.3 Content Creation and Curation

MARFORRES SMEs had a large amount of content created, but no central repository or cohesive message to direct the study of the content. The NPS team worked with SMEs to create a "knowledge management project" that would eventually become a consolidated online course.

In the first phase of development, the NPS team applied a 5-step AGILE (align, get set, iterate and implement, leverage and evaluate) design methodology to guide MARFORRES (SMEs) in converting their existing PowerPoint presentations into embedded HTML Image files within the learning management system. This design

fridaya@uef.fi

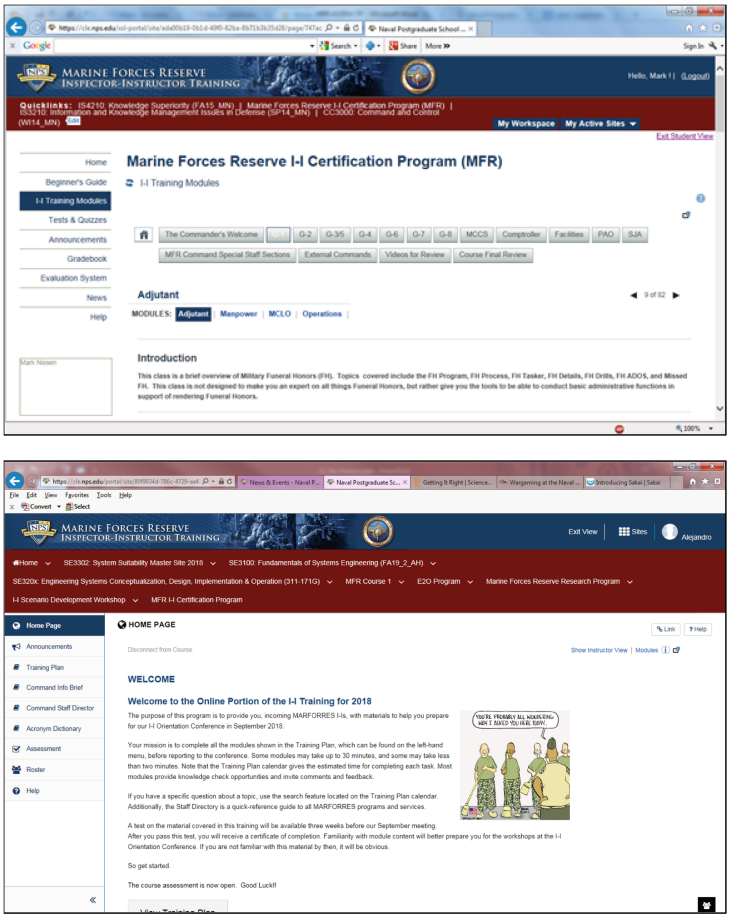


Fig. 7. Prototype (top chart) and latest (bottom chart) versions of the I-I web-based distance learning modules.

methodology focuses on the user to create useful experiences. From these source documents, the team created narratives for the presentations within the Sakai environment. Users have access to these presentations at any time, thereby meeting the "retain organizational knowledge" goal to some degree.

This first prototyping approach identified key obstacles for developing the environment. The primary challenge was that the MARFORRES SMEs had learned one way to provide instruction, which involved a heavy amount of textual information that had to be read from a PowerPoint presentation. Turning the resident face-to-face presentation into narrated online presentations did not improve training and knowledge transfer. Secondly, the time required for SMEs to become designers and developers of content was a major burden for an already overtasked staff.

An additional challenge was the unfamiliarity of MARFORRES leadership with the new approach for delivering educational information. The uncertainty with the new

fridaya@uef.fi

method raised concerns about the necessity for a new approach and the likelihood of success. Deviation from the norm was a substantial cultural shift for the command. The first two years of the I-I Certification Training project included educating leadership on the benefits of applying a more pedagogically sound approach to training than had been formerly used. Change in the MARFORRES leadership during the project phases also required time and effort in these informational and educational efforts.

Simultaneously, the NPS team determined that a curriculum review and update of I-I training content was necessary. The team facilitated a one-week workshop hosted at the Naval Postgraduate School with selected MARFORRES SMEs and leadership. The goal of workshop was to identify the major components (courses) of the training curriculum and identify a detailed approach for the subsequent development of distance learning modules and I-I workshop training sessions. At the end of the workshop, the primary curriculum topics had been identified and were configured to support two components of training: online distance learning modules and face-to-face instruction.

To facilitate effective knowledge transfer, participants are required to complete the DL modules before attending the MARFORRES Annual I-I Orientation Conference. This preliminary requirement ensured that all I-Is had a common entry-level knowledge. The Conference would provide the opportunity to apply this knowledge during practical exercise scenarios, which would expand and further develop the incoming I-Is essential knowledge and skills.

5.4 Scenario Methods: Practical Exercises

Early discussion with SME and leaders made it clear that MARFORRES is a learning organization that greatly values experiential data. Nissen's [15] five dimensions of knowledge flow show that OJT has high power (thick purple line) with individuals. The challenge for the design team was how to inject experience in the I-I education program.

There are tremendous efforts to understand human behavior: human—machine interfaces, technology transition and adoption, and work force effectiveness and efficiency [18]. Human dynamics in context of an organization is among one of the most complex to study. Consequently, Commanders expend time and energy to fathom holistic methods to explore these most difficult of problems. The team turned to time-proven techniques of systems thinking and scenario methods to develop this critical component for educating the I-I cadre.

A scenario describes a possible future situation [19]. Once defined, scenarios provide a vehicle to study and understand various futures while not attempting to predict the future [20]. Scenario methodology is the set of assumptions, rules, and tools that enable exploration of future conditions or circumstances. Scholars relate subsets of scenario methods that include scenario building and planning, scenario analysis, and scenario or systems thinking [21, 22]. Scenario techniques present a new class of methods of inquiry that produces fresh ideas [23]. Using a scenario approach to create new experiences for the I-I increases the knowledge about the reserve culture.

The application of scenario techniques to develop I-I knowledgebase has a number of advantages that are useful for the I-I's own understanding about the duties they will perform. Scenarios examine a subject's response to different circumstances [20].

fridaya@uef.fi

Providing a range of behavior from various stimulants increases the likelihood for responding to situations with a better understanding of the overall picture, especially if there is an application of multiple comparison correction techniques [24]. Varying the set of circumstances has residual value; it opens the mind to more creative thinking and increases the mind's pliability to new situations. Gharajedaghi [25] explains that the ability to see multidimensional relationships of a problem is the most potent principle of systems thinking, which reflects those in scenario thinking. Most importantly, scenarios identify weaknesses or shortfalls relative to a baseline. Processes in applying scenario methods underscore a communications loop between the conflict situation, an individual's knowledge, and the ability to connect the two.

The NPS team developed a one week workshop with MARFORRES HQ staff and selected I-I cadre to serve as facilitators. During the workshop, the team builds scenarios for specific situations that the MARFORRES SME have determined as the most troublesome for I-I cadre. The I-I facilitators administer these practical exercises during Annual I-I Orientation Conference.

Chermack [21] introduces a process to construct a scenario. While other scenario practitioners such as Meitzner and Reger [20], Perla and Curry [26], Lindgren and Bandhold [27], and Maack [28] offer modified descriptions, the differences are negligible. We apply Chermack's approach for developing a scenario. The process consists of five phases that are grouped into two stages: (1) Scenario building, and (2) Scenario deployment. Preparation, exploration, and development comprise activities in Stage 1. Stage 2 includes implementation and assessment. The elements of Stage 1 highlight the need to define the purpose of the scenario and its objectives. It requires an exploration or research about the audience and the situation. Finally, the scenario must be created to meet the purpose and objectives that had been articulated. Stage 2 introduces the scenario to the audience. Assessment comes in two flavors, rating the goodness of the scenario and evaluating the results from the audience. Figure 8 illustrates this process.

Construction of a scenario system highlights key elements in each phase. We describe these steps.

- *Preparation*: The overarching purpose of any scenario is to provide an I-I with experience about a specific reserve situation. The objectives is to create an awareness about an I-I's duties as a whole.
- *Exploration*: Scenario builders must gain an understanding about the type of individuals who will be assigned to MARFORRES. Their nominal background about operational doctrine, reserves, and administration are useful information. Operational and simulation subject matter experts (SME) provide a baseline solution for different tasks.
- *Development*: In past years, MARFORRES has selected five reserve-unique situations: force protection, training, personnel finance, logistics support, and immersion in reserve culture. Separate vignettes (or scenario injects) in the scenario can tease out more nuanced responses from the I-I.
- *Implementation*: There is one facilitator team (usually two persons) for each practical exercise. The MARFORRES HQ staff support the discussing in the implementation stage. The G-3/7 provide the venue, equipment, or material that is required for the exercise. There is a timeline associated with any given exercise. At times a stressor scenarios can be introduced during the exercise.

fridaya@uef.fi

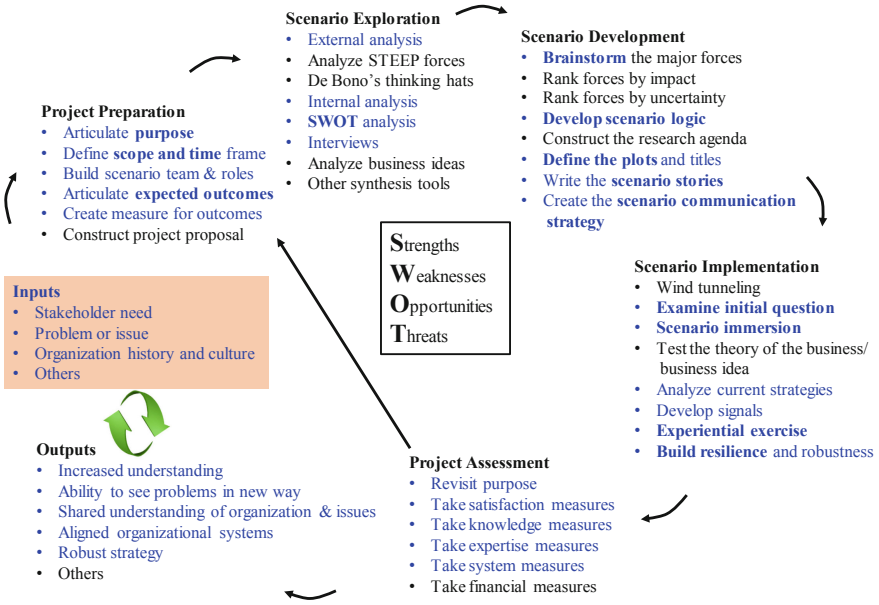


Fig. 8. A systems approach for creating performance-based scenarios [21].

- *Assessment*: There is no formal assessment of the responses that the audience gives in the exercise. Facilitators discuss best practices for the situation. Additionally, the MARFORRES staff surveys I-I attendees to gain feedback about the exercises, as well as the whole Orientation experience.

5.5 Configuration Control Board

The previous sections have discussed the major components of the I-I cadre's education program. However, the program does not remain in stasis. Feedback from the I-I attendees, change in operational mission, the emphasis of a new MARFORRES Commander, and other factors can influence the educational needs, and in turn the design of the program. To systematically manage the design of the I-I education program, the NPS team recommended formation of a configuration control board (CCB). Members of a CCB usually include the Program Designer and Developer, Evaluator, User Representative (I-I cadre), Technical Support, Program Operators or Instructors, Budget Personnel, Program Manager. This list is not comprehensive, but attempts to cover critical personnel for modifying the system.

The CCB prioritizes, approves, plans, and integrates requests for changes and repairs to the current program (i.e., during project execution after requirements and/or design freeze) and future versions. Any member of the Command may submit a Change Request (CR) to the CCB. The CCB Chair maintains a Change Control Database to track requests. Such changes must include the following information. (a) Definition of the change. (b) Subsystems changed, (c) Unchanged subsystems that are affected, (d) Documentation to be updated, (e) Priority, (f) Dependencies,

fridaya@uef.fi

(g) Cost/savings (if any), (h) Effort, (i) Requested completion date. The CCB process is held at least once a year. If circumstances warrant, the CCB Chair can call for additional meetings.

5.6 A Flow Diagram of an Integrated I-I Education Program

The different components of the I-I education program operates as a complete system. As such, there are interactions among the components, as well as processes that must be in place for the overall program to be successful. Figure 9 is a top-level flow diagram that graphically represents the physical route or flow of objects (e.g. people, materials, vehicles, or communication) that are associated with a process in the I-I education program [13].

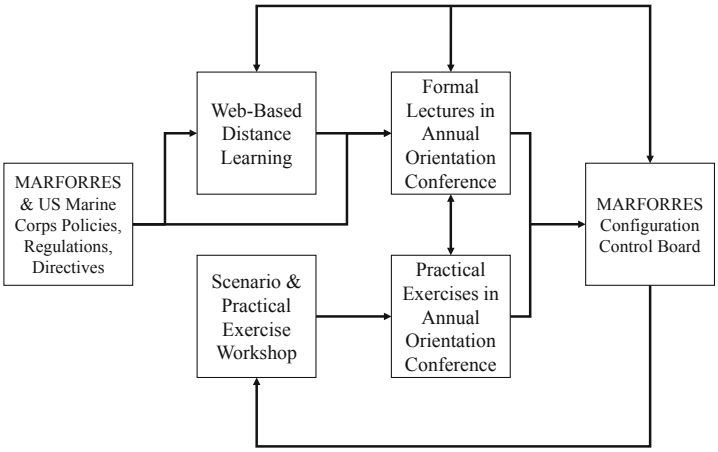


Fig. 9. A top-level flow diagram of the I-I education systems.

This flow diagram is an integrated view of the I-I education program. It offers a simplified process to build the necessary knowledge that the I-I cadre require. The process begins with a study of authoritative sources to inform the DL site and the lectures presented in the I-I Annual I-I Orientation Conference. The I-I attendees are directed to complete the web-based DL instruction prior to arriving to the Conference. Before the Conference, the NPS and MARFORRES team hold a scenario and practical exercise workshop to develop specific scenarios that the CCB identifies. During the Conference, the MARFORRES HQ staff provide formal lectures about MARFORRES and reserve processes. Interspersed in the Conference schedule are practical exercises that reinforce the formal lectures and the wed-based DL modules. After the Conference, MARFORRES hosts a CCB to discuss feedback about the Conference, formal lectures, practical exercises, and web-based DL. The CCB also initiates ideation for the next year's theme or topics for the practical exercises. Results from the CCB influence revisions in the web-based DL, as well as the formal lectures in the Conference. If there are changes in Marine Corps or MARFORRES processes after the CCB, the CCB Chair may call another session to assist in updating the program components.

fridaya@uef.fi

6 Initial Response to the New Education Program

MARFORRES has implemented this education program in various forms. A prototype was employed in 2014, which did not include a practical exercise element. Refinements in the succeeding years employed a prototype of some practical exercises, and then a formal process to develop them. In 2018, MARFORRES established a CCB, which completed the vision for the program and set the stage to employ the entire I-I education program in FY19.

The previous education program for the I-I was decidedly not meeting the expectations of the MARFORRES leaders, thus prompting a search for improvements. Surveys regarding the past program were poor. Response from the most recent Annual I-I Orientation Conference in September 2018 is the most representative. In this version of the program, all components of the education system is present, except for the CCB. This survey includes 111 data from 168 attendees, which is approximately a 70% response rate from the overall attendees. The exit survey asks a respondent, on a scale of 1 to 6, the level of agreement with a given statement (Fig. 10).

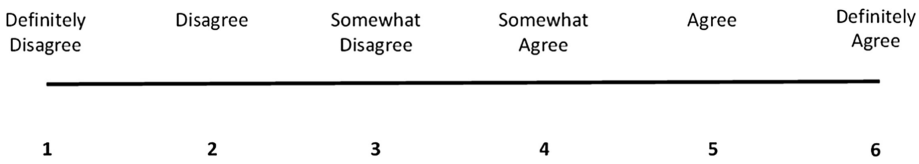


Fig. 10. Survey scale for degree of agreement with specified statement.

Three key statements upon which the team and MARFORRES leaders focused are captured in Table 2. The median response to each statement is 5, "Agree." Additionally, each formal lecture in the Conference received a median response of 5 or 6. Only one lecture received a median response of 4, "Somewhat Agree." In each of the three statements, less than 6% responded with a 3 or less. Results from the October 2018 CCB are not available at the time of this writing, but these early signs of positive feedback are encouraging.

Table 2. Responses to components of the new, integrated I-I education program.

Survey Statement	Median Score or Response
The I-I Distance Learning course was helpful in orienting me to the nature and scope of work that I am assuming as an I-I.	5 = Agree
The Practical Exercises introduced new facets or expanded my perspective about my I-I duties.	5 = Agree
In general the I-I Annual I-I Orientation Conference achieved the "What's In It For Me" concept.	5 = Agree

fridaya@uef.fi

After their maturation, two new components (Web-Based DL and Practical Exercise) to build the I-I knowledgebase are receiving good reviews. Anecdotally, I-Is volunteered ideas for new practical exercise. Other I-Is offered a recommendation to make completion of the web-based DL modules mandatory for their entire command.

7 Conclusions

This paper is a broad description of an education design that integrates formal lectures, internet technology, and scenario methods. To innovate a unique design for the challenges in the MARFORRES organization, the NPS and MARFORRES team employed systems engineering processes and tools in the design framework, which has resulted in a new education program that fuses web-based DL, formal lectures, and scenario methods expressed as practical exercises. The underpinning of each component is based on applicable theory, evidentiary support, and system engineering design concepts. MARFORRES has successfully implemented this innovative instructional design.

Initial responses to the I-I education program are strongly positive. These preliminary results point to a return on investment that supports continued study and development. The team is prepared to pursue new technologies and methods that will enhance and accelerate the development of the I-I knowledgebase. For instance, MARFORRES has voiced a desire to embed interactive scenarios or practical exercises into the online portal. Such areas of improvement will be informed by the results of the CCB, thereby providing a valid trace of requirements for the education system.

Future projects based upon the SE approach will incorporate a careful and thorough analysis of the prerequisite conditions that need to be in place in order to shape condition for program success. Activities include stakeholder engagement, academic partnerships with practitioners, and resources to implement the program over multiple iterations. The team is also reviewing the efficacy of heuristic tools to evaluate the long term and force wide impact of this approach to education. Most importantly, the team believes that this approach for developing an effective educational program is repeatable and scalable. A near term effort involves the MARFORRES personnel division (G-1). Many personnel actions in reserve units are unique. Education and training of administrative tasks for active duty Marines do not include conditions in the reserve units. As a result, incoming I-I are challenged when performing the critical task of deploying Marines to augment active duty forces. The G-1 chief has requested an instructional design for incoming administrative personnel that is similar to the I-I education program.

There are different areas throughout MARFORRES and organizations with comparable conditions that can benefit from the systems engineering model that was used to develop the I-I education program. While previous education designers have referenced system engineering, the NPS team directly implemented techniques from systems engineering to create a complete integrated design solution for the engineering problem that the MARFORRES education issues presented. The partnership that NPS and MARFORRES has forged promises new ideas and innovations in education design that other practitioners can follow in their own search for customized education programs, or as a means to adjust existing programs to meet the demands of new situations.

fridaya@uef.fi

References

1. Nissen, M., McGuiness, R., Davis, A.: Accelerating Knowledge Flow through Asynchronous Learning Technologies. Technical Report. Naval Postgraduate School, Monterey (2014)
2. McGuiness, R.: Presentation to Naval Postgraduate School Faculty and Students: Marine Corps Thesis Research Working Group (TRWG), Monterey, CA (2013)
3. House of Representatives Committee on Armed Forces.: Title 10, United States Code—Armed Forces (As Amended Through January 7, 2011). In: House of Representatives, vol. III, Washington, DC (2011)
4. Davis, A.: Summarized Exit Survey Responses from MARFORRES Annual I-I Orientation Conference, New Orleans, LA (2018)
5. Powley, N., Borrego, A.: Presentation to MARFORRES Leadership: Cultural Assessment, New Orleans, LA (2018)
6. Blanchard, B., Fabrycky, W.: Systems Engineering and Analysis, 5th edn. Pearson Prentice Hall, Upper Saddle River (2011)
7. Gaff, J.G.: Toward Faculty Renewal. Jossey-Bass Publishers, San Francisco (1975)
8. Reigeluth, C.M.: Principles of educational system design: designing and implementing new models of schooling. Salisbury, D.F. (ed.) Int. J. Educ. Resour. **19**, 99–196 (1993)
9. Reigeluth, C.M.: Educational system design. In: Kovalchik, A., Dawson, K. (eds.) Education and Technology: An Encyclopedia. ABC-CLIO, Santa Barbara (2004)
10. Dick, W., Carey, L.: The Systematic Design of Instruction, 4th edn. HarperCollins, New York (1996)
11. Sehanovic, J., Krelja-Kurelovic, E.: Interactive Web-Based Learning. Education Technology. https://www.scribd.com/document/388654399/w49. Accessed 11 Oct 2018
12. Lister, M.C.: Trends in the design of e-learning and online learning. MERLOT J.Online Learn. Teach. **10**(4), 671–679 (2014)
13. Buede, D.M., Miller, W.D.: The Engineering Design of Systems, Models and Methods, 3rd edn. Wiley, Hoboken (2016)
14. Cook, D.A.: Web-based learning: pros, cons, and controversies. Clin. Med. **7**(1), 37–42 (2007)
15. Nissen, M.E.: Harnessing Dynamic Knowledge Principles in the Technology-Driven World. IGI Global, Hershey (2014)
16. Dalton, J.: AR, VR and 3-D can make workers better. Ind. Syst. Eng. **50**(9), 42–47 (2018)
17. Hadjerrouit, S.: Creating web-based learning systems: an evolutionary development methodology. In: Proceedings of the 2006 Informing Science and IT Education Joint Conference, 119–144 (2006)
18. Brown, M.E.: A Grave New World: Security Challenges in the 21st Century. Georgetown University Press, Washington, DC (2003)
19. von Reibnitz, U.: Scenario Techniques. McGraw-Hill Publishing Co., New York (1988)
20. Mietzner, D., Reger, G.: Advantages and disadvantages of scenario approaches for strategic foresight. Int. J. Technol. Intell. Plann. **1**(2), 220–239 (2005)
21. Chermack, T.J.: Scenario Planning in Organizations: How to Create, Use, and Assess Scenarios. Berrett-Koehler Publishers, Inc, San Francisco (2011)
22. Bradfield, R., Wright, G., Burta, G., Cairns, G., Van Der Heijden, K.: The origins and evolution of scenario techniques in long range business planning. Futures **37**, 795–812 (2005)
23. Ramireza, R., Mukherjeeb, M., Vezzolic, S., Kramerd, A.M.: Scenario as a scholarly methodology to produce interesting research. Futures **71**, 70–87 (2015)

fridaya@uef.fi

24. Hoad, A.K., Monks, T.: A note on the use of multiple comparison scenario techniques in education and practice. In: Proceedings of the 2011 Winter Simulation Conference, 3899–3909 (2011)
25. Gharajedaghi, J.: Systems Thinking, Managing Chaos and Complexity: A Platform for Designing Business Architecture. Elsevier, Burlington (2006)
26. Perla, P.P., Curry, J.: The Art of Wargaming: A Guide for Professionals and Hobbyists. U.S. Naval Institute Press, Annapolis (2011)
27. Lindgren, M., Bandhold, H.: Scenario Planning: The Link between Future and Strategy. Palgrave MacMillan, Basingstoke (2009)
28. Maack, J.N.: Scenario analysis: a tool for task managers. In: Social Analysis, Selected Tools and Techniques, Social Development Family of the World Bank, Washington, D.C. (2001)

fridaya@uef.fi

Application of Web 2.0 Tools in the Pedagogical Process of Mathematics: A Case Study with Students of Basic Education

Diego Avila-Pesantez[1,2(✉)], Mónica Vaca-Cárdenas[2], L. Miriam Avila P.[2], and Leticia Vaca-Cárdenas[2]

[1] Unidad Educativa Pacífico Villagómez, Guano, Ecuador
davila@espoch.edu.ec
[2] Escuela Superior Politécnica de Chimborazo, Riobamba, Ecuador

Abstract. This work was focused on the study of Web 2.0 Open Source tools by using the methodology design: Selection, Design, and Implementation (SDI) for the pedagogical development of Mathematics teaching to eighth-grade students of basic education. The most suitable tools applied in the educational and pedagogical development were determined. The research was conducted in two scenarios: the first one, by only using traditional learning with teacher guidance in the classroom; and the second one, by using Web 2.0 tools as a didactic and pedagogical resource for the reinforcement of guided knowledge. As a result of this analysis, it was determined that the Moodle platform was the most suitable tool for this school environment, combining with several web 2.0 tools such as Edmodo, Edpuzzle, and exercise simulators. Through them, improvement was achieved in the assimilation of knowledge, application of experiences, and intellectual development of students.

Keywords: Information and Communication Technologies (ICT) · Selection, Design, and Implementation (SDI) Methodology · Web 2.0 tools · Pedagogical process · Learning

1 Introduction

At the present time, the Internet and a large number of available tools for Information and Communication Technologies (ICT) in the media have sparked a revolution in all areas of knowledge and have transformed the outlook of education that cannot be left behind, when facing processes that the current society requires [1]. This trend is spreading vertiginously with the arrival of the Internet of Things (IoT) [2–4]. In the XXI century, information is considered to be a power and its management is facilitated through the use of technology [5]. ICT in education provides multiple advantages, on the one hand it improves the learning approach and on the other hand it creates new forms of interaction and dissemination [6–8]. The internet and ICT include several tools used for knowledge and communication; they work with an infinity of information which is created, managed, stored and communicated [9, 10].

Today there are many Open Source Web Tools. Its appearance began with Web 1.0, which consisted of only static pages with text and images that contained information;

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1097–1116, 2019.
https://doi.org/10.1007/978-3-030-22868-2_73

fridaya@uef.fi

however, it did not allow interaction. The Web became popular and improved, giving rise to Web 2.0 or Social Web; a technology with server-side applications, which are programs that host web pages and allow the user to interact with them [11, 12]. XHTML, CSS, and Javascript standards appear in Web 2.0. In addition, these technologies allow users to participate actively and collaboratively. Subsequently, the Web 3.0 known as the Semantic Web [13] appears, which is characterized by the use of artificial intelligence, it uses large databases on accessible content that can be offered on the Web through multiple search engines and the well-known geospatial web. The geospatial web refers to a three-dimensional design, where a user can see the real image of a building and the location of a place or space. All these types of technologies are being used in games, virtual reality (VR) models [14–16], and multimedia solutions [17]. Even though, the trend is towards Web 3.0 and Web 4.0, important tools in Web 2.0 exist to enhance pedagogical processes [18].

The pedagogical processes are the set of situations that are designed and organized by teachers in a systematic and logical way, through which, the learning proposed in the didactic unit, class, or learning session will be developed [19]. There are barriers that prevent the successful implementation of ICT and Web 2.0 tools in the pedagogical process of Basic General Education (BGE) areas in Ecuador. The main limitation for using these resources and technological materials lies with the teachers, because many of them are not updated and do not know how they work. Both teachers and students of the new era face great challenges in the educational area. The "Digital Natives" [20, 21] have a natural practice with the use of technology and communication tools; nevertheless, they demonstrate deficiencies in the execution of tasks that lead to improving their academic performance [22, 23].

It should be noted that the use of ICT and Web 2.0 tools in the teaching and learning process of any area, by themselves, will not improve 100% student knowledge and development. To achieve the tool support objective of improving the understanding of contents, firstly, it is necessary to outline the objective or the skill that is wanted to be attained with the implementation of Web 2.0 tools [11, 24]. In addition, it is necessary that an alignment and relationship between the educational objective and Web 2.0 tools designed for the content exist, that is why the Selection-Design-Implementation (SDI) Methodology is designed [18]. This methodology helps teachers in any area of BGE to choose and apply the resource that best suits a particular process, as support material and academic reinforcement for their students, because it raises criteria for both the teacher, who will make a choice; as well as, technical criteria for the Web 2.0 tool that will be selected [11, 24].

Emphasizing the study of Mathematics as a basic subject of BGE, the methodology of Grangel et al. [18]. has been applied. In order for the students' learning to be meaningful and lasting, it is essential to change the traditional schemes on which the teacher guides their class for interactive pedagogical processes, through the use of technological means that are centered on Web 2.0 tools [25, 26], such as: Moodle, Edpuzzle, Google Sites, Kahoot, Wordpress, Youtube, Prezi, Google docs, Google Drive, Edmodo, Facebook, among others. These technological means, besides being motivators for the students, they also improve the assimilation of knowledge and consolidation of constructivist learnings [18, 27].

fridaya@uef.fi

Thus, the main objective of this article is to make a study of the existing Web 2.0 tools that can be used in the educational field, as well as to present the results of the application of the SDI methodology applied to the subject of Mathematics with the students of eighth year of Superior Basic Education of the Educational Unit "Pacífico Villagómez".

1.1 Importance of Web 2.0 in the Academic World

Marín et al. [28] describe Web 2.0 tools like a computer technology, which promotes the flow of information, thereby facilitating access to it in an organized, participatory and effective manner, through technological tools. On the other hand, Vaquerizo [29] refers to Web 2.0 as a collaborative web where the content is created by users, which has meant a revolution in the dissemination of information, since it implies greater participation and relationship between a technological aspect and a social aspect. Both authors agree that Web 2.0 is a collaborative platform in which the user is no longer a mere reader of web pages, but the user is a participatory, collaborative, and cooperative entity, and resource- builder. These technologies help in the development of tasks used in different areas, and one of them is education, focused on teaching and learning processes that must be permanently adapted to the characteristics of these technological tools [3].

The use of Web 2.0 tools in the educational process establishes that students interact and build their own knowledge. In addition, these tools allow teachers and students to get trained by using blogs, WebQuests, wikis, virtual course platforms, exercise simulators, evaluation platforms, among others. Teachers can organize learning strategies in the pedagogical process of different areas of knowledge. Through these technological didactic resources, a more dynamic and interactive learning can be achieved, focused on the new requirements of the educational era. In Fig. 1, the characteristics of Web 2.0 are displayed.

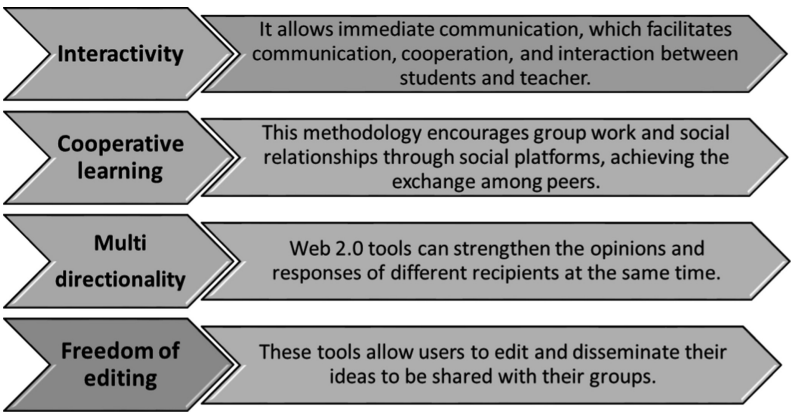


Fig. 1. Particularities of Web 2.0 tools.

Web 2.0 provides asynchronous and synchronous technological resources that constitute the main advantage of the network. These media which is based on social

fridaya@uef.fi

networks can provide feedback and place discussion forums using themes proposed by the teacher, in order to support the development of better competencies. In addition, Web 2.0 allows teachers to reinforce and encourage learning by means of virtual platforms, through the open participation of all students in certain processes within the subjects. The teacher can achieve interaction among their students, form discussion debates, share ideas through chat, send activities through tasks and lessons, and download resources such as videos, simulators, exercises, and others. These components will permit to obtain constructivist and meaningful learning. The classification of Web 2.0 tools is shown in Fig. 2 in 5 groups of agents.

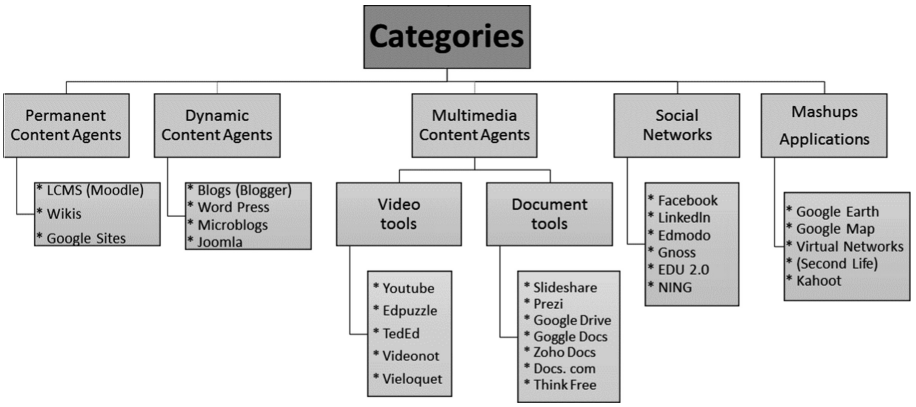


Fig. 2. Classification of Web 2.0 tools [18].

The learning session is composed of pedagogical processes that are the set of situations that each teacher designs and organizes in a logical sequence to develop outlined learnings that are raised in the didactic unit or in the learning session [19]. The link that is established between the pedagogical process and the action of the students, is participation; which is specifically distinguished by its social, cognitive and affective nature. Within this aspect, the relationship between the teacher and the student is presented, since, through the teacher's pedagogical orientations, the student is the one who organizes his/her knowledge in function of the contents of study. For that purpose, the teacher is the one who must make use of a series of didactic resources and mainly of technological tools that are available in Web 2.0 [4] (See Fig. 3).

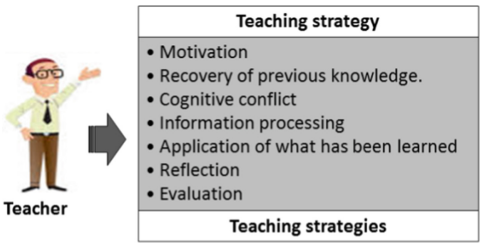


Fig. 3. Pedagogical processes in teaching

fridaya@uef.fi

In order for the pedagogical process to reach an effective concretion in a real way, the form in which the teachers guide their students is important; and the students in turn assimilate, act and participate directly in achieving the new learning. Teachers require a reflective attitude towards their teaching practices, which allows them to provide spaces for integration, participation, and active development to improve the quality of learning. To achieve learning through cognitive processes there is a range of activities linked to the condition of students, who must accomplish them to assimilate knowledge effectively. Those activities complement the teaching activities carried out by the teacher (see Fig. 4).

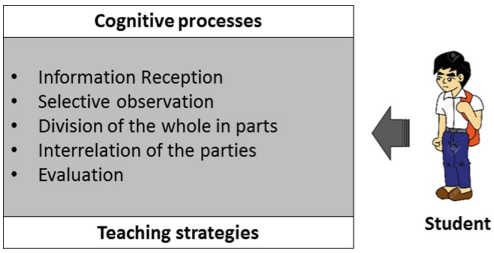


Fig. 4. Cognitive processes in the student learning.

The teacher is the fundamental factor in the learning process; so that, it is important to bear in mind certain characteristics that strengthen the teacher's pedagogical action inside and outside the classroom [5]. The main characteristics of the teacher are displayed in Fig. 5.

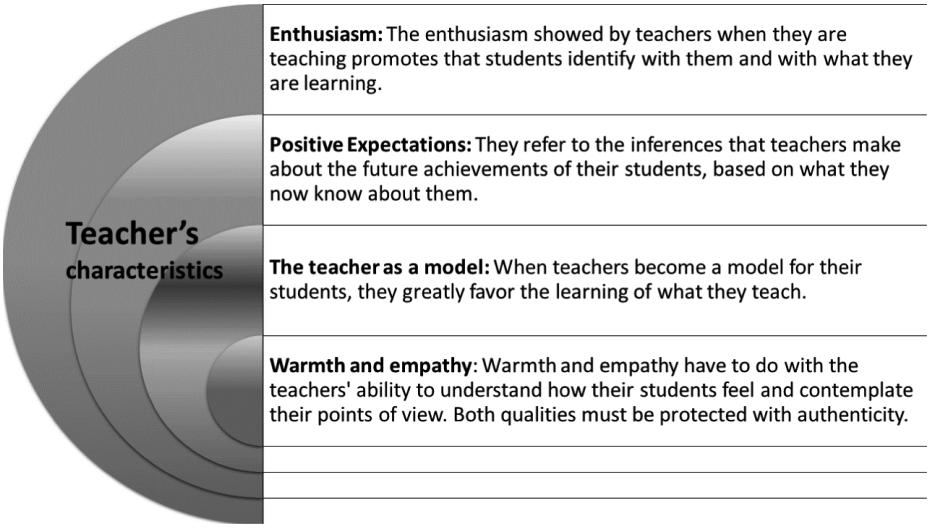


Fig. 5. Particularities of the teacher.

fridaya@uef.fi

2 SDI Methodology to Select Web 2.0 Tools

There is a wide range of technological tools available on the Internet; therefore, it is necessary to choose a methodology to select the most suitable web resources for each teacher in the basic areas. With this background, the SDI Methodology was chosen. This is based on three phases: (a) Establishment of the ideal web tool through the selection criteria; (b) Designing of the teacher's experience, in which the didactic guide is presented incorporating the tool in the pedagogical process; and (c) Implementation of the tool in the educational environment, where the resource design is executed and its functionality is assessed in the pedagogical process of the respective subject (see Fig. 6). Next, the phases are detailed.

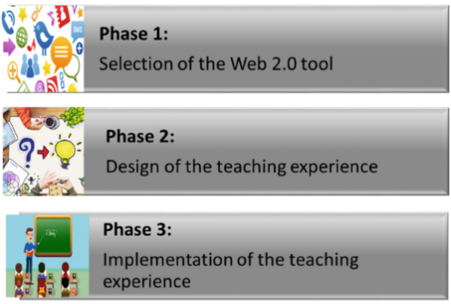


Fig. 6. Adaptation of the SDI Methodology phases [18].

2.1 Phase 1: Selection of the Web 2.0 Tool

The selection of the Web 2.0 tool is based on three steps:

(a) To define the teaching objectives: These should be raised from the main objective of the subject, so students achieve a better learning, by using Web 2.0 technologies.
(b) To define the characteristics of the subject: The characteristics of the contents must be established, theoretical or practical; the number of students; and the profile of the students to whom this subject is dedicated.
(c) To define the technical requirements: These requirements must be set out taking into account the established objectives and the characteristics of each subject, such as: easiness of use (usability), portability, file management, multimedia support, access control or accessibility, interactivity, type of license, collaboration, navigability, flexibility.

2.2 Phase 2: Design of the Teaching Experience

In this phase, the teacher of the subject develops the activities in which the 2.0 technology will be used. That is, the teacher must prepare the teaching guide that contains the objectives, competences, activities, resources, evaluation and specific bibliographic

fridaya@uef.fi

sources to achieve a better integration of students with the contents. Once this guide is designed, the environment is created through the selected tool, adapting it to the activity proposed in the guide. This is elaborated taking into account the characteristics of the students, the subject, the security and the access that the students must have for its later use. In addition, the teacher must introduce the contents in the environment, validate its operation through the configuration of user roles, and especially seek information about help and conditions for the use of the selected Web 2.0 tool.

2.3 Phase 3: Implementation of the Experience

In this stage the execution of the teaching experience and student evaluation on their performance in the class is carried out, taking into account the organization, cooperation and the individual or collaborative work of the student using the Web 2.0 tools in the teaching and learning process.

3 Implementation of the Web 2.0 Tool in the Pedagogical Mathematical Process

The proposal outlines the analysis of Web 2.0 technologies that are adapted to the teaching objectives and to the students of eighth grade of basic education in the Mathematics subject. Students have a series of different skills and interests, so it is necessary to change the way teachers think and especially the way they teach and learn. Once, the phases of the SDI methodology are detailed, each one of them will be implemented in a practical way.

Phase 1: The selection of the Web 2.0 tool in the pedagogical process of mathematics contemplates the following activities:

(a) For the definition of the objectives in the pedagogical process, Web 2.0 tools have been selected for the topics of Statistics and Probability, taking into account four established objectives: learning by doing, learning by interacting, learning by searching and learning by sharing (see Fig. 7).
(b) For the definition of the characteristics of the subject in the pedagogical process of mathematics, the following parameters were taken into account (see Fig. 8).
(c) For the evaluation of the appropriate tool, the Linker scale of 1–5 was used with each one of the criteria or technical requirements outlined in phase 1 of the proposed methodology. This score assessment was based on the work carried out by [29–39] (see Table 1).

Eight technical criteria have been taken into account to evaluate each one of the tools, such as Usability, Interactivity, Functionality, Flexibility, Multimedia Support, Access Control, License Type, and File Management. Then, they were classified by category, discarding the tools after considering the easiness of use and dynamism for the student. The results of the selection of Web 2.0 tools are displayed in Table 2.

fridaya@uef.fi

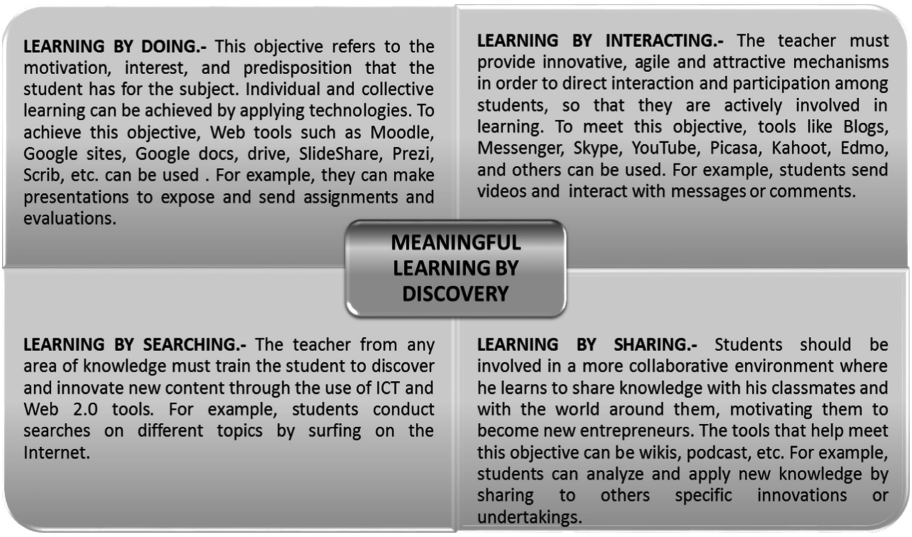


Fig. 7. Objective approach in the pedagogical process.

Teacher's field	Characteristics of the environment	Characteristics of the subject	Characteristics of the group
The profile of the teacher who guides the pedagogical process must be taken into account. It refers to the teachers characterization (engineer, graduate, doctor, technologist, lawyer, etc.)	Before applying the Web tool, one must identify the context in which the students develop, see if the subject taught by the teacher is a seminar, a workshop, etc.	At this point, it will be defined in the subject has a practical or theoretical focus, if it is a compulsory, optional or complementary subject, within their study curriculum.	The number of students and the type of groupings to be carried out in order to work cooperatively with the students must be taken into account.

Fig. 8. Definition of characteristics of the Mathematics subject.

For each evaluation criterion the grade obtained with the Likert scale was averaged. As a result, it was determined that Moodle is the platform that best adapts to the teacher's objectives due to its characteristics with a greater degree of usability, accessibility and easy navigability for the student and the teacher. In addition, it allows the incorporation of many Web 2.0 resources such as: videos, images, documents, exercise simulators, evaluations, task completion, and collaborative communication.

Phase 2: The Teaching Guide design in the area of mathematics is detailed in Table 3.

fridaya@uef.fi

Phase 3: Web 2.0 tools were used in the pedagogical process to measure the achievements of students and improve motivation and cooperation of autonomous learning.

Table 1. Technical selection criteria for Web 2.0 tools

Indicators to be evaluated	Tools	Relevance degree					Total
		Very high	High	Medium	Low	Very low	
Usability Accessibility, navigability, look & feel (design and image)	Moodle	5					5
	Google sites		4				4
	Edpuzzle		4				4
	Google Driver		4				4
	Edmodo		4				4
	Kahoot		4				4
Interactivity asynchronous/synchronous tool Cooperation and collaboration tool.	Moodle	5					5
	Google sites			3			3
	Edpuzzle	5					5
	Google Driver	5					5
	Edmodo	5					5
	Kahoot	5					5
Functionality Efficiency and effectiveness Portability	Moodle	5					5
	Google sites		4				4
	Edpuzzle	5					5
	Google Driver	5					5
	Edmodo	5					5
	Kahoot	5					5
Flexibility Technological or technical adaptability Pedagogical and didactic adaptability	Moodle	5					5
	Google sites	5					5
	Edpuzzle			3			3
	Google Driver	5					5
	Edmodo	5					5
	Kahoot	5					5
Multimedia Support Integration of images, photos, videos, sounds, animations, etc.	Moodle	5					5
	Google sites	5					5
	Edpuzzle	5					5
	Google Driver	5					5
	Edmodo	5					5
	Kahoot	5					5

(*continued*)

fridaya@uef.fi

Table 1. (*continued*)

Indicators to be evaluated	Tools	Relevance degree					Total
		Very high	High	Medium	Low	Very low	
Access Control It requires self-identification of security for entry	Moodle	5					5
	Google sites					1	1
	Edpuzzle	5					5
	Google Driver	5					5
	Edmodo	5					5
	Kahoot	5					5
License Type It is commercial or free	Moodle	5					5
	Google sites	5					5
	Edpuzzle	5					5
	Google Driver	5					5
	Edmodo	5					5
	Kahoot	5					5
File Management Access to documents and review of works and tasks.	Moodle	5					5
	Google sites		4				4
	Edpuzzle	5					5
	Google Driver	5					5
	Edmodo	5					5
	Kahoot	5					5

Table 2. Summary of Web 2.0 tools selection for its implementation in the mathematical course

Tool	Criteria to be evaluated								Point total	Average
	Usability	Interactivity	Functionality	Flexibility	Multimedia Support	Access Control	License Type	File Management		
Moodle	5	5	5	5	5	5	5	5	40	**5,00**
Google sites	4	3	4	5	5	1	5	4	31	**3,87**
Edpuzzle	4	5	4	3	5	5	5	5	36	**4,50**
Google Driver	4	5	5	5	5	5	5	5	39	**4,87**
Edmodo	4	5	5	5	5	5	5	5	39	**4,87**
Kahoot	4	5	5	5	5	5	5	5	39	**4,87**

fridaya@uef.fi

Table 3. Components of the mathematics teaching guide

Subject Data			
Subject Name	Mathematics		
Subject	Mathematics using Web 2.0 tools		
Block	No. 6. Statistics and Probability Relations and functions		
Curriculum	2018-2019		
Name of the Educational Unit	Pacífico Villagómez		
Quarter	II	**Class**	8th "A"
Hours per week	6 h per week		
Theoretical hours	15	**Classroom**	8th "A"
Practical hours	15	**Place**	Computer center of the educational unit

Teacher's Data	
Teacher's name	Luis Zambrano
Area of knowledge	Education sciences
Phone	0985323048
Email	zambranosilva5@hotmail.com
Tutorial schedule	13H30 a 15H30 teacher's room

Subject Description

2.1. Presentation

Mathematics subject using the Web 2.0 tools aims to provide 8th year students of BGE with the necessary knowledge about the creation of data tables and statistical graphs for the development of their professional activity. Since this is one of the most important subjects of the curriculum map and in which students present more problems; it is important to find the best way to make mathematics easy and motivating. Mathematics is related to geometry and measurement and statistics in each one of the proposed blocks, that is why mathematics is in every moment of life.

Competencies or skills with performance criteria

2.2. Subject Competencies

- Basic knowledge about the use of Web 2.0 tools in some mathematics topics.
- To recognize ordered pairs with integers and place them in the Cartesian plane.
- To interpret and construct data tables and graphs related to different areas of daily life.
- To collect, analyze, organize and represent statistical data related to different areas of daily life.
- To extract representative information from a collective based on statistical parameters.
- To calculate and contrast absolute and accumulated frequencies of a series of graphical data.
- To recognize the importance of collective work in the execution of tasks and studies.

Contents

2.3. Contents according to the curriculum (Theoretical contents)

1. Data tables and Cartesian graphics
1.1. Data Tables 1.2. Cartesian coordinates
1.3. Cartesian graphics

(*continued*)

fridaya@uef.fi

Table 3. (*continued*)

Contents			
2.3. Contents according to the curriculum (Theoretical contents)			
2. Statistical studies 2.1. Statistical variables. Frequencies 2.2. Absolute and relative frequency **3. Statistical tables and graphs** 3.1. Statistical Tables 3.2. Statistical graphs 3.3. Description of randomized experiments			
Teaching methodology			
Formative activities			
Activities	Teacher's work	Student's work	Graded points
Theoretical class	Exposition about the topic of study using the learning cycle method. To help students in doubts they might have through the realization of questions about the explained topic.	To attend punctually to each one of the classes. To have their workbook about the proposed topics in case of doubts. To study and reinforce the knowledge guided by the teacher.	10
Resolution of theoretical exercises	To set out several exercises of each one of the topics. Once the exercises have been explained, the teacher proceeds to bring up new tasks for students to develop and ask if there have any doubt.	Active and voluntary participation to develop the exercises.	10
Partial and quarter evaluations	To make evaluation rubrics either in paper or with the help of one of the Web 2.0 tools	Revise for the partial evaluation or the quarter evaluation, according to the case	10
Practices using Web 2.0 tools	To elaborate the topic using Web 2.0 tools according to the theme, the number of students, and the objective to be achieved with 8th-year students.	To use the web tool designed by the teacher as a pedagogical means to reinforce the theoretical contents. To help their classmates and interact with them by using web tools.	10
			10

fridaya@uef.fi

Evaluation

Instrument	Criteria to be evaluated	Weight (points)	Competences or skills
Test elaborated by using a Web 2.0 tool, proposed for the topic.	Answer the form of questions related to the creation of data tables.	10 points	To interpret and construct data tables and graphs related to different areas of daily life
	To see the videos planned in the Edpuzzle platform for the 8th A group of mathematics which includes a questionnaire about the topic (See Fig. 9).		To recognize ordered pairs with integers and place them in the Cartesian plane.
Laboratory practices	To do all the activities about the topic of tables and the cartesian plan together under the supervision of the guide or teacher of the subject. These activities are included in each one of the themes and must be done in cooperation with a classmate (See Fig. 10).	10 points	To develop critical and autonomous skills in students. Group work and team collaboration to develop activities. Commitment to working competitively on the proposed themes.
Formative evaluation	To perform the test type quiz located in the site: www.Aprendiendomate8medivierto.com, which consists of 10 questions about the topic of the creation of data tables, location, and identification of the points in the Cartesian plane. To use the Kahoot tool as a resource for the reinforcement of contents studied in Statistics and probability (See Fig. 11).	10 points	Individual development about their skills and competencies using Web 2.0 tools.

fridaya@uef.fi

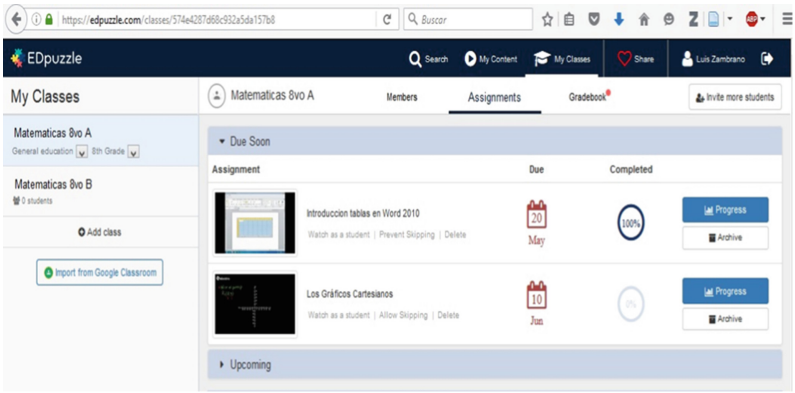


Fig. 9. Using Edpuzzle for students' evaluation.

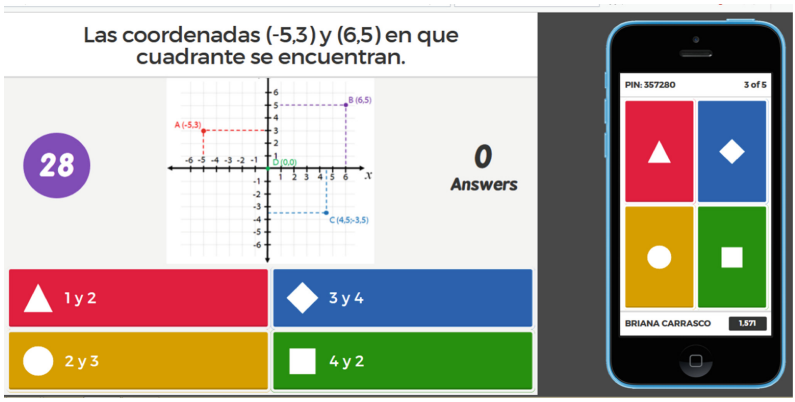


Fig. 10. Activities about the cartesian plan

Fig. 11. Students work on the Web 2.0 tool in the classroom.

fridaya@uef.fi

4 Results and Discussion

The present work implemented the SDI methodology to determine the best Web 2.0 tool applied in the learning process of Mathematics in the eighth year of Basic Education in the Educative Unit Pacífico Villagómez. The participants in this research were 30 students (W = 45.55%, M = 54.45%) with ages between 11 and 12 years old, who completed a survey about the problems and difficulties of the use and application of Web 2.0 tools in the development of school activities. This methodology was focused on an instrument based on analysis, recognition of shortcomings and proposal of improvements; in order to optimize the management and virtual communication within the institution of middle school level. In addition, the competencies of teachers were strengthened, tending to satisfy training needs based on educational innovations that provide coverage, flexibility, and appropriateness; in order to achieve personal development of students by using technological tools.

This is a quantitative research based on a contextualized and relational approach, characterized by using practical techniques and strategies to direct and guide students in the proper use of Web 2.0 tools in the Educational Unit. For experimentation, a survey designed by the authors was applied. It was structured with 12 questions in which each participant could choose only one of them, and all had the purpose of identifying the level of applicability of the web tools within the academic activity. With the obtained data, a correlational analysis was performed based on the collected results. Consequently, the same survey was applied twice in the pre-test and post-test process (see Table 4).

Table 4. Comparative table of scenarios; before and after the application of Web 2.0 tools

No	Questions	Before			After		
		Always	Sometimes	Never	Always	Sometimes	Never
1	Have you used any Web 2.0 tool in the presentation of tasks?	5	4	21	26	3	1
2	Have you used Web 2.0 tools such as: Facebook, Youtube, Google Drive, Blogger, Moodle in the learning of mathematics?	3	1	26	23	4	3
3	When you connect to the Internet, do you maintain an interactive, dynamic, and cooperative process with your classmates in educational aspects?	4	6	2.0	24	3	3
4	Have you received training from teachers for the proper use of Web 2.0 tools in the educational process?	1	1	28	28	1	1
5	Have the teachers of your institution used Web 2.0 tools in any topic of the study of mathematics?	2	4	24	24	1	5
6	Do the educational orientations focus on the use of Web 2.0 tools to reach new knowledge in some basic areas and expand your intelligence?	6	2	22	27	2	1

(*continued*)

fridaya@uef.fi

Table 4. (*continued*)

No	Questions	Before			After		
		Always	Sometimes	Never	Always	Sometimes	Never
7	Do you organize your tasks through the use of technological tools?	1	3	26	26	3	1
8	Do you consider that Web 2.0 tools permit to improve the learning of mathematics?	4	3	23	21	3	6
9	Do teachers who work in the eighth year of basic education perform evaluation processes by using a Web 2.0 tool?	0	1	29	2.0	8	2
10	Have you used Web 2.0 tools for the development of three-dimensional models, geometric figures and more contents of geometry and statistics?	1	2	27	23	4	3
11	Do you put into practice your initiatives and creativity to interact through Web 2.0 tools as a mechanism to develop learning?	3	7	2.0	25	4	1
12	Do teachers use applications and Web 2.0 tools to achieve meaningful learning in some subjects?	2	2	26	22	7	1
	Total	**32**	**36**	**292**	**289**	**43**	**28**
	Percentage	**9%**	**10%**	**81%**	**80%**	**12%**	**8%**

A summary of the results of the survey applied to the 30 students is shown next. The Cronbach's alpha was employed to determine the reliability coefficient, obtaining the value of 0.97 (see Table 4, pre-test section). With the total collected data, the deficiency of the use of Web 2.0 tools can be confirmed, limiting the development of the students' skills and abilities in Mathematics subject. Additionally, it was determined that the majority of students do not use these tools in a meaningful way, due to ignorance and lack of use and application.

After using the Web 2.0 tools selected and implemented in the classroom, the relationship with the applied survey could be evidenced. It is concluded that 80% of responses were "ALWAYS" in the different questions of the questionnaire, 12% answered "SOMETIMES" and 8% "NEVER" Those results show that using these tools has strengthened the learning processes of students (see Table 4, post-test section). Furthermore, the improvement of the pedagogical process of Mathematics was evidenced, by using the chi-square statistic, with a confidence level of 95% and with 2 degrees of freedom; obtaining a value of 424.18. These results show that the study and use of Web 2.0 tools significantly influenced the pedagogical process of the students in this case study. As a result, it is essential to continue executing motivating strategies and activities so that students develop digital skills in different spaces by making use of technological means to create a more collaborative learning through the interaction and development of teachers and students. Figure 12 shows a synthesis of the obtained results.

fridaya@uef.fi

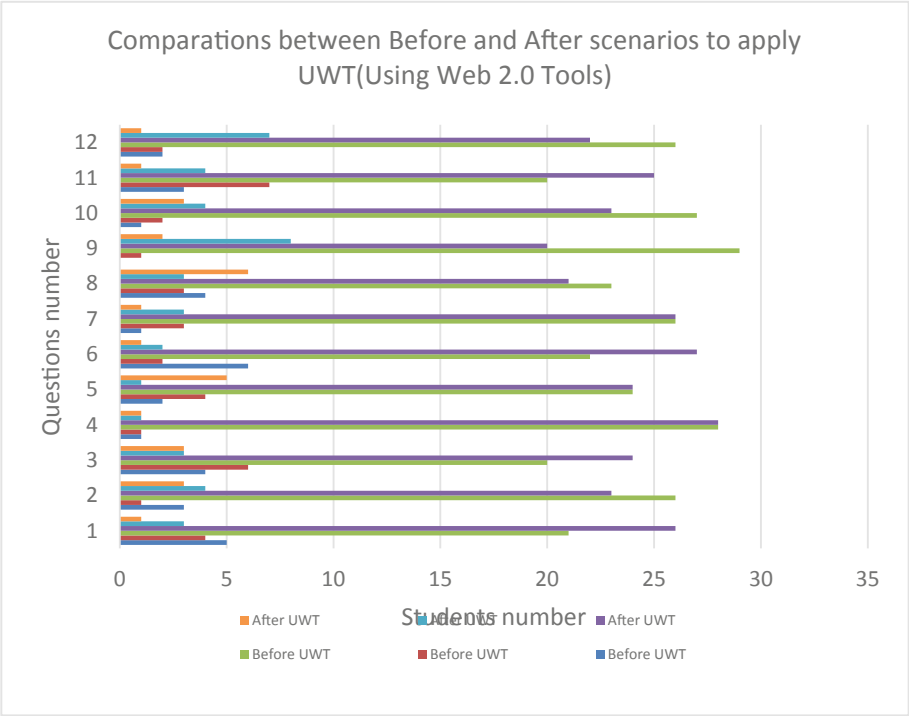


Fig. 12. Result comparison of surveys conducted to students

5 Conclusions

The use of Web 2.0 technologies applied practically and methodologically in the teaching and learning process of eighth-year students of Basic General Education has given positive results because they allowed a theoretical-practical approach to the contents of Statistics and probability (Mathematics of eighth year of basic education). These tools improved the attitude and skills of the students in the way they learn. According to the students' experiences at the time of implementing the methodology, there were Web 2.0 tools that motivated them significantly, such as Edmodo, Edpuzzle, and the exercise simulators. Those Web 2.0 tools got the students' attention, facilitating the learning of contents and achieving the goal of learning by doing, as the constructivism pedagogical current promotes it.

In addition, the design of the SDI methodology allowed teachers to select and adapt a tool according to the specific topic of the unit to be studied and their group of students, in order to achieve a better cooperation in the teaching and learning process. Therefore, it is necessary that teachers update and innovate their pedagogical process through the appropriate use of Web 2.0 tools as a didactic means of support of the learning process of Mathematics. For that purpose, initiative, creativity and mainly the application of new methodological strategies must be put into play; so that, students develop practically and interactively a meaningful learning. All of that, through the use

fridaya@uef.fi

of videos, forums, discussions, analysis, and reflections, as well as the realization of individual and group work in order to strengthen the learning process.

Finally, it is important that teachers structure evaluation activities by using Web 2.0 tools, to identify the students' level of learning and can take the necessary corrective actions, providing pedagogical feedback in the classroom about the contents that require it.

References

1. Hirshman, E.: Technology: reshaping economy and educational goals. Educ. Rev. **2**(January/February), 8–9 (2016)
2. Atzori, L., Iera, A., Morabito, G.: The internet of things: a survey. Comput. Networks **54**(15), 2787–2805 (2010)
3. Uskov, V., Pandey, A., Bakken, J.P., Margapuri, V.S.: Smart engineering education: the ontology of Internet-of-Things applications. In: 2016 IEEE Global Engineering Education Conference (EDUCON), pp. 476–481 (2016)
4. Yang, Y., Yu, K.: Construction of distance education classroom in architecture specialty based on internet of things technology. Int. J. Emerg. Technol. Learn. **11**(5), 56–61 (2016)
5. Adu, E.O., Galloway, G.: Information and Communication Technologies (ICT) and teacher education preparation in South Africa : implications for 21st century classroom-based Practice. J. Commun. **6**(2), 242–247 (2015)
6. Forcheri P., Molfino, M.T.: ICT as a tool for learning to learn (2000)
7. Bai, Y., Mo, D., Zhang, L., Boswell, M., Rozelle, S.: The impact of integrating ICT with teaching: evidence from a randomized controlled trial in rural Schools in China. Comput. Educ. **96**, 1–14 (2016)
8. Vaca-Cárdenas, L., McHaney, R.: Business teachers' ICT skills and uses in an IoT classroom. J. Educ. Learn. **11**(4), 394–403 (2017)
9. Hubackova, S., Klimova, B.F.: Integration of ICT in lifelong education. Procedia Soc. Behav. Sci. **116**, 3593–3597 (2014)
10. Vajargah, N., Jhani, K., Azadmanesh, S.: Application of ICTs in teaching and learning at university level: the case of Shahid Beheshti University. Turkish Online J. Educ. Technol. **9**(2), 33–39 (2010)
11. Rojas, Y.: La Importancia de las Herramientas Web 2.0 en la Educación. prezi.com, May 2011.
12. McHaney, R.: The New Digital Shoreline: How Web 2.0 and Millennials are Revolutionizing Higher Education. Stylus Publishing, Herndon (2011)
13. Liz, R.E.: Blended-Learning Afectivo Y Las Herramientas Interactivas De La Web 3.0 , vol. I (2014)
14. Grimon, F., Meza, J., Vaca-Cardenas, M., Monguet, J.M.: Research and trends in the studies of collective intelligence from 2012 to 2015. Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng. LNICST **180**(1), 181–187 (2017)
15. Avila-Pesantez, D., Vaca-Cardenas, L., Rivera, L.A., Zuniga, L., Miriam Avila, L.: ATHYNOS: Helping Children with Dyspraxia Through an Augmented Reality Serious Game. In: 2018 International Conference on eDemocracy & eGovernment (ICEDEG), pp. 286–290 (2018)

fridaya@uef.fi

16. Vaca Cárdenas, L., Olmedo Vizueta, D., Tavernise, A., Gabriele, L., Bertacchini, F., Pantano, P., Bilotta, E.: Darwin has come back to the Galápagos islands: an educational journey to discover biological evolution. In: EDULEARN14 Proceedings, pp. 6088–6095 (2014)
17. Cabrera, S.: Herramientas web.
18. Grangel, R., Campos, C., Rebollo, C., Remolar, I., Palomero, S.: Metodología para seleccionar tecnologías Web 2.0 para la docencia (2012)
19. R. Calderón, ¿Qué son los procesos pedagógicos ?|Intel Engage. ¿Qué son los procesos pedagógicos? October 2012.
20. Prensky, M.: Digital immigrants, digital natives. On the Horizon **9**(5), 1–6 (2001)
21. Rikhye, R., Cook, S., Berge, Z.L.: Digital natives vs. digital immigrants: myth or reality. Int. J. Inst. Technol. Distance Learn. **6**(2), 3–10 (2009)
22. Palfrey, J., Gasser, U.: Born Digital. Basic Book, New York (2008)
23. Vaca-Cárdenas, L., Bertacchini, F., Tavernise, A., Gabriele, L., Valenti, A., Olmedo, D., Pantano, P., Bilotta, E.: Coding with Scratch: the design of an educational setting for elementary pre-service teachers. In: 2015 International Conference on Interactive Collaborative Learning (ICL), pp. 1171–1177 (2015)
24. Vossen, G., Hagemann, S.: Unleashing Web 2.0: From Concepts to Creativity. Elsevier, Boston (2010)
25. Hernández Arteaga, S.M., Alvarado Pérez, I., Luna, J.C.: Creatividad e innovación: competencias genéricas o transversales en la formación profesional. Rev. Virtual Univ. Católica del Norte **44**, 135–151 (2015)
26. Goeman, K., Elen, J., Pynoo, B., van Braak, J.: Time for action! ICT integration in formal education: key findings from a region-wide follow-up monitor. TechTrends **59**(5), 1–32 (2015)
27. Belloch, C., Abad Gómez, R., Fuentes Esparrel, J.A., Liz, R.E., Grangel, R., Campos, C., Rebollo, C., Remolar, I., Palomero, S.: La web 2.0 como herramienta didáctica de apoyo en el proceso de enseñanza aprendizaje: aplicación del blog en los estudios de Bellas Artes (2012)
28. Watson, M.S., Hewson, S.M.: Securing protection standards for Canada's marine protected areas. Mar. Policy **95**, 117–122 (2018)
29. Vaquerizo García, A.E., Renedo Mena, M.B., García, V.: Aprendizaje colaborativo en grupo: Herramientas Web 2.0. In: Jornadas de Enseñanza Universitaria de la Informática (2009)
30. Marín, V.I., Zawacki-Richter, O., Garcías, A.P., Salinas, J.: Tendencias en el ámbito de la Tecnología Educativa en Iberoamérica: 20 años de la revista Edutec. Edutec. Rev. Electrónica Tecnol. Educ., 357 (2017)
31. Gómez, E.H., Vidal, F.M.: Estrategias de aprendizaje basadas en entornos virtuales en educación secundaria. Rev. científica electrónica Educ. y Comun. en la Soc. del Conoc. **15** (2015)
32. Yanez, P.: El proceso de aprendizaje: fases y elementos fundamentales. Revista San Gregorio, 70–81 (2016)
33. Cela, K., Fuentes, W., Alonso, C., Sánchez, F.: Evaluación de herramientas web 2.0, estilos de aprendizaje y su aplicación en el ámbito educativo. J. Learn. Styles **3** (2010)
34. Cuadrado, V., Carballo, J.: Análisis de las aportaciones de los blogs educativos al logro de la competencia digital. Rev. Investig. en Educ., 83–90 (2010)
35. Fumero, A., Roca, G., Sáez Vacas, F.: Web 2.0, 12th edn. Fundación Orange, Madrid (2007)
36. Guerra, J.G.: Web 2.0 presente y futuro de las aplicaciones. Perspect. s **4**(4), 60–64 (2017)
37. Hung, E.S., Calderón, C.A.: Los cibermedios en América Latina y la Web 2.0. Comun. Rev. científica Iberoam. Comun. y Educ., 125–131 (2011)

fridaya@uef.fi

38. Rodríguez-Martínez, R., Codina, L., Pedraza-Jiménez, R.: Indicadores para la evaluación de la calidad en cibermedios: análisis de la interacción y de la adopción de la Web 2.0. Rev. española Doc. científica **35**, 61–93 (2012)
39. Tamayo, C.F.M.: Plataformas virtuales como recurso para la enseñanza en la universidad: análisis, evaluación y propuesta de integración de moodle con herramientas de la web 2.0. Universidad Complutense de Madrid (2012)

fridaya@uef.fi

Designing an Educational Multimedia System for Supporting Learning Difficulties in Arabic

Moutaz Saleh(✉) and Jihad M. Alja'am

Computer Science and Engineering Department, Qatar University,
22713 Doha, Qatar
{moutaz.saleh, jaam}@qu.edu.qa

Abstract. Nowadays, schools are highly encouraged to use information technology in unconventional manners and also be creative in combining the use of various tools, while maintaining high standards. In spite of the latent advantage for the usage of assistive technology for children with learning difficulties (LD), many of the software applications, tools, and websites that those children deal with are designed without consideration of their required needs, making these software less effective or completely inaccessible for them. This paper presents an educational multimedia system to help children overcome their Arabic learning problems. The research work aims at proposing a practical educational strategy for automatically converting Arabic language text into an appropriate personalized multimedia content that would significantly enhance understanding, communication, and thinking skills for children with LD in elementary schools. The system will be also compared and evaluated with other existing educational multimedia systems.

Keywords: Educational system · Multimedia content · Assistive technology · Knowledge extraction · Text mining · Arabic language

1 Introduction

There has been a remarkable increase for children with learning difficulties (LD) around the globe [1, 2]. Such children need thorough attention especially during the early education years. To promote dynamism in education, public schools are highly encouraged to use information technology and tools in unconventional manners and be creative in combining the use of various tools, while maintaining high standards. However, the current methods of teaching in many public elementary schools in the Arab world are still legacy, somehow obsolete and need proper improvements. One major cause for such problem is the lack of well-designed resources that would effectively help teachers to present focused materials for students with LD in schools.

In fact, an important number of the school children are experiencing problems in a traditional classroom learning context which may interfere with their Arabic literacy development skills including reading, writing, and understating. This issue is particularly challenging for children with moderate understanding who have been integrated recently in the public elementary schools and require dedicated attention.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1117–1128, 2019.
https://doi.org/10.1007/978-3-030-22868-2_74

fridaya@uef.fi

Recently, Increasing attention has been drawn in the Human-Computer Interaction (HCI) community towards the design and development of accessible computer applications for children with LD. Due to better healthcare and education, the learning quality of for such children has been obviously improved especially that these children grow up and live in the community while surrounded by computer applications and software. This paper presents an adaptive multimedia learning system for assisting children with Arabic LD to enhance their understanding, communications, and thinking skills. The proposed system dynamically generates multimedia contents which can be effectively used by instructors to provide children with LD the opportunity to learn with contents designed to suit their intellectual needs, and maximize the education process through reducing the administrative and teaching overheads. The paper is organized as follows: Sect. 2 introduces related works, Sect. 3 describes the educational multimedia development process showing both Arabic text processing and its knowledge extraction, Sect. 4 shows the system operation, Sect. 4 compares and evaluates the proposed system with other existing educational multimedia systems, while Sect. 5 concludes the paper.

2 Related Work

Researchers in special education are aware of the benefit of using computer technology in helping children with LD [3]. The claim is that technology can help such children increase confidence and motivation through creative activities with immediate feedback and self-paced learning. There have been many efforts in developing multimedia learning systems in the literature. Cheng et al. [3] developed an Arabic e-learning system with built-in materials. The system chooses a personalized educational content for each student based on his initial knowledge and understanding. The authors in [4] proposed an edutainment system for teaching Arabic called "ArabicTutor". The demonstrated material shows an Arabic text with its synonyms in different settings with multimedia images. The work in [5] designed a flashcard system to teach stories for children in enjoyable manner. The teacher chooses a story and requests the student to organize its flashcards sequentially. Saleh et al. [6–9] proposed a computerized system that helps children with LD learn in an enjoyable manner using multimedia with fixed content. They developed several multimedia-based tutorials with puzzles and games that greatly enhance the learning skills of these children. The interactive educational multimedia tutorials were designed on the basis of Mayer's cognitive theory [10]. The system links the tutorials and exercises with the required teaching objectives of each child's preset personalized study plan. The system provides dynamicity in offering such multimedia contents when considering children's abilities and speed of interaction. Designing multimedia tutorials for children with LD is challenging. In fact, children have different means of acquiring knowledge even for those who are ranked at the same academic level. Hence, personalized study plan for each child with a specific knowledge database that contains the concepts, along with the corresponding multimedia contents to teach these concepts, were built. Eventually, the tutorials were developed with preset aims, and designed the tutorials by selecting adequate multimedia elements. Though many educational systems for children with LD have been introduced in the

fridaya@uef.fi

literature, none of the works so far could adaptively convert Arabic language text into an appropriate personalized multimedia content that would significantly enhance the learning process for such children.

A Text-to-picture system is a system that automatically converts natural language text into pictures representing the meaning of that text. The pictures can be static illustrations such as images or dynamic illustrations such as animations. For instance, the Story Picturing Engine system [11] uses is a pipeline of three processes namely: story processing and image selection, estimation of similarity, and reinforcement-based rank. The estimation of similarity between pairs of images is based on their visual and lexical features. The images are ranked based on a mutual reinforcement approach and most highly-ranked images are saved. In spite of its high correctness and performance, the story picturing engine only retrieves one image for a given story and discards many features such as the temporal or spatial relations. Contrariwise, Text-to-Picture Synthesis system [12] associates several images to each extracted main phrase and shows the story as a sequence of related images. It considers the text-to-picture conversion problem as an optimization process but still inherits the drawbacks of text-to-picture systems despite its performance.

In Word2Image system [13], various techniques including correlation analysis, semantic and visual clustering is used to produce sets of images from Flickr. To make the representative image more understandable, a large version of the original sized image could be shown when the user places the mouse over it. Although, Word2Image translate concepts into its visual counterpart with sets of high quality, precise, diverse and representative images, it inherits the problems related to input language such as restricting the input to single words only. Chat with illustration (CWI) [14] is a visual assisted instant messaging scheme which presents users visual messages associated with textual message automatically [14]. The CWI scheme consists of Image database construction, dialogue analysis, image selection and ranking, and visual dialogue summarization. Despite the accuracy and the intuitiveness of CWI, it relies on making excessive preparation of image resource. Furthermore, only few unpicturable concepts, such as verbs, adjectives, fixed phrases and interrogatives are considered and labeled to images manually which makes the chat directed more to concrete concepts only. Another drawback of this system is its restricted language processing abilities and grammatical and logical relations. WordsEye [15] is a text-to-scene system that can convert input text into representative 3D static scenes automatically using a linguistic analyzer and a scene depicter. Although WordsEye has attained a good degree of success for describing the scenes, it is still not interactive and does not exploit the user's feedbacks. Moreover, WordsEye relies on its huge offline rule-base and data repositories containing different geometric shape which lacks an automatic annotation task. CONFUCIUS [16] is a multi-modal text-to-animation conversion system that can generate animation from a single input sentence containing an action verb and synchronize it with speech [16]. It is composed of several modules with different tasks to accomplish such as Knowledgebase, Language processor, Media allocator, Animation engine, Text-to-speech engine, Story narrator, and Synchronizer. Generally, CONFUCIUS is not interactive in a sense that it does not let the user modify the generated animation. In addition, only a single sentences in each input is allowed and only a limited format of the input sentences is permitted.

fridaya@uef.fi

3 Proposed System

Our proposed system aims at improving the learning process of the Arabic language for children with LD. The system aim is to generate personalized multimedia contents dynamically through providing an automatic mechanism to query for a desired topic based on semantic content analysis and ontology. The system will statistically learn the preferred customization per student and will generate automatically customized tutorials. The dynamic tutorials would become more time efficient due to the machine learning process of the automated system. The detailed architecture of the system is shown in Fig. 1. The system will be able to perform Arabic text processing in a specific story domain, extract the main concepts and link them automatically to multimedia contents. Hence, images and short video clips will be presented to children that describe the suitable meaning of the Arabic terms exist in the text. The system utilizes deep knowledge in Arabic text processing including morphological analysis, ontology, logical rules and discretization.

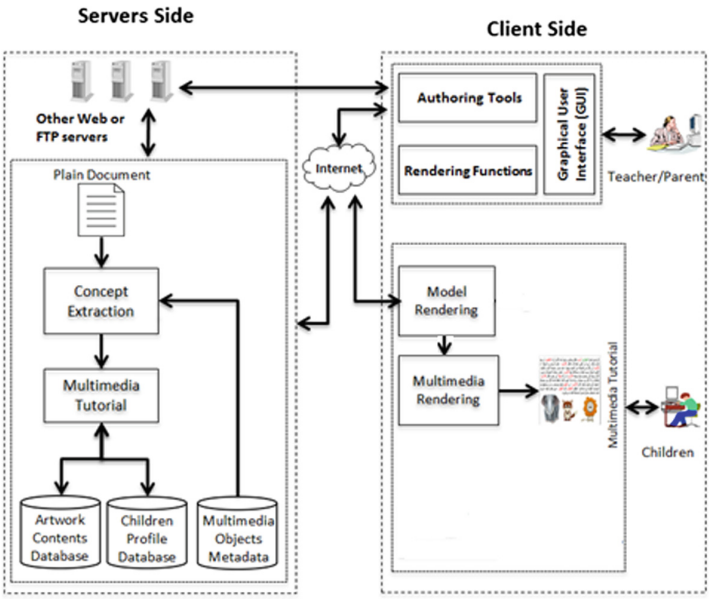


Fig. 1. Proposed system architecture

3.1 Arabic Text Processing

The automatic processing of text written in the Arabic language faces a number of challenges beyond what is expected for English. We focus here on three aspects: orthographic ambiguity, morphological richness, and syntactic freedom. First, Arabic is written with an orthography in which short vowels (and some additional phonological features) are optionally written as diacritical marks, resulting in around three different

fridaya@uef.fi

core readings per word on average. For example, the word بين byn can be diacritized as bayn 'between', bay∼ana 'he explained', bay ∼in 'clear', and biyin∼ 'in Yen'. Second, the Arabic language has with words to gender, number, people, etc. in addition to attachable clitics that include the definite article, some prepositions and possessive pronouns. For example the single Arabic word وسيكتبونها wsyktbwnhA consists of six morphemes and stands for 'and they will write it': w+s+y+ktb+wn+hA. Third, Arabic language has a freedom in verb position alternations and case endings. For instance, the verb is mainly at the start of the sentence, while nouns have case endings especially in poetic form. Therefore, we need to ensure that Arabic text processing is completely orthographically and morphologically disambiguate the text in addition to producing a syntactic parse of it. After a full disambiguation and a parse is produced, we still need to determine the appropriate linguistic features to use in the proposed system. For morphological and orthographic disambiguation, we use the MADAMIRA system [17]. As for syntactic analysis, we use the CATIB dependency parser [18]. Figure 2(a) shows the performance of MADAMIRA on the sentence الثعلب الاحمر يطارد الارنب الابيض, while Fig. 2(b) shows the parse tree related to the sentence.

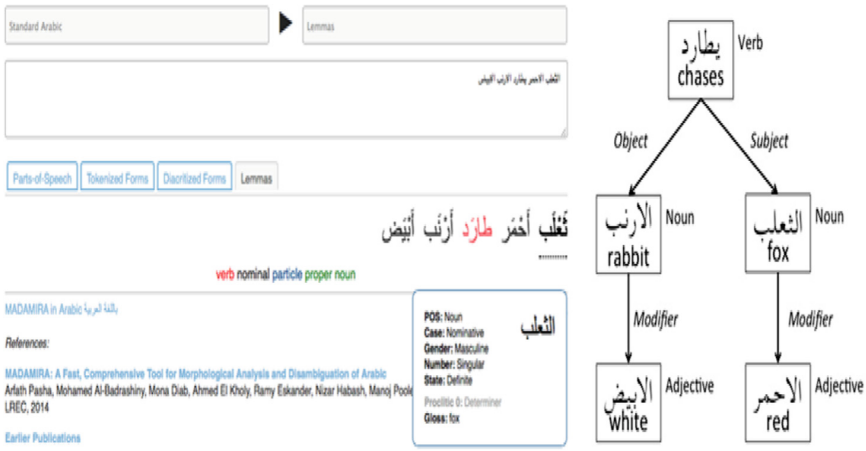


Fig. 2. Madamira analyzer (a); CATiB parser (b)

In this section, we present some of the most important linguistic features to be used:

- Diacritic: Diacritical marks are primarily used to disambiguate the words phonologically and thus morphologically and semantically. The map from diacritized text to phoneme sequences can be accomplished with a limited set of phonotization rules [19]. Since, as pointed out above, the final case-marking diacritics are not as well predicted as the lexical diacritics, we will ignore the final diacritics, which is a common practice in modern Arabic produced by native speakers.
- Tokenizations/Lemmas/Stems: Given the morphological complexity of Arabic words, there are a number of ways to tokenize them [20]. Some tokenizations can be shallow and focus on separating different clitics, while others go deeper to lemmas and stems. Lemmas are citation forms of words that abstract away from

fridaya@uef.fi

morphological variations. Stems are a particular way to separate prefixes and suffixes from the core of a word. For highly regular affixational languages, stems and lemmas may be interchangeable. But for Arabic, where more nouns have irregular plurals than sound plurals [21], stemming may not be the best option. We plan to compare the use of different tokenization schemes, lemmas and stems in the context of our work. These "token" possibilities can be used in the search for multimedia entries as well as in constructing frequency models for determining the most important terms in a story. However, we think lemmas are the appropriate form to use for ontology construction as they abstract all the inflection features, but maintain important derivational morphology features.

- Part-of-Speech: Part-of-speech information, at different levels of granularity [22] can be used in various parts of the platform to rank terms and to model the degree of complexity of the text; this is in addition to the use in syntactic parsing.
- English glosses: While English glosses are not exactly Arabic features, the fact they are produced in MADAMIRA, paired with the lemmas, make them a low hanging fruit to use for search and ontological modeling. The English glosses also allow easy linking to existing English resources.
- Syntactic subphrases: The syntactic dependency tree can be used to construct complex search terms consisting of bigrams or higher order n-grams. For example, in the sentence (shown in Arabic word order and morphology to highlight the point) 'the-fox the-red chases the-rabbit the-white', the dependency tree allows for easy construction of terms such as "fox red" and "fox chases" but not the less connected "red chases". The syntactic tree also allows us to explore the mapping of the sentence into predicate argument structure. In the example above, we know from the dependency tree that a chase event took place, the fox is the chaser and the rabbit is the chasee.

3.2 Educational Ontology Development

Generating ontologies is a very active field of research. In this work, we will use a methodology that involves designing the global structure of the ontology; building a hierarchical taxonomy, validating the ontology against a domain coverage, and finally creating a knowledge base on the selected domain. The ontology will be evaluated in its ability to provide the right elements. We discuss these outcomes with the instructors who are the domain experts. The ontology will then be refined iteratively following this process. It should be designed to fit the needs of children with learning difficulties. Consulting with special education experts and school teachers, we will evaluate the ontology by composing query questions and check the retrieved answers (i.e., competency questions). For instance, "where does the Camel live?", "which animals live with the camel", "Is the Camel a carnivore or herbivore?", "which animals live in the desert?", "What is the best food for the Camel?", "How many days can a camel live without water?", and so on. We will validate that the ontology answers contain the needed information. We will set also a mechanism to maintain the ontology and check its integrity and its expansion. Moreover, the concepts of the ontology should be well defined. What they are, their properties, and information they should provide, need to be indicated. For instance, carnivore-related term can include, "animal name" as "lion",

fridaya@uef.fi

"location" as "forest", "food" as "chair". Initially we will build a list of terms covering the domains (i.e., animals or food). These terms are extracted manually from the stories of children. In our first prototype we have annotated 15 stories. We can continue this process and annotate 50 other stories. We can also automate this process. The objective is to have a comprehensive set of concepts logical relations and properties even though some redundancy may exist. We can refine these terms at a later stage. We will then develop the class hierarchy (superclass and subclass), give the definitions of these concepts and set their properties (e.g., slots) using both top-down approach and bottom-up approach. In this work, we designed an ontology of animal classes, depicted in Fig. 3, being widely used in children stories.

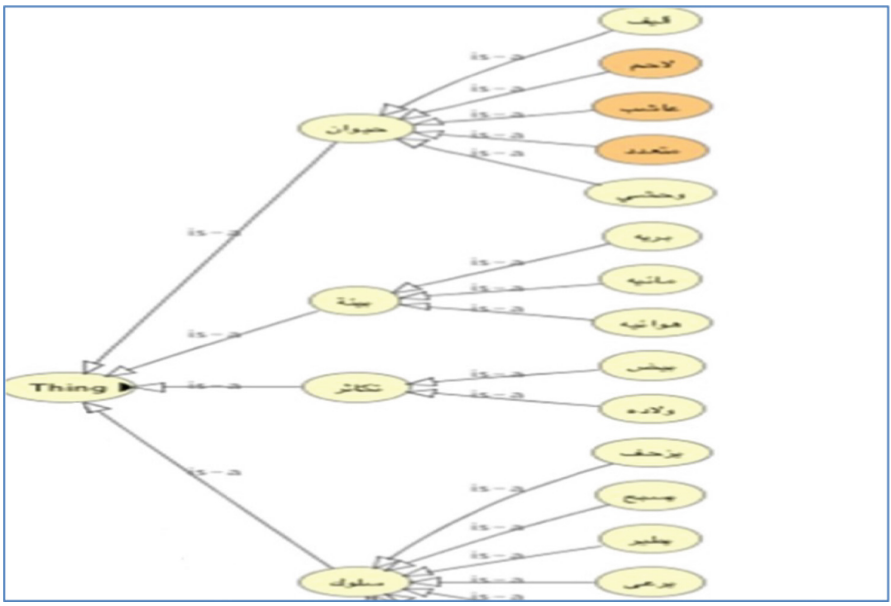


Fig. 3. An example of Arabic animal ontology

3.3 Multimedia Contents Generation

To generate the multimedia elements, we will link the terms and concepts of our ontology using a dependency graph. The ontology terms are the nodes or the vertices of the graph. The ontology can be expanded incrementally. We have proposed a model given in Fig. 4 in our previous work. The children can see the Arabic words with the corresponding image. The instructor questions are considered as semantic queries that will be addressed to the ontology which provides the responses that need to be assessed. As a result, Fig. 5 shows the output of processing an Arabic sentence and its automatic conversion to multimedia elements.

fridaya@uef.fi

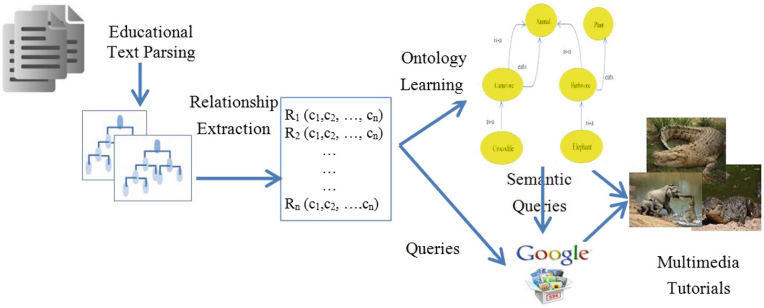


Fig. 4. An overview of dynamic multimedia-based tutorial model

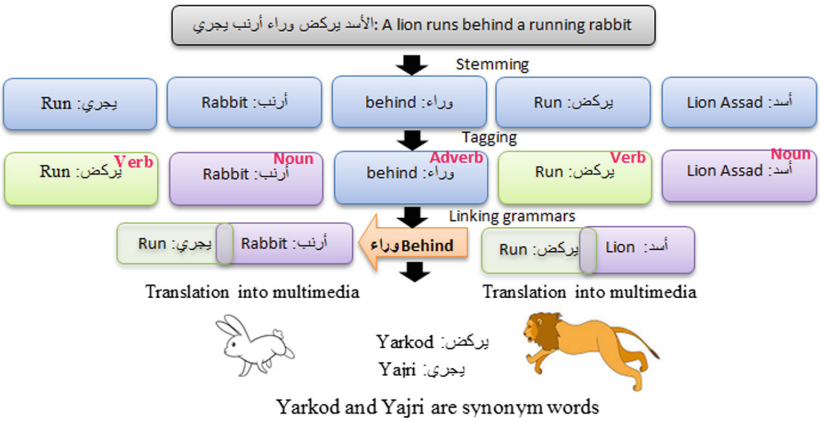


Fig. 5. An Arabic sentence with multimedia mapping

The ontology will be evaluated from both intrinsic and extrinsic aspects. For intrinsic evaluation, we want to verify the following: (1) Is the ontology modeling free of obvious inconsistencies and errors?; (2) Is the underlying ontology building structurally sound, appropriately modular, and easy-to-follow?; (3) Is the ontology designed and implemented according to sound principles of logical, semantic, and ontological analysis?; and (4) Does the ontology offer support for manipulating all desired changes? One other hand, extrinsic evaluation will focus on the usefulness of the ontology for performing knowledge management tasks. This essentially can be done through: (1) involving user domain specialists to assess the quality of the ontology in terms of the acceptability of its relationships and concepts; (2) comparing our ontology obtained results against different existing tools such as Arabic; (3) measuring the capability of the ontology to correctly provide dynamic comprehension for certain user queries.

fridaya@uef.fi

4 Comparative Study

From previous related work, we see several attempts at illustrating text with multimedia pictures to help for better understanding and communication. Accordingly, the common processing steps for most text-to-picture multimedia systems and approaches are illustrated in Fig. 6 below.

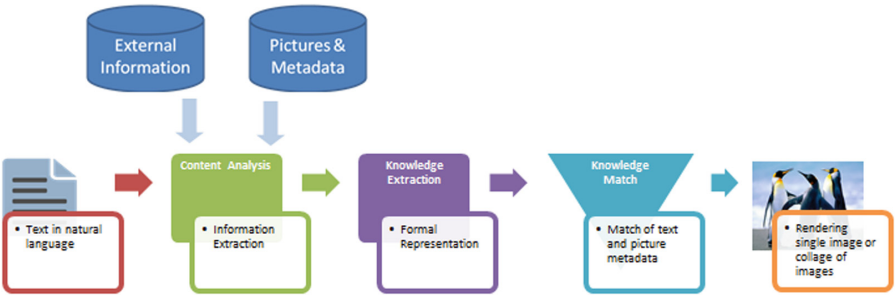


Fig. 6. Text-to-picture processing pipeline

To do a comparative study of the above surveyed text-to-picture multimedia systems, we will first summarize its technical details, text processing and rendering features based on the following criteria:

- Input: Input text can be word(s), sentence(s), paragraph(s), story, query, speech, script, and regular expression.
- Output: Determines the output modality (e.g. 2D picture collage, 3D, speech) of the system.
- Methodology: Determines the method used by the system (data-driven, rule based, and multi-agent).
- Knowledge resource: Text resource determines the exploited knowledge resource (e.g. lexicons, WordNet, etc.) and it help to add related context to the input and image resource determines the visual resources (e.g. Google images, Flickr…) and it enables automated images retrieval.
- Image ranking: Determines the method of image clustering, selection and ranking.
- Image layout: Determines the image placement and layout.

Following Table 1 gives an overall comparison between the proposed system and the previously reviewed text-to-picture multimedia systems. The supported features of each system are listed on the table header.

fridaya@uef.fi

Table 1. Comparison of text-to-pictures multimedia systems

System/year/interactive/adaptive	Input/output	Domain/methodology	Syntactic/semantic analysis	Knowledge resource	Image ranking/layout
Story picturing engine/2006/no/Yes	Paragraph/2D picture	General/data-driven	Bag-of-Words/semantic association	WordNet, Terragalleria, AMICO, ARTKids	Mutual reinforcement analysis/none
Text-to-picture synthesizer/2007, 2009/no/yes	Typed text/2D picture collage	General/data-driven	POS-tagging/association, semantic role labeling	Google Image, Google Web	largest cluster/A, B, or C region using a linear-chain CRF
Word2Image/2008/no/yes	Words being pictured/2D picture collage	General/data- driven	Bag-of-Words/semantic association	Flickr	Images clustering, saliency-based rank/images resize to cluster ranking score
Chat with Illustrator/2014/no/no	Typed short text	General/data-driven	Bag-of-Words/semantic association	Flickr	Image clustering, saliency-based rank/size, location as optimization problem
WordsEye/2011/no/no	Story/3D pictures	General/rule-based	Statistical parsing/dependency parsing	FrameNet, VigNet, WordNet, 3D Objects	None/depiction rules for 3D poses
CONFUCIUS/2006/no/ no	1 Sentence / 3D animation, audio	General/rule-based	Dependency parsing/lexical visual semantic	WordNet, LCS database, FrameNet, VerbNet	None/animation engine
Proposed system/2018/yes/yes	Story/2D pictures	Children stories/data-driven	Bag-of-Words/hierarchical association	Arabic WordNet, OWL/Scribd	Image clustering, maximum conceptual graph intersection/none

5 Conclusions

We proposed an adaptive multimedia learning system for assisting children with Arabic LD to enhance their understanding skills of Arabic vocabulary and sentences. The system dynamically generates multimedia tutorials by using text processing, relationship extraction, ontology building, and online extraction of multimedia elements. The system can be effectively used by instructors to provide children with LD the potential to learn with personalized contents designed to suit their needs. This would maximize the learning process and reduce the administrative and teaching overheads for instructors. The system can be adopted in a school environment as well as at home where parents can support their children by revising the contents they learned in schools. As future work, the aim is to develop a complete assistive platform for children with Arabic learning difficulties in an innovative manner. Such proposed platform would allow children to read, write, and understand Arabic vocabulary, grammar and sentences from multimedia elements and facial expression through the use of a haptic device to physically guide children's hand along a sequence of strokes with reference handwriting, and a realistic animated talking face to pronounce Arabic vocabulary with gestures. The platform will adopt new algorithms to automatically understand and extract concepts from standard Arabic text and represent them in

fridaya@uef.fi

images and speech. Teachers can generate adaptive and personalized multimedia-based tutorials to suit the needs of the children.

Acknowledgment. This work was made possible by NPRP grant #10-0205-170346 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

1. Lee Swanson, H., Harris, K.R., Graham, S.: Handbook of Learning Disabilities, 1st edn. Guilford Press, New York (2014). ISBN-13 978-1462518685
2. Flanagan, D.P., Alfonso, V.C.: Essentials of Specific Learning Disability Identification, 1st edn. Wiley, Hoboken (2010). ISBN-13 978-0470587607
3. Cheng, I., Basu, A., Goebel, R.: Interactive multimedia for adaptive online education. IEEE Multimedia **16**(1), 16–25 (2006)
4. Erradi, A., et al.: ArabicTutor: a multimedia m-learning platform for learning Arabic spelling and vocabulary. In: IEEE International Conference on Multimedia Computing and Systems, pp. 833–838 (2012)
5. Jumail, W., et al.: A guided digital storytelling prototype system using illustrated flashcards, pp. 1–6 (2010)
6. Saleh, M., Alja'am, J.M.: A fully accessible Arabic learning platform for assisting children with intellectual challenges. In: 14th International Conference on Computers Helping People with Special Needs, Lectures Notes in Computer Sciences, (ICCHP'2014), pp. 1–8, Paris, France (2014)
7. Dandashi, A., Saad, S., Alja'am, J.M., Mustafa, M.S.: The multimedia-based learning system improved cognitive skills and motivation of disabled children with a very high rate. J. Educ. Technol. Soc. **18**, 1–8 (2014). ISSN 1436–4522
8. Karkar, A.G., Saleh, M., Saad, S., Al Ja'am, J.M.: An Arabic ontology-based learning system for children with intellectual challenges. In: IEEE Global Engineering Education Conference, pp. 670–675 (2014)
9. Saleh, M.S., Alja'am, J.M., Karime, A., El Saddik, A.: MeMaPad: an edutainment system for assisting children with moderate intellectual and learning disability. In: Proceedings of the International Conference on Technology for Helping People with Special Needs, pp. 69–74, Riyadh, Saudi Arabia, 18–20 Feb 2013
10. Mayer, R.E.: Multimedia Learning. Cambridge University Press, New York (2001)
11. Goldberg, A.B., Rosin, J., Zhu, X., Dyer, C.R.: Toward text-to-picture synthesis. In: Proceedings of NIPS 2009 Symposium on Assistive Machine Learning for People with Disabilities (2009)
12. Zhu, X., Goldberg, A.B., Eldawy, M., Dyer, C.R., Strock, B.: A text-to-picture synthesis system for augmenting communication. In: Proceedings of the 22nd National Conference on Artificial Intelligence, Vol. 2 (2007)
13. Li, H., Tang, J., Li, G., Chua, T.-S.: Word2Image: towards visual interpretation of words. In: MM'08 – Proceedings of the 6th ACM International Conference on Multimedia, pp. 813–816 (2008)
14. Jiang, Y., Liu, J., Lu, H.: Chat with illustration. Multimedia Syst. **22**, 5–16 (2014)
15. Coyne, B., Sproat, R.: Wordseye: an automatic text-to-scene conversion system. In Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (2001)

fridaya@uef.fi

16. Eunice, M.M.: Automatic Conversion of Natural Language to 3D Animation. University of Ulster (2006)
17. Pasha, A., Al-Badrashiny, M., Diab, M., El Kholy, A., Eskander, R., Habash, N., Pooleery, M., Rambow, O., Roth, R.M.: MADAMIRA: a fast, comprehensive tool for morphological analysis and disambiguation of Arabic. In: Proceedings of the Language Resources and Evaluation Conference, Reykjavik, Iceland (2014)
18. Marton, Y., Habash, N., Rambow, O.: Dependency parsing of modern standard Arabic with lexical and inflectional features. Comput. Linguist. **39**(1), 161–194 (2013)
19. Biadsy, F., Habash, N., Hirschberg, J.: Improving the Arabic pronunciation dictionary for phone and word recognition with linguistically based pronunciation rules. In: Proceedings of the North American Association for Computational Linguistics, Boulder (2009)
20. El Kholy, A., Habash, N.: Orthographic and morphological processing for English-Arabic statistical machine translation. Mach. Transl. J., Spec. Issue Arabic Mach. Transl. **26**, 25–45 (2012)
21. Alkuhlani, S., Habash, N.: A corpus for modeling morpho-syntactic agreement in Arabic: gender, number and rationality. In: Proceedings of Conference of the Association for Computational Linguistics (ACL), Portland (2011)
22. Habash, N.: Introduction to Arabic Natural Language Processing, Synthesis Lectures on Human Language Technologies. Graeme Hirst (ed.). Morgan and Claypool Publishers (2010)

fridaya@uef.fi

Straight Line Detection Through Sub-pixel Hough Transform

Guillermo J. Bergues(✉), Clemar Schürrer, and Nancy Brambilla

Universidad Tecnológica Nacional Facultad Regional Córdoba, Córdoba, Argentina
gjbergues@gmail.com,cschurrer@gmail.com,nancybrambilla@gmail.com

Abstract. In this study, we present the sub-pixel Hough Transform, which is used to detect line positions at sub-pixel level. A comparison with the conventional Hough Transform was performed. As a result, we obtained an increase in detection resolution.

Keywords: Hough transform · Sub-pixel line detection · Metrology

1 Purpose

The edges and lines that delimit objects in images are the most important features in pattern recognition. This natural aptitude of human beings is of great interest in the area of procedure automation via computer vision, especially when we aim at obtaining a certain measurement (of position or form) [1].

For a computer vision system to be able to emulate the measurements made by a specialized operator, it is necessary to recognize the global image based on simple patterns (lines, edges, circles). For example, one of the most interesting features of this approach is the ability to recognize the straight segments present in the graduated reticles of high-precision optical measuring instruments [2].

In most applications for line or contour detection, an edge detector is used. There are many edge detection methods; however, the most common ones have been designed based on discrete derivatives applied in the image seen as a surface [3]. Nevertheless, the position of the detected edges is approximate due to loss of information in the procedure. This is a serious problem if we are attempting at obtaining an exact measurement.

Given that modern vision systems require high precision and accuracy, the development of sub-pixel techniques for detecting lines and edges is one of the most interesting topics in dimensional metrology [4,5]. In view of said matter, this paper presents the implementation of a line detection algorithm based on the Hough Transform. We introduce key modifications in the Hough transform in order to detect lines at sub-pixel level.

The conventional Hough transform [6] detects at a maximum resolution of 0.5 pixels. To improve this resolution, we analyse a new parametric space given by $\rho(G) = x \cdot cos(\theta(G)) + y \cdot sin(\theta(G))$, where (G) is each image gray level. H_i accumulators are created, corresponding to each gray level, which are then

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1129–1137, 2019.
https://doi.org/10.1007/978-3-030-22868-2_75

fridaya@uef.fi

derived into a four-dimensional density function. In its maximum peaks, this density function presents the positional information of each of the lines at sub-pixel level.

The overall algorithm performance was studied under a rigorous simulation of the lines present in the reticle of a high precision optical instrument (Nikon 6B Autocollimator) that includes the noise analysis.

2 Background: Conventional Hough Transform

The Conventional Hough Transform (CHT) is a technique used to isolate features within an image. The curves in the image are parametrized as straight lines, polynomials and circles [6]. To achieve the objective, the CHT transforms the edged binary image's discrete space, defined by pixels, into the parametric space in function with the variables θ and ρ. These two parameters define a line through $\rho = x \cdot cos(\theta) + y \cdot sen(\theta)$. In this way, a set of properties is developed between the parametric and the image plane. These properties allow us to solve the problem of detecting a line in a binary image. Figure 1 explains how the collinear points of the binary plane are transformed into a set of concurrent sinusoidal curves in the parametric plane.

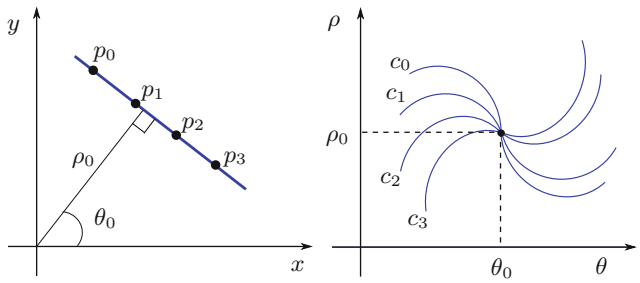


Fig. 1. Transformation of collinear points from a binary image (x, y) to the parametric plane (θ, ρ).

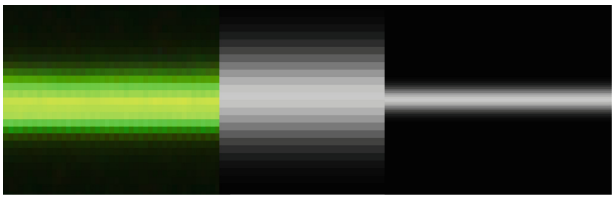
2.1 Edge Detectors

The CHT algorithm works with an edge detector filter which allows the image plane to be transformed into a binary image plane. The most commonly used filters are Sobel and Canny, which include some pre-processing tasks to reduce the image noise [7,8]. Being the mostly used filter, Canny poses a disadvantage: in addition to removing the noise, it blurs the image, causing the finest details to be lost [9, p. 176]. Therefore, Canny's Gaussian filter causes the CHT not to accurately detect line position, which generates a greater uncertainty in the measurement.

fridaya@uef.fi

2.2 Simulation

To determine the algorithms' detection accuracy, we used a simulation which represents the scale segments of a Nikon 6D autocollimator (see Fig. 2). In Fig. 2a, we can observe a line segment extracted from a real captured image and in Fig. 2b, a simulated segment. As a result, the real lines of both the instrument's cross and scale are represented (see Fig. 2c), [4]. And we introduce noise in every image using a signal-to-noise ratio (S/N) as a measure of it. The use of this simulated data enabled the analysis of each simulated detector's performance.



(a) Real Line. (b) Simulated Line. (c) Simulated image.

Fig. 2. Segment of the real line (a) and the simulated line (b). Simulated image (c).

The equation used to generate the Gaussian line in an image I formed by a matrix (N_x, N_y) was the following:

$$I(i,j) = round(A.e^{\frac{-(m-j)^2}{2.\Delta^2}}),\ i_0 = \frac{N_y}{2} \in N, \tag{1}$$

The values of c (position of the simulated line) to be estimated are $m = i_0 + k$ with $k = 1/100$. The *round* function is the round to the nearest integer function. The width of the line given by Δ allows us to build a line according to the number of pixels that the line in the captured images occupies (using a BASLER ACA 1600/20gc camera as a model). Where:

$$0 \leq I(i,j) \leq 255\ e\ I(i,j) \in N, \tag{2}$$

2.3 CHT Performance

The Sobel filter is used in pre-processing because it is the one that is best suited for this type of images [4]. On the other hand, taking into account the sub-pixel centroid values c' estimated for each detector as a linear function of the simulated centroids c, the Pearson correlation coefficient r [10] is used as a quality estimation measurement.

In this first study, the line was moved at pixel level in the simulation image with integer values of $k = 1, ..., 100$ on (1) because the CHT doesn't detect at sub-pixel level. The results for a $S/N = 23\,\text{dB}$ are shown in Fig. 3; and in Table 1, the detection under noise and its behaviour can be observed, using the Pearson coefficient as a measure.

fridaya@uef.fi

Table 1. D_{H1}: Pearson coefficient values for the different noise levels studied.

S/N	r_H
1 dB	0.18736
5 dB	0.47941
10 dB	0.95070
18 dB	0.999
23 dB	0.99999

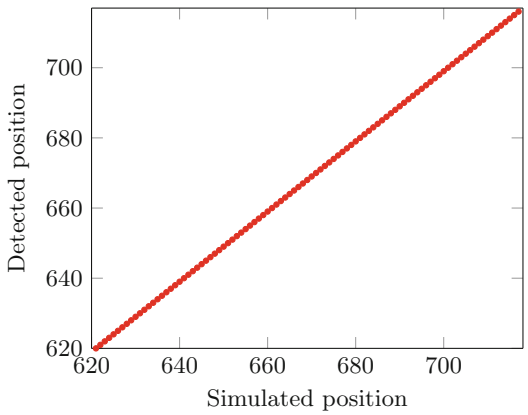


Fig. 3. CHT detector: Detected line position vs. simulated position (spacing between lines equals 1 pixel) $r \approx 0.999$, (with $S/N = 23\,\text{dB}$)

2.4 CHT Resolution

Although the above results are correct for some applications, the CHT algorithm detects at sub-pixel level with a maximum resolution of 0.5 pixels (at sub-pixel level, $k = 1/100$, there is not detection). On the other hand, if line thickness is formed by an even number of pixels, the transform chooses one of the two, according to its weight, introducing errors in the position measurement. This problem is schematized in Fig. 4, where the representation of the captured line of Fig. 2a is shown, with its respective central detection starting from the two edges that emerge from the filter.

3 Method

In order to improve the resolution of 0.5 pixels, it is not enough to apply any of the conventional detection methods that derive from the CHT, since similar results would be obtained again. This is so because in the CHT, the information contained in the gray levels is almost completely lost when the thresholding is performed to obtain the binary image (Sobel in our case). The pre-processing

fridaya@uef.fi

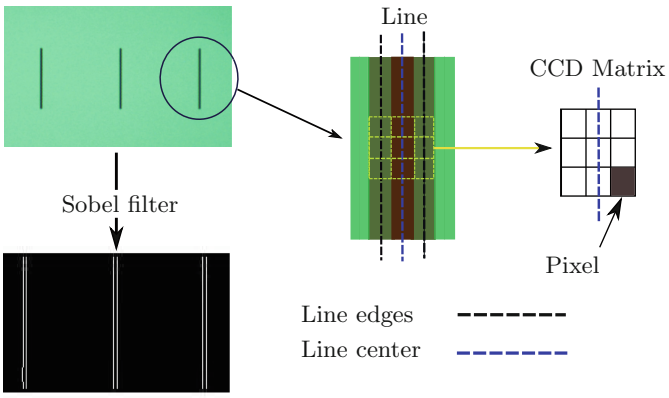


Fig. 4. Sub-pixel value of the detected line position. The pixel's central value of the CCD grid is taken from the edges detected by the Sobel filter. Maximum resolution value: 0.5 pixels.

tasks, both thresholding and edge detection, are necessary in the CHT, but they can be ignored if a new parametric space is created. This new space has to include the intensity value given by each grey level [11]. This space, called "gray Hough parameter counting space", allows us to originally detect linear bands (of a single grey level) at the different intensity levels of the image. Each band is defined as a parallel set of lines with the same orientation and grey level, whose main characteristics to obtain are: gray level, orientation, total number of pixels and length.

This highly useful approach for applications such as robotics and industrial processes is not suitable for measurements where the line position must be obtained with sub-pixel precision (not the band, but the central position of the band). On the other hand, the lines analysed are made up, in most cases, by several grey levels at the same time.

3.1 Sub-pixel Hough Transform

Based on $\rho = x \cdot cos(\theta) + y \cdot sen(\theta)$, with an image space given by (x, y, G), in which G is the gray level corresponding to each point (x, y) of the image, a new parametric space is formed:

$$\rho(G) = x \cdot cos(\theta(G)) + y \cdot sen(\theta(G)), \tag{3}$$

This parametric space defines a mapping space $f : (x, y, G) \rightarrow H(\rho, \theta, G)$ which builds H_i accumulators that correspond to each gray level. $2^{n^\circ bits}$ accumulators -unified in a 4D matrix- are created, being n the quantization bits number of the camera.

The $4D$ matrix behaves as an easy-to-understand density function: its maximums define the position of the lines in the space (x, y). In Fig. 5, as an example,

fridaya@uef.fi

the local maximum values of the H_n accumulator are shown. Black color responds to the region most likely to encounter a straight line for the gray level n.

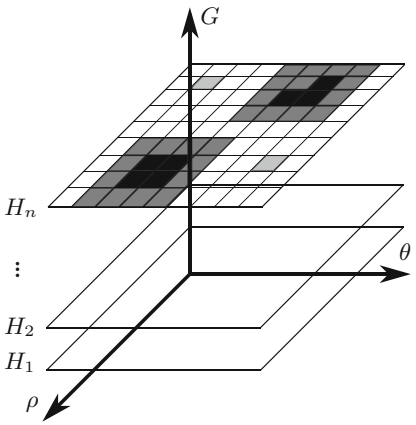


Fig. 5. Gray Hough parameter counting space. There is a sequence of n matrices that form the $4D$ matrix. Each cell of matrix H_n corresponds to a counting element for the gray level n.

It is important to emphasize that the CHT works with a single accumulator H that includes all the information corresponding to the number of lines in the image and, therefore, all the information that is discarded in the thresholding stage is lost. The new parametric space does not lose relevant information for line positioning.

The following algorithm summarizes the procedure:

1. Initialize each accumulator $H(\rho, \theta, G)$ a 0.
2. For each pixel (x, y, G) y $\theta_j = 0° \rightarrow 179°$.
3. Calculate $\rho_j(G_i) = x_i \cdot cos(\theta_j(G_i)) + y_i \cdot sen(\theta_j(G_i))$.
4. $H(\rho_j, \theta_j, G_i) = H(\rho_j, \theta_j, G_i) + 1$.
5. Create the final matrix $H_T = H_1 + ... + H_n$.
6. Compute the points with the highest degree of probability in the parametric plane.
7. Map these points onto the image plane and obtain the density function center that is defined by them.

From a geometric structure point of view, the lines are composed of several connected parallel straight lines, which have an equal θ orientation, have different distances ρ_i and different gray levels (see Fig. 2b). Once the transformation to gray levels is performed, each of these lines is mapped onto a local maximum in its respective accumulator H_i (as in accumulator H_n of Fig. 5). These maximum values are forming the peaks shown in Fig. 6, which are then searched to obtain the center of the line in the pixels plane (x, y).

fridaya@uef.fi

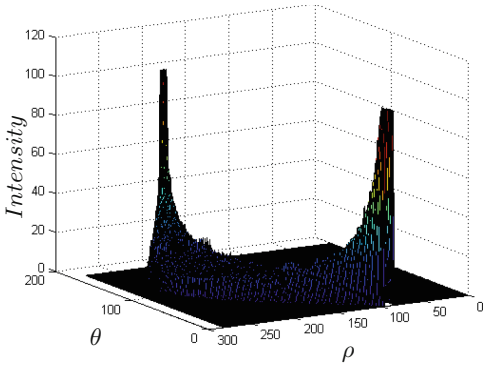


Fig. 6. Density function of final parametric plane. It is obtained from the sum $H_T = H_1 + ... + H_n$. The peaks present in the density function correspond to the searched lines' centers.

3.2 Results

Table 2 shows the results of the Pearson coefficients calculated for the different noise levels studied. Figure 7 shows the relationship between the simulated positions and those obtained by the Hough sub-pixel detector (with $S/N = 23\,\text{dB}$). Each noise level has a similar curve. On the other hand, Fig. 8 is a summary of the sub-pixel detector study. It shows the relationship between the signal-to-noise ratio S/N and its r value.

Table 2. Sub-pixel Hough transform: Pearson coefficient values for the different noise levels studied.

S/N	r_H
13 dB	0.9494
14 dB	0.9512
15 dB	0.9518
16 dB	0.9579
17 dB	0.9645
18 dB	0.9669
19 dB	0.9748
20 dB	0.9804
21 dB	0.9811
22 dB	0.9821
23 dB	0.9846

fridaya@uef.fi

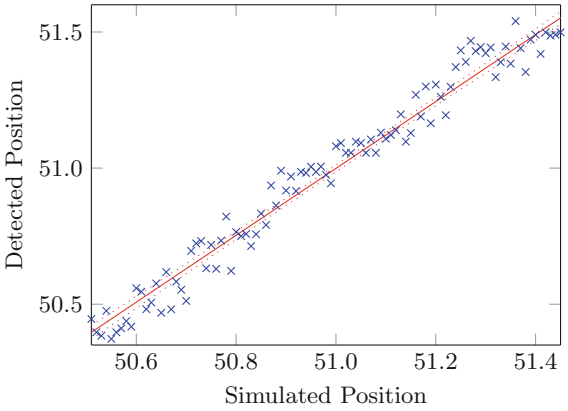


Fig. 7. Sub-pixel Hough Detection: Detected position vs. Simulated position. $r = 0.9846$ ($S/N = 23\,\mathrm{dB}$).

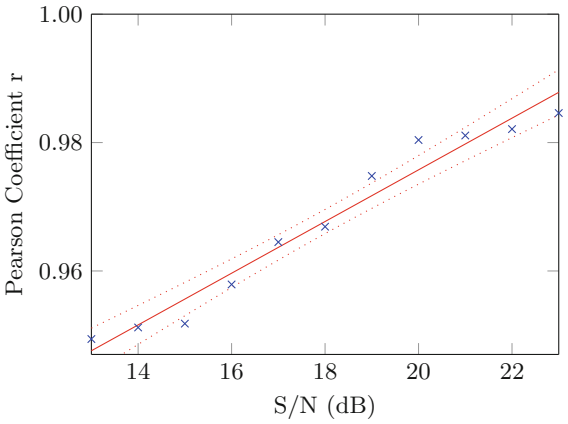


Fig. 8. Sub-pixel Hough detection: r vs. S/N.

4 Conclusions

After analysing Fig. 8, it is clearly seen that, as the S/N increases, sub-pixel detection is better (r approaches the ideal value 1). There is a fairly linear growth that stops in 20 dB and remains on a slightly positive slope's plateau. This behaviour is so due to the fact that the density function represents a part of all the information on the positions of the lines. The detector has an operating stop around the value of $r = 0.9850$.

Another important fact that emerges from the analysis of the curves is that the dispersion degree of the data, according to the regression line, is very good. The average quadratic dispersion (s) is equal to 0.05 pixels using a BASLER ACA 1600 20gc camera as a model for the simulation.

fridaya@uef.fi

As a main result, this new algorithm allows for line detection at sub-pixel level, being this 10 times better in resolution than the CHT (which detects at a maximum resolution of 0.5 pixels). Moreover, this new algorithm does not depend on an odd pixel structure like the CHT. On the other hand, this maximum value is tied to the characteristics of the camera (n° of bits and resolution). If these latter values increase, the new parametric space can be more detailed and therefore a higher resolution can be obtained.

References

1. Fernandes, L., Oliveira, M., Silva, R.: Uncertainty propagation: avoiding the expensive sampling process for real-time image-based measurements. Comput. Stat. Data Anal. **52**(7), 3852–3876 (2008)
2. Bergues, G., Ames, G., Canali, L., Schurrer, C., Flesia, A.: External visual interface for a Nikon 6D autocollimator. In: Proceedings of the 2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), pp. 35–39, May 2014
3. Maini, R., Aggarwal, H.: Study and comparison of various image edge detection techniques. Int. J. Image Process. (IJIP) **3**(1), 1–11 (2009)
4. Flesia, A.G., Ames, G., Bergues, G., Canali, L., Schurrer, C.: Subpixel straight lines detection for measuring through machine vision. In: Proceedings of the 2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), pp. 402–406, May 2014
5. Fabijanska, A., Sankowski, D.: Computer vision system for high temperature measurements of surface properties. Mach. Vis. Appl. **20**(6), 411–421 (2009)
6. Duda, R.O., Hart, P.E.: Use of the Hough transformation to detect lines and curves in pictures. ACM **15**(1), 11–15 (1972)
7. Senthilkumaran, N., Rajesh, R.: Edge detection techniques for image segmentation-a survey of soft computing approaches. Int. J. Recent Trends Eng. **1**(2), 250 (2009)
8. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **PAMI–8**(6), 679–698 (1986)
9. Gonzalez, R., Woods, R.E.: Digital Image Processing, pp. 176–178. Prentice Hall, Upper Saddle River (2002)
10. Niven, E.B., Deutsch, C.V.: Calculating a robust correlation coefficient and quantifying its uncertainty. Comput. Geosci. **40**, 1–9 (2012)
11. Lo, R.C., Tsai, W.H.: Gray-scale Hough transform for thick line detection in gray-scale images. Pattern Recogn. **28**(5), 647–661 (1995)

fridaya@uef.fi

Microcontroller-Based Speed Control Using Sliding Mode Control in Synchronize Reluctance Motor

Wei-Lung Mao[1(✉)], Gia-Rong Liu[1], Chao-Ting Chu[2], and Chung-Wen Hung[1]

[1] National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan
wlmao@yuntech.edu.tw

[2] Chunghwa Telecom Laboratories, Chunghwa Telecom Co., Ltd., No. 99, Dianyan Road, Yangmei District, Taoyuan City 32661, Taiwan, R. O. C

Abstract. This paper proposes the implementation of integral sliding mode control for a synchronous reluctance motor (SynRM) using microcontroller-based architecture. The technology of microcontroller system has been quite mature, and the price is much cheaper than the conventional dSPACE control system. The integral sliding mode control algorithm is designed for speed control such that it can alleviate the chattering phenomenon in the presence of unmodeled system dynamics and external disturbances. The SMC method, MTC algorithm and PWM calculations are implemented in MSP432 platform. The microcontroller dsPIC30F4011 can process the three-phase PWM output and the motor speed data information. The motor speed data is transferred to the MSP432 using the UART interface. Both microcontroller-based and dSPACE control architectures are implemented and compared. Experimental results show that the proposed method is robust and cheap for speed control of SynRM system.

Keywords: Microcontroller-based control · Synchronous reluctance motor (SynRM) · Integral sliding mode control · dSPACE control

1 Introduction

A microcontroller is a small computer system which plays a vital role in industrial automation and motion control applications [1, 2]. The synchronous reluctance motor (SynRM) [3] has simple and robust structure that is suitable and is used in high-speed and high-precision environments. The ease of control and excellent performance of the SynRM motors will ensure that their applications continue to grow in the future. This research aims to control the speed of SynRM motor using a microcontroller system for cost and flexible performances. Various speed control methods are applied and can be classified into self-tuning PID method and sliding mode control (SMC) method. The SMC has become a popular research topic in robust controllers [4, 5]. However, a critical problem with sliding mode controllers is that the system includes chatter. Recently, the sliding mode controller is frequently being combined with a neural network to overcome this problem [4].

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1138–1149, 2019.
https://doi.org/10.1007/978-3-030-22868-2_76

fridaya@uef.fi

A complementary sliding mode controller with a recursive neural network applied on a three-axis platform was proposed to accurately estimate the system error and lump uncertainty [5]. Niu et al. applied a sliding mode neural network estimator [6] with the Lyapunov function in the delay system, which produced a good output response. A sliding mode controller with a neural network estimator algorithm was used [7, 8] which the estimator quickly tracked errors in the maglev system. When external interference occurred, the estimator could respond accurately and return to the command position efficiently. One major criticism of the neural network estimator is its long calculation time. An integral sliding mode controller is designed to deal with these difficulties.

In consideration of the controller price, some researcher [9–11] proposed microcontroller based architecture to control motor system. Šetka et al. [9] proposed motor controller designed for robotics using microcontroller with integrated EtherCAT, in which the experimental results have satisfactory response. Taveiros et al. [10] published conventional and adaptive strategies applied to DC motor control using microcontroller PIC18F4520. Since DC motor does not need the complex control method, the PIC18F4520 microcontroller can provide compliance with efficiency. In [11], a sun tracking system based on stepper motor, and microcontroller 16f877a had been proposed by Awad et al. In this research, an integral sliding mode controller is designed and implemented for SynRM system using microcontroller-based architecture. The stepper motor also does not need complex PWM control. This research proposes the integration of microcontrollers MSP432 and dsPIC30F4011 to control SynRM, in which the speed controller uses SMC method to track command speed. In the experiments, it is shown that our method has satisfactory performance compared with traditional dsPACE 1104 control card architecture.

2 The SynRM Motor System Model

Figure 1 shows the d − q axis equivalent architecture for the SRM. The voltage equations are expressed as:

$$V_{ds} = R_s i_{ds} - \omega_r L_{qs} i_{qs} + L_{ds} \frac{di_{ds}}{dt} \tag{1}$$

$$V_{qs} = R_s i_{qs} + \omega_r L_{ds} i_{ds} + L_{qs} \frac{di_{qs}}{dt} \tag{2}$$

where V_{ds}, V_{qs} are the d-q axis voltages, i_{ds}, i_{qs} are the d-q axis currents, L_{ds}, L_{qs} are the inductances, R_s is the copper loss resistor, and ω_r is the speed for SRM.

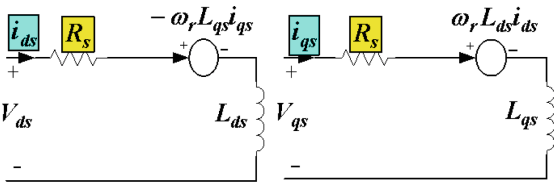


Fig. 1. d-q axis equivalent architecture for the SRM.

fridaya@uef.fi

Figure 2 shows the architecture of SRM system, and the dynamic equation of the motor is represented as

$$T_e = J_m \frac{d\omega_r}{dt} + B_m \omega_r + T_L \tag{3}$$

where T_e is the torque of SRM, T_L is the external load, J_m is the moment of inertia, and B_m is the coefficient of friction.

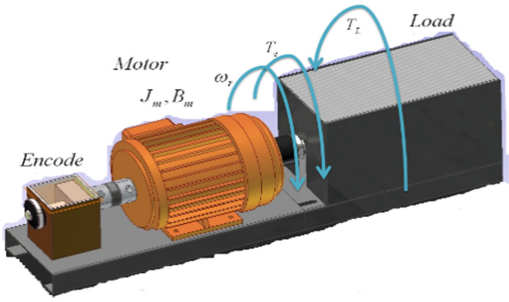


Fig. 2. The architecture of SRM system.

The corresponding electromagnetic torque equation is given as:

$$T_e = \frac{3}{4} P(L_{ds} - L_{qs}) i_{ds} i_{qs} \tag{4}$$

where P is the pole number in the SRM. The maximum torque control (MTC) scheme is applied [12, 13] here, and the torque current commands are given as:

$$i_{ds} = \sqrt{\frac{|T_e|}{\frac{3P}{4}(L_d - L_q)}} \tag{5}$$

$$i_{qs} = \mathrm{sgn}(T_e) \sqrt{\frac{|T_e|}{\frac{3P}{4}(L_d - L_q)}} \tag{6}$$

Figure 3 illustrates the architecture of our proposed speed control for a SynRM system. The dynamic equation of SRM is expressed as

$$\begin{bmatrix} \dot{\theta} \\ \dot{\omega}_r \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{B_m}{J_m} \end{bmatrix} \begin{bmatrix} \theta \\ \omega_r \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{J_m} \end{bmatrix} T_e + \begin{bmatrix} 0 \\ T_L \end{bmatrix} \tag{7}$$

fridaya@uef.fi

Therefore, we have

$$\begin{aligned}\dot{\omega}_r &= -\frac{B_m}{J_m}\omega_r + \frac{1}{J_m}(T_e - T_L) \\ &= A_r\omega_r + B_r(T_e - T_L) = a_r\omega_r + b_r u(t) + E\end{aligned} \tag{8.1}$$

$$y_{motor} = \omega_r \tag{8.2}$$

With $A_r = -J_m^{-1}B_m = a_r + \Delta a_r$, $B_r = J_m^{-1} = b_r + \Delta b_r$, $u = T_e$, $E = (\Delta a_r\omega_r + \Delta b_r u(t) - b_r T_L)$

where Δ is the uncertainty, and E is the lump uncertainty. We assume that the parameter E is bounded.

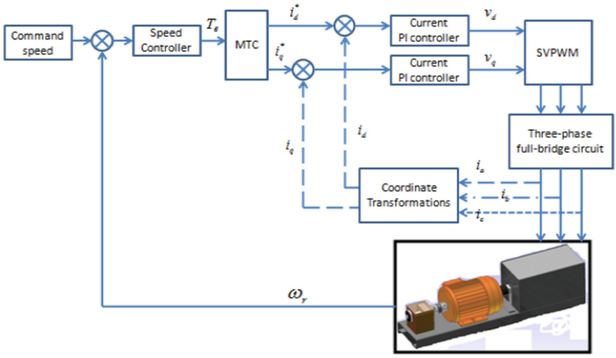


Fig. 3. System model of SRM.

3 Design of Integral Sliding Mode Controller

The output speed error is defined as:

$$e(t) = \omega_r^* - \omega_r \tag{9}$$

We define sliding function as

$$S(t) = e(t) + c_2\int_0^t e(t)dt \tag{10}$$

where $c_2 > 0$. Define the Lyapunov candidate function as

$$V = \frac{1}{2}S^2 \tag{11}$$

fridaya@uef.fi

Hence

$$\dot{V} = S(\dot{e}(t) + c_2 e) = S\left(\dot{\omega}_r^* - a_r \omega_r - b_r u(t) - E + c_2 e\right) \tag{12}$$

We have

$$u = \frac{1}{b}\left[\dot{\omega}_r^* - a_r \omega_r + c_2 e + k_v S + \varepsilon \mathrm{sgn}(S)\right] \tag{13}$$

with $k_v > 0$.

Let $|E| \leq \varepsilon$ and substitute Eq. (13) into Eq. (12), we have

$$\dot{V} \leq -(k_v S^2 - |E||S| + \varepsilon|S|) \leq 0 \tag{14}$$

Equation (15) shows that the Lyapunov function is stable and convergent. Defining uniform continuity function $\xi(t)$ as

$$\xi(t) = (k_v S^2 - |E||S| + \varepsilon|S|) \tag{15}$$

Integrating the Eq. (16) with respect to time, it yields

$$\int_0^t \xi(\tau)d\tau \leq V(S(0)) - V(S(t)) \tag{16}$$

Since $V(S(t))$ and $V(S(0))$ are bounded, so it can be seen

$$\lim_{t\to\infty} \int_0^t \xi(\tau)d\tau < \infty \tag{17}$$

With the use of Barbalat lemma [14], it shows

$$\lim_{t\to\infty} \xi(\tau) = 0 \tag{18}$$

Therefore, the designed SynRM speed control system is stable, and the error will converge to zero asymptotically. The stability of our developed SMC system is guaranteed.

4 Microcontroller-Base Architecture

This research realizes the motor control technology using the lower-priced microcontroller. The architecture diagram of the dual-microcontroller motor control is shown in Fig. 4. It consists of the two microcontrollers, those are MSP432 and dsPIC30F4011.

fridaya@uef.fi

The SMC and MTC algorithms are implemented in MSP432 chip. The functions of PWM and position feedback are developed in dsPIC30 chip. In our developed motor control system, it also includes the Hall current sensing components, inverter module, synchronous reluctance motors, and magnetic power electric actuators.

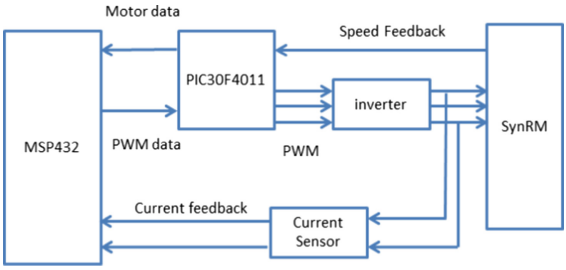


Fig. 4. Architecture of dual chip motor controller.

4.1 dsPIC30F4011 Digital Controller

The microcontroller dsPIC30F4011 is produced by Microchip Corporation. This series of microcontrollers are widely used in power control, signal processing, system control, etc. Figure 5 shows the dsPIC30F4011 pin map.

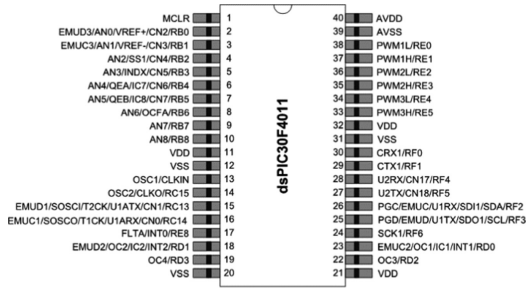


Fig. 5. dsPIC30F4011 pin assignment.

The function of quadrature encoder interface (QEI) is used as optical encoder. It can detect position and speed of the motor system. QEI function includes three input channels, which are QEA, QEB and index pulse (Index). It reads QEA and QEB signals as the positioning calculation. At the same time, the difference of the two-phase voltage signal is used to increase the positioning accuracy to four times calculating the current position. The direction judgment diagram is shown in Fig. 6.

The PWM module simplifies multiple simultaneous pulse width modulated outputs and supports four control applications for three-phase AC induction motors, switched reluctance motors, brushless DC motors, and uninterruptible power supply systems. It has six PWM output input pins and three duty cycle generators. When any pair of

fridaya@uef.fi

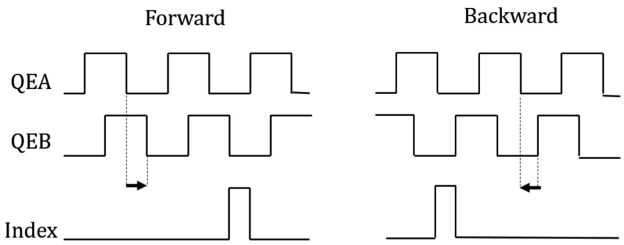


Fig. 6. Positioning encoder interface direction judgment method.

PWM output pins is complementary mode, PWM can provide idle time generation. In Fig. 7, the complementary PWM ensures that one set upper and lower arm PWMs will not generate a high state at the same time, and the dead time of the PWM can be set to protect the power crystal [15], so it is very suitable for application in three phase motor control.

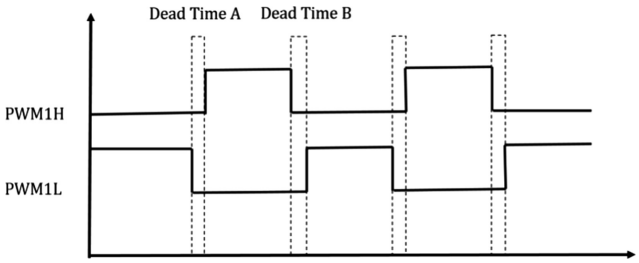


Fig. 7. The idle time of the motor control PWM module generates an output signal.

4.2 MSP432 Processor

The microcontroller MSP432 is produced by Texas Instruments Inc. as shown in Fig. 8. It is a low-power and high-performance 32-bit microcontroller. The ARM® Cortex®-M4F MCU optimizes the power performance, in which startup and standby power consumption are only 95μA and 850nA, respectively. It integrates digital signal processing (DSP) engine and the floating-point computing core (FPU) in the ARM Cortex-M4F that enables a wide range of high-performance applications such as signal conditioning and sensor processing. It is feasible to develop the SMC control algorithm using C language on this platform.

4.3 dsPIC30F4011 and MSP432 Integrated Development

In this paper, the SMC method, MTC algorithm, and PWM calculations are implemented in MSP432 platform. The dsPIC30F4011 is utilized for three-phase PWM output and motor speed revolution data. The data is transferred to the MSP432 for

fridaya@uef.fi

Fig. 8. Microcontroller MSP432.

calculation by the UART interface. The MSP432 converts the calculated control command into three-phase PWM data, and then sends to dsPIC30F4011 through UART interface. It can directly control the PWM position and then output the three-phase voltage of the motor to make the system achieve effective motor control.

4.4 Comparison of dSPACE Control and Dual Microcontrollers

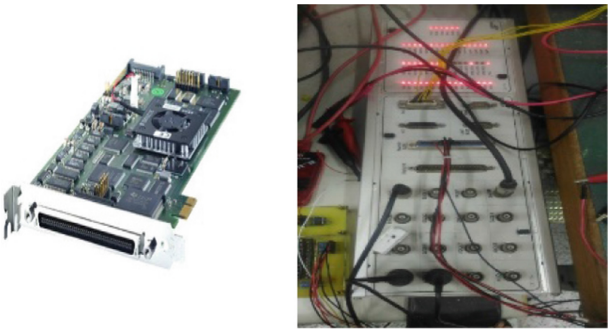
In this research, the dSPACE control method is also implemented using DS1104 controller board (dSPACE, Inc., Germany) for performance comparison. The cost of the DS1104 control card with the signal converter group is about NTD 55,000, and the dual microcontroller is about NTD 1000. The prices between the two systems are very large. In terms of software, the DS1104 control is designed with Simulink in MATLAB, and the microcontroller-based system is using C language. Simulink needs to design the functional blocks, and C language have to learn a lot of programming concepts. The DS1104 control card is equipped in a desktop computer, so it cannot be carried around. On the other hand, the microcontroller control is compact and portable.

The comparisons between the dSPACE control and microcontroller-based structure are shown in Table 1. Figure 9(a) and 9(b) show the DS1104 control card and signal converter, respectively. Figure 10 shows the dual microcontroller structure.

Table 1. Comparisons between dSPACE control and dual microcontrollers

	dSPACE controller	Dual chip controller
Price	NTD 55000	NTD 1000
Software design	MATLAB/simulink	C language
Hardware	Inconvenient to carry	Easy to carry
Application	Low	High

fridaya@uef.fi



(a) (b)

Fig. 9. (a) DS1104 control card, (b) DS1104 control card signal converter.

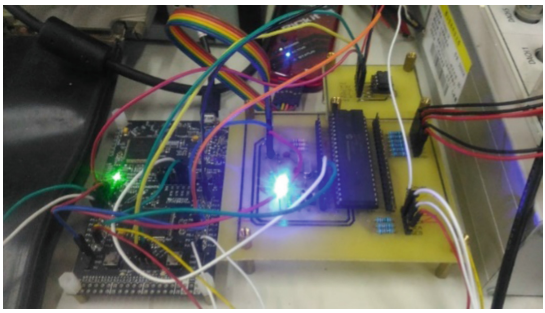


Fig. 10. Dual microcontroller structure.

Table 2. Case of experimental results

	ω_{ref} command speed
Parameters	$\omega_{ref} = 600rpm$, 0–10 s $\omega_{ref} = 800rpm$ at 10 s

5 Experimental Results

In order to compare the performance of our proposed controller, the experimental conditions are shown in Table 2.

The experimental setup for the implementation of the proposed SynRM motor system is depicted in Fig. 11. Figure 12 is the experimental responses using dSPACE control method. The reference speed command is set to 600 rpm at 0 s, and changed to 800 rpm at 10 s, when the dsPACE control card with sliding mode controller is developed. Figure 12(a)–(d) are the speed response, control command, speed error and the sliding function responses. Figure 13 is the experimental results using

fridaya@uef.fi

microcontroller-based architecture. The speed command is the same as the above conditions. Figure 13(a)–(d) are the speed response, control command, speed error and the sliding function responses. The robust control property of our developed system can be observed clearly. The speed errors of our microcontroller system can decrease significantly, and the controlled system can catch up the system command swiftly.

Fig. 11. SynRM control system setup.

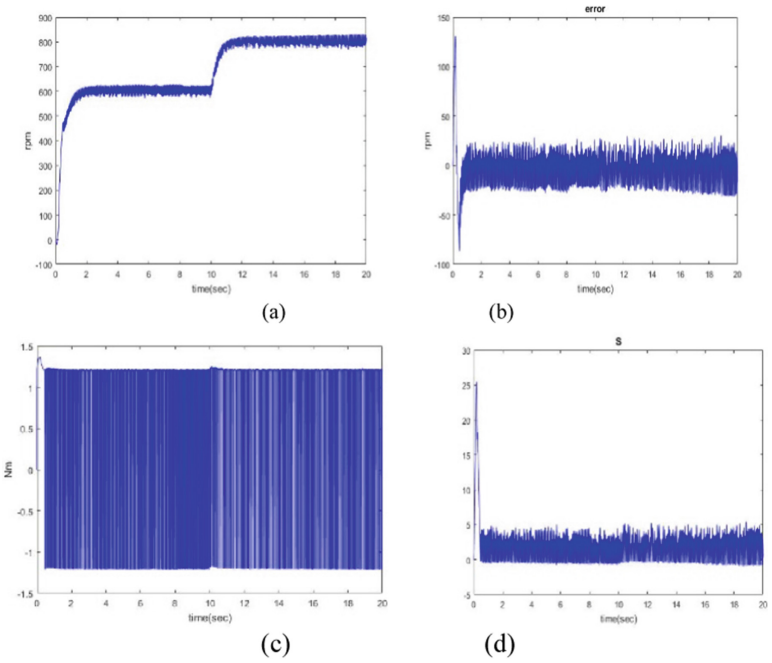


Fig. 12. dSPACE structure using integral sliding mode controller (a) speed response, (b) control value, (c) speed error, (d) sliding function.

fridaya@uef.fi

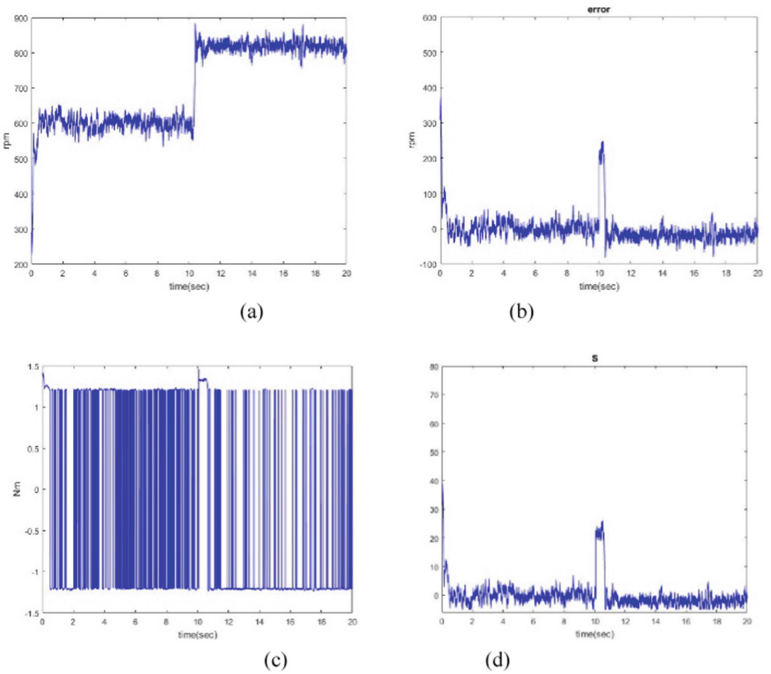


Fig. 13. Microcontroller-based structure using integral sliding mode controller, (a) speed response, (b) control value, (c) speed error, (d) sliding function.

6 Conclusion

This paper proposes microcontroller-based control and dSPACE control architectures with integral sliding mode controller for SymRM speed control. The experimental results show that sliding mode control can have faster response and provide the robustness. For microcontroller-based control, it shows that the proposed SMC scheme can provide superior speed tracking capability for SynRM system in the presence of model uncertainties and external disturbances. The microcontrollers can effectively and simply generate the PWM signal and floating-point computing capability. The microcontroller-based structure can be implemented effectively and the feasibility can be provided in this research. Therefore, the proposed SMC strategy can be widely implemented in industrial applications.

Acknowledgment. The authors would like to thank the Ministry of Science and Technology of the Republic of China, Taiwan, for financially supporting this research under Contract No. MOST 107-2221-E-224-040-.

fridaya@uef.fi

References

1. Reddy, P.B., El-Refaie, A.M., Zou, M., Pan, D., Alexander, J., Tapadia, P., Grace, N.K., Huh, K.K., Johnson, F.: Performance testing and analysis of synchronous reluctance motor utilizing dual-phase magnetic material. IEEE Trans. Ind. Appl. **54**(3), 2193–2201 (2018)
2. Chu, C.T., Wang, Y.K., Liang, H.L.: SISO PID-fuzzy controller for BLDC motor speed control base on low cost MSP430 solution. In: 5th International Symposium on Next-Generation Electronics (ISNE), pp. 1–2, Taiwan (2016)
3. Chu, C.T., Chiang, H.K., Chung, K.C., Lin T.C.: Self-tuning single input fuzzy controller for a synchronous reluctance motor. In: International Symposium on Next-Generation Electronics (ISNE), pp. 1–4, Taiwan (2014)
4. Lin, F.J., Shyu, K.K., Wai, R.J.: Recurrent-fuzzy-neural-network sliding-mode controlled motor-Toggle servomechanism. IEEE/ASME Trans. Mechatron. **6**(4), 453–466 (2001)
5. Lin, F.J., Chen, S.Y., Shyu, K.K., Liu, Y.H.: Intelligent complementary sliding-mode control for LUSMs-based x-y-z motion control stage. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **57**(7), 1626–1640 (2010)
6. Niu, Y., Wang, X., Lam, J., Ho, D.W.C.: Sliding-mode control for nonlinear state-delayed systems using neural-network approximation. IEE Proc. Control Theor. Appl. **150**(3), 233–239 (2003)
7. Lin, F.J., Chen, S.Y., Shyu, K.K.: Robust dynamic sliding-mode control using adaptive RENN for magnetic levitation system. IEEE Trans. Neural Networks **20**(6), 938–951 (2009)
8. Li, X., Yu, W.: Neural sliding mode control for magnetic levitation systems. In: IEEE International Conference on Control Applications (CCA), pp. 615–619, (2010)
9. Šetka, V., Tolar, D.: Motor controller designed for robotics based on microcontroller with integrated EtherCAT. In: 19th International Carpathian Control Conference (ICCC), pp. 289–294 (2018)
10. Taveiros, F.E.V., Barros, L.S., Barros, C.M.V.: Conventional and adaptive strategies applied to DC motor control using microcontroller PIC18F4520. In: Simposio Brasileiro de Sistemas Eletricos (SBSE), pp. 1–6 (2018)
11. Awad, H., Aziz, S.A.: Development and control of a sun tracking system based on stepper motor and 16f877a microcontroller. In: Nineteenth International Middle East Power Systems Conference (MEPCON), pp. 1–6 (2017)
12. Chen, C.A., Chiang, H.K., Lin, B.R.: The novel adaptive sliding mode controller for synchronous reluctance motor drive. Eur. Trans. Electr. Power **20**(8), 1058–1070 (2010)
13. Chiang, H.K., Tseng, C.H.: Design and implementation of a grey sliding mode controller for synchronous reluctance motor drive. Control Eng. Pract. **12**(2), 155–163 (2004)
14. Narendra, K.S., Annaswamy, A.M.: Stable Adaptive Systems. Prentice-Hall, Englewood Cliffs (1988)
15. Broeck, H.W.V.D.: Analysis and realization of a pulsewidth modulator based on voltage-space vector. IEEE Trans. Ind. Appl. **24**(1), 142–150 (1988)

fridaya@uef.fi

A Breath Switch for Secure Password Input

Natsuki Sayama, Naoka Komatsu, Kana Sawanobori, and Manabu Okamoto(✉)

Kanagawa Institute of Technology, Atsugi, Kanagawa, Japan
manabu@nw.kanagawa-it.ac.jp

Abstract. One of the most popular authentication methods in information security is the use of passwords. However, people who cannot use typical keyboards, like those having physical disabilities, cannot input their passwords efficiently. With assistive technologies, they can move a cursor and/or hold a pointer in their mouth to enter characters shown on a screen keyboard. However, if someone is next to them and is looking at the screen, they can easily see the characters being entered. We propose a breath-based password input system, in which characters can be entered via exhalation. This will allow people with severely limiting physical disabilities to securely input their secret information.

Keywords: Assistive technology · Authentication · Password

1 Introduction

ID/password authentication is required by most personal services on the internet and several applications on our desktops and smartphones. However, we also need to consider people with physical disabilities who cannot securely input their ID/password as they cannot use a typical input keyboard.

These people use assistive technologies like on-screen keyboards that help the user use keys on a software keyboard. Clicking the button on an on-screen keyboard matrix on the PC's screen works exactly like pressing a physical keyboard key. Moreover, they can press keys on this keyboard with the help of a "mouth stick" [2] that they can hold in their mouths. Alternatively, they can use a trackball with their feet, chins, elbows, or heads instead of their hands. Various assistive technologies like these have been proposed for people with physical disabilities and put to practical use. Other supporting systems include voice recognition systems that use their speech as input, or Gaze-input, which allows people to move a cursor with their eyes.

However, these support technologies do not allow discretion. You can press a key on an on-screen keyboard [1] using a mouse stick, but a bystander can observe all their input actions such as selections and clicks. Moreover, as their actions are slower when using these input devices, it is easy for a bystander to learn what they input. For voice recognition inputs, a bystander can hear the password itself being spoken.

Thus, it is difficult for someone with physical disabilities to discretely enter secret information like passwords or PINs or credit card numbers. In this paper, a method to input secret information, especially passwords, for people with physical disabilities is

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1150–1154, 2019.
https://doi.org/10.1007/978-3-030-22868-2_77

fridaya@uef.fi

proposed. Specifically, we have developed a “breath switch” that can recognize human breath; a password can be input by exhaling for a long or short period of time.

2 Basic Idea

This paper presents the basic idea, which we proposed at SAI2018 [3] and SAI2016 [4]. The method based on this idea can be provided as a PC application which only requires the use of a few keys on the keyboard. The basic structure of the method proposed here is identical to the methods we proposed in [3] and [4], so we quote the basic usage below from our own papers.

An 8 × 8 random matrix table is shown to the user who wishes to input the password. This table consists of the characters a–z, A–Z, 0 (symbolized by “_” in this matrix to distinguish it from O and o), 1–9, and #$, as shown on the left in Fig. 1. Importantly, the sequence of characters is randomized each time.

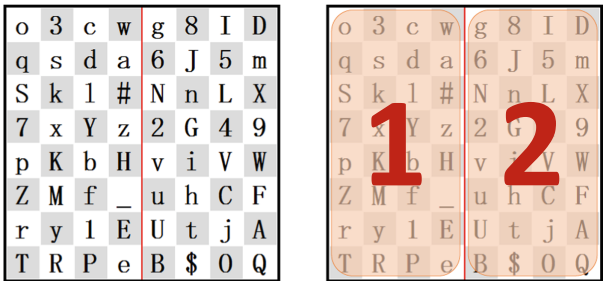


Fig. 1. First selection.

A red line divides the matrix table into left and right sections. The user needs to select the section which contains the letter they want to input. To select the section, the user needs to perform an action such as pressing a key corresponding to the desired section. Here, for the sake of convenience, the user press the key “1” to select the left section, and the key “2” to select the right section. In the example shown in Fig. 1, when the user wants to input “p,” which is the initial character of a password, they press the key that corresponds to the left section, i.e., the “1” key. Then, each section moves left or right, and spaces are created between them.

Then, each small section is subdivided into smaller subsections, as is shown in Fig. 2. The user selects the section in which the letter of their password to be input is present, and each subsection moves up and down and spaces are created between them. In this way, the sections are continuously reduced until the user isolates the letter of their password to be input.

Here, to input the first letter of their password “p,” the user’s button actions correspond to 1-2-1-1-1-1. A new 8 × 8 random matrix table is displayed on the screen so the user can input the next letter of the password. The sequence of the characters in the new matrix is not same as those in the previous matrix. When the user finishes their

fridaya@uef.fi

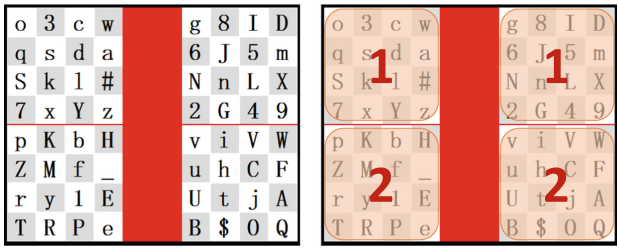


Fig. 2. Second selection.

input, they only need to perform a termination action such as pressing the Enter key to complete the authentication.

We described the method using the "1", "2", and "Enter" keys as an example for the sake of simplicity. Of course, three different input keys can be used; the only requirement is that the sections be selected via a two-key input method. In fact, we use the breath length for these choices as a two-key input device.

3 Breath Switch User Test

We adopted a breath switch as an input device to perform section selection. The breath switch (see Fig. 3) is a switch that can recognize human breath; it turns "ON" upon inhalation, meaning that any person can use this device as long as they breathe.

Fig. 3. Breath switch.

As breathing can be performed without sound, it is a secure form of input. We designed the breath switch to select either Sect. 1 or Sect. 2 based on the duration of the user's breath.

We conducted user tests on the switch by implementing the software for testing. The user tests were conducted as follows. These tests were also conducted in [3–5]. The conditions are similar.

– The subjects were six university students with no physical disabilities.

fridaya@uef.fi

– Passwords of four, eight, and twelve characters were created.
– The recall ability of a password affected the experiments. Therefore, the passwords were set as “kait” (four characters), “kaitpass” (eight characters), and “kaitpassword” (twelve characters).
– We measured the time it took to input a password by having the subjects input a password three times every alternative week.

The time between the initial display of the character matrix, and the pressing of the “Enter” key were measured to gauge the completion of input.

– The subjects were allowed to practice the inputting the password ten times before each test.
– If a mistake was made during password input, we let the user try again, and the input time for that test was not measured.

The average measured input time of the user tests is shown in Table 1. This method requires more time than the time taken by people without disabilities to input a password using a normal keyboard because we need to breathe six times to input each character of a password. Further, time is spent searching for the character in an 8×8 matrix. However, many subjects demonstrated increasing efficiency trends. Although the number of subjects were limited, we plan to obtain more data in the future.

Table 1. Average measured input time.

	4 characters	**8** characters	**12** characters
Average time (s)	111.87	168.32	157.04

4 Advantages and Future Work

This method can counter shoulder hacking, as it is difficult for a bystander who can see the display to guess the password simply from the screen, which does not reveal any hints. To verify the actual password guessing attack, we carried out a shoulder-hacking test. The subject input a letter of the alphabet using our method. After the input was made, a bystander tried to guess the input letter. Six pairs of students without disabilities participated in the test. As a result, no bystanders could correctly guess the input. However, it was necessary to train the subjects to breathe quietly so that there was no sound.

In this paper, we use a breath switch, but other, more silent assistive devices may be more effective as breathing creates some low sounds. The effectiveness of the proposed system can be confirmed using long-term user tests and attack tests. This will be addressed in a future study.

Acknowledgment. This work was supported by JSPS KAKENHI Grant Number JP17K00194.

fridaya@uef.fi

References

1. Microsoft Screen Keyboard: https://www.microsoft.com/ja-jp/enable/products/windows7/onscreenkeyboard.aspx. Accessed 29 Aug 2018
2. Soft Fit Mouse Stick: https://toksoamor-com.ssl-xserver.jp/32_427.html. Accessed 29 Aug 2018
3. Momose, R., Shibata, M., Okamoto, M.: Password input method using simple device. In: SAI Computing Conference (SAI) 2018, IEEE, London, UK (2018)
4. Ito, A., Kumazawa, Y., Okamoto, M.: Input password method for handicapped people. In: SAI Computing Conference (SAI) 2016, pp. 1306–1308. IEEE, London, UK (2016)
5. Okamoto, M.: Input password using simple input device. In: The Future Technologies Conference (FTC), Vancouver, Canada (2018)

fridaya@uef.fi

Don't Sweat It: Mobile Instruments for Clinical Diagnosis

Hellema Ibrahim[1,2(✉)] and Perry Xiao[1,2]

[1] London South Bank University, London, UK
hellema@protonmail.com
[2] Biox Systems Limited, London, UK

Abstract. This paper presents the development of a mobile closed chamber skin barrier function analyser for dermatological and cosmetic applications. Skin barrier function is evaluated by the measurement of Transepidermal Water Loss (TEWL). A closed chamber analyser design was chosen for its improved ability to limit environmental interference compared to open chamber devices. Theoretical and practical experiments are presented to demonstrate different methods of TEWL computations. Finally, the accuracy of the device is compared with existing instruments to highlight the importance of TEWL calculation optimisation. Calibration studies are ongoing.

Keywords: Instrument · Dermatology · TEWL · Trans-epidermal water loss · Closed-chamber · Diffusion

1 Purpose

The skin is the largest organ of the human body. It is responsible for thermoregulation and permits the sensation of touch, as well as protecting against dehydration, poison and microbes. Multiple non-invasive methods have been investigated as a surrogate marker for skin barrier function, including imaging of skin water distribution, measurement of hydration, surface pH level, and skin surface lipids [1]. Transepidermal Water Loss (TEWL) has been a recognised method for quantifying skin barrier function for over fifty years. TEWL is defined as evaporation of water through the stratum corneum (SC) of the skin [2]. TEWL measurements are used in dermatological applications such as the formal study of the irritancy of chemicals, susceptibility for skin irritation and contact dermatitis [3,4]. A reduction in the integrity of the skin barrier is also linked to a number of skin diseases and food allergies. There is therefore a demand for an objective, quantitative, measurement-based means of diagnosing human skin diseases using TEWL. At present, advances in fundamental and applied research are creating powerful diagnostic tools potentially capable of this function. However, there is a technology gap between measurements that work well in the laboratory setting and diagnostic devices capable of being used in clinical settings that are usually less well controlled environments. The aim of

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1155–1169, 2019.
https://doi.org/10.1007/978-3-030-22868-2_78

fridaya@uef.fi

this project is to develop a mobile version of the Aquaflux, an instrument to measure TEWL using the condenser chamber method, to be used for dermatological and cosmetic applications. Requirements for such a device are the inclusion of an embedded user interface, the ability to act as an independent device, being able to perform consistently in an uncontrolled environment, and being energy efficient. This article focuses on our current prototype, aiming to optimise the gradient calculation within a closed, non-ventilated chamber.

2 Background/Significance

The measure of TEWL can be performed with an open-chamber, closed unventilated chamber or condenser chamber. Open chambers measure the gradient of humidity between two sensors. They rely on Fick's first law of diffusion:

$$J = -D \times \frac{\partial C}{\partial y} \tag{1}$$

with D representing the coefficient of diffusion in the medium. The main limitation of open chambers is their sensitivity to the environment. Indeed, their measurement can be disrupted by room air currents as minute as the opening of a door, window or someone breathing [5].

Closed chamber devices measure the gradient of humidity between two time points. This gradient can be linearised for a short period of time [6]. Limitations of this linearisation are the ability of the air to accept the water diffusion from the skin if the air is not saturated, and only being able to work within the operating conditions of the sensor. To perform linearisation, one must set t: time, k: slope of the rising humidity during measurement, and C: the concentration of water in air, or humidity. The concentration of water in air is also the mass in the chamber volume:

$$C = C_{i_{n_i}} + k \times t \tag{2}$$

where k is proportional to the TEWL. k depends on measurement conditions such as Skin Surface Water Loss (SSWL), the ambient temperature and humidity.

In the condenser chamber method, a condenser is a peltier-cooled surface at the far end of the chamber and removes water vapour by freezing it. Mathematical modelling of this chamber can be found in [7].

Previous publications have compared the three methodologies. Elkeeb et al. [8] found measurement capacity with the Aquaflux condenser chamber to be significantly higher than the Tewameter TM210 open chamber Vapometer closed chamber. The same observation was noted by Farahmand et al. [9], that the condenser chamber was more sensitive than the open chamber Tewameter TM210 or closed chamber Vapometer, as well as being able to demonstrate tape stripping effect. According to Farahmand [9], the Aquaflux condenser chamber showed a strong correlation with the open chamber Tewameter TM210 and the closed chamber Vapometer SWL-2. The difference in absolute measurement values was likely due to varying calibration methods. In conclusion, all of these instruments

fridaya@uef.fi

have advantages and disadvantages. Some measurement probes are portable with independent wifi or bluetooth functions. However, as all skin measurements require follow up measurements in laboratory settings, current mobile devices are not truly 'mobile'.

3 Method

3.1 Development Method

A closed chamber was chosen for the prototype because TEWL measurement is less impacted by the environment using this chamber design. Development of a portable instrument was made using waterfall project methodology (Fig. 1). After testing the prototype, a new version was developed using waterfall and incremental model processes. Thus, this engineering project used spiral model methodology. Each work package used a traditional V model of developing and testing the software/hardware units prior to integration. When design questions arose, experimental tests and simulations were used. Experimental tests were used for the Relative Humidity (RH) values near the skin. Simulations were run on ANSYS and compared to practical experimental results. Verification was done by calibration and comparison with the existing non-mobile Aquaflux. Practical experiments were conducted to test the prototypes against the simulations. These were considered integral to prototype design to give real-life validity to the measurements and to identify otherwise unaccounted for confounding variables. Development of the prototype is ongoing.

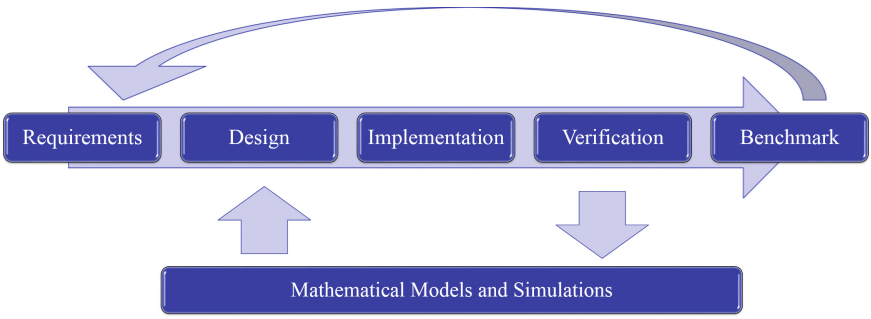


Fig. 1. Development methodology

3.2 Measurement Procedure

TEWL guidelines [3] provide strict measurement methods. For instance, the participant's TEWL can be affected by stress, exercise, spicy food and sweating. Skin measurements were performed in triplicate at the same site for consistency (further increased frequencies of 5–6 measurements were trialled but did not

fridaya@uef.fi

lead to improved consistency). Flux of water through the skin (i.e. TEWL) was measured at two anatomical sites, the volar forearm and palm (representing the minimum and maximum expected TEWL regions respectively) [10]. Whilst other sites may be used (e.g. the cheek or the leg), it is preferable to choose a site that is easily accessible at any time of the year that is acceptable to participants. To improve reproducibility and accuracy, a mark was drawn on the skin around the position of the chamber to ensure identical probe placement on repeat measurements.

In order to compare with the simulations, the orientation of the probe was accounted for. The measurements are performed as per Fig. 2.

Fig. 2. How to perform TEWL measurements

3.3 Data Collection

A probe application time of 30 s was chosen because quick measurements would be necessary for any prototype to be translated into practical use. In order to limit the impact of the environment, a quick measurement is ideal. Although measurements of one or two minutes offer improved stabilisation of readings, the chamber needs to rest for longer in order to come back to pre-reading conditions. Whilst gradient calculation could potentially be better performed using two sensors to calculate a space gradient, a first step is to model using a single sensor to assess accuracy and feasibility. Also, several sensors on I2C bus protocol would likely require significantly more processing.

Data is saved onto an SD card as this was the quickest possible data capture method (every 0.02 s). The data is then analysed on Matlab for calculation optimisation. It is sampled every 50 items when compared to simulations with a 1 s timestep.

4 Results

4.1 Prototype Presentation

The prototype was designed to be big enough to accommodate a Relative Humidity & Temperature (RHT) sensor (SHT31 from Sensirion). A typical chamber is

fridaya@uef.fi

10 mm diameter and 20 mm length [11]. In our prototype, the chamber is 15 mm in diameter and 17 mm in depth. The chamber is at the extremity of the probe and doesn't interfere with the closing mechanism of the casing. This necessitates having the chamber on a cap, separate from the rest of the device casing (Figs. 3 and 4)

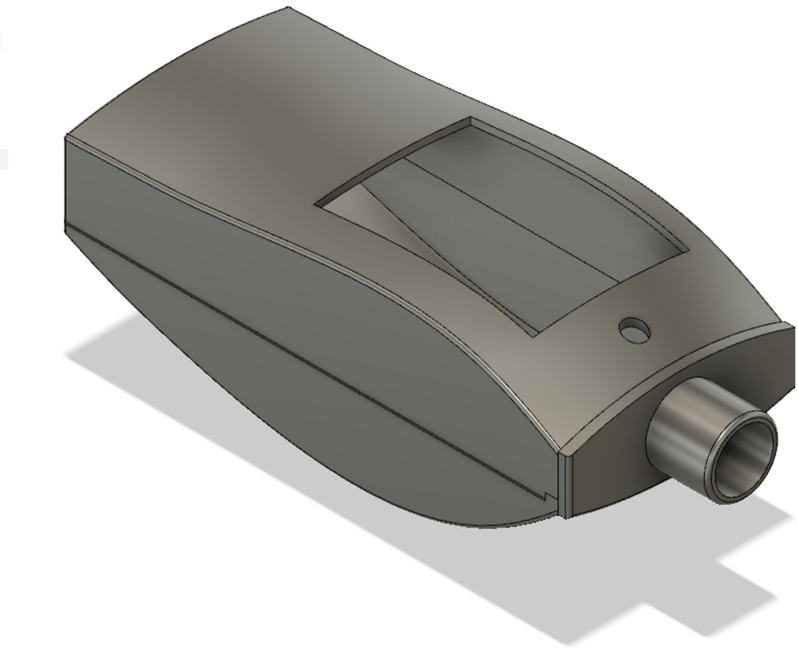


Fig. 3. Chamber 7.1 casing design

The casing is large enough to contain the Printed Circuit Board (PCB); 130 mm × 52 mm. The instrument was designed with Fusion 360 and 3D printed by extrusion with a RepRap Anet A8. Due to constraints from the 3D printer, the smallest surface thickness requirement was 1.5 mm. Due to the requirement of the closing mechanism, the current prototype casing is larger than could theoretically be achieved.

There is a touch screen to observe TEWL measurement in real time. This allows the user to see if the measurement is occurring as anticipated or if the chamber has reached saturation. In addition, the prototype has an SD card and bluetooth capability in order to save the measurements and to monitor from a BLE capable device. BLE has been chosen for energy saving. On the chamber, which acts as a Generic Attribute (GATT) server, the characteristics are RH and temperature sent as double precision numbers, as well as the battery level. A chamber reader was also developed for Android operating system. The GATT client can read these characteristics, subscribe to them, and be informed of any

fridaya@uef.fi

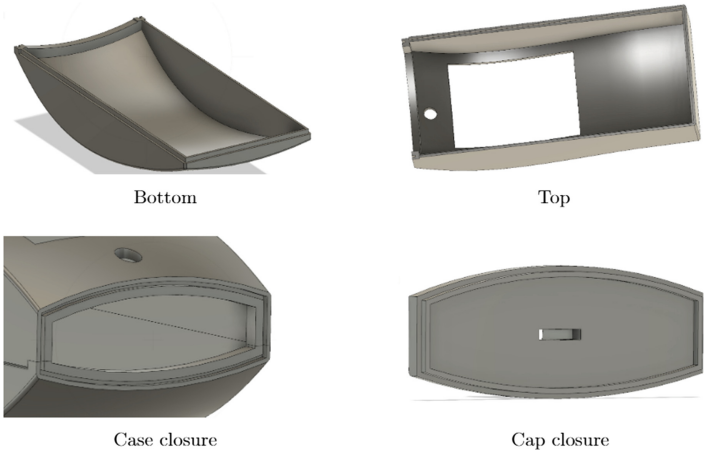


Fig. 4. Chamber 7.1 casing closure design with 3 parts: top, bottom, cap

changes. This prototype uses a tinyTILE microprocessor. The battery is a 3.7 V LiPo battery. Autonomy tests were conducted with two LiPo batteries of the same model and the chamber lasted 21h10 and 20h30 respectively (Table 1).

Table 1. Autonomy comparison of 2 batteries

Battery 1	Battery 2
20h30	21h10
19h50	21h10
21h	
20h45	
Mean = 20h30	21h10

4.2 Measurements

In our measurement protocol, the skin was acclimatised for one hour in a climate-controlled room with a temperature of 20–21 °C and RH of 45–46%.

TEWL measurements on the palm are particularly difficult due to the abundance of sweat glands. Sweat interferes with TEWL: it causes SSWL that affects the measurements. This can manifest as a steep early slope on the humidity measurement. To reduce the flux, especially for the palm, we use a small cap on the chamber (see Fig. 5). We also measure the RHT on the VF with the small cap. Unfortunately, measurements were not as consistent for the VF using the small cap. The results presented are for a surface area of $1.88692 \times 10^{-3} \mathrm{m}^2$. The TEWL is also measured at the same location site, in triplicate, to reduce skin and instrument errors.

fridaya@uef.fi

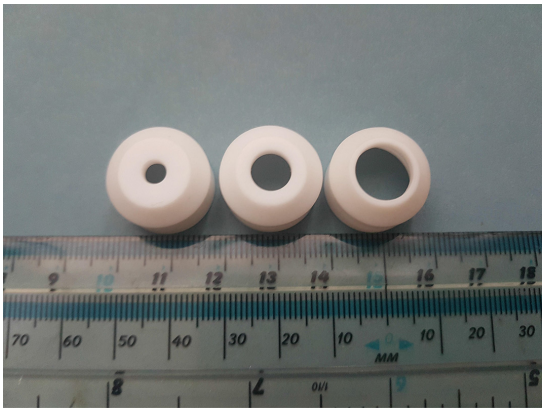


Fig. 5. Caps for different flux

The temperature slowly increased during the measurement session (Fig. 6). For the VF the temperature increased by 2 °C. Therefore the relative humidity decreased by 4%. This could be accounted for by room temperature variation. It

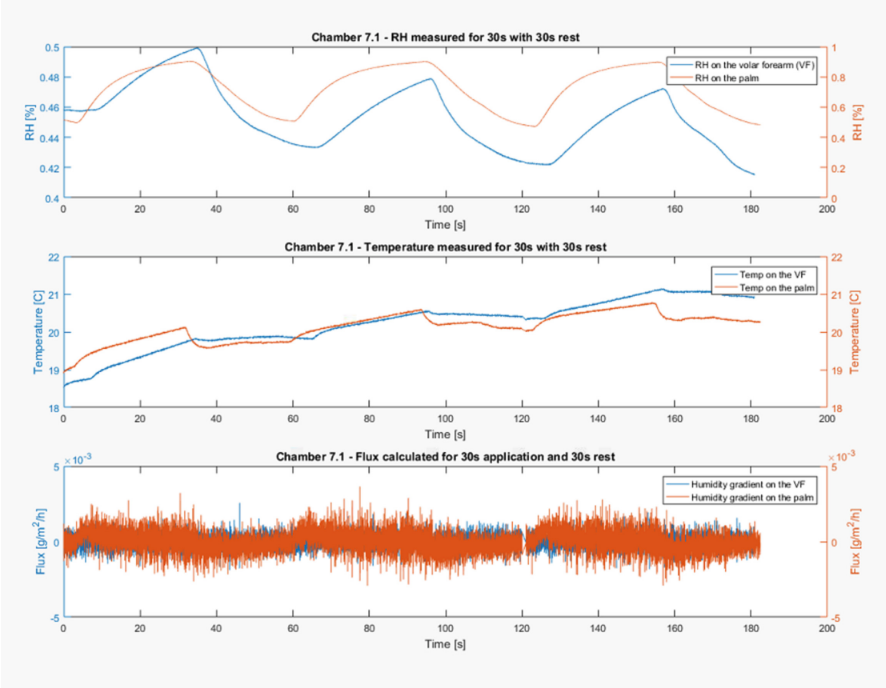


Fig. 6. Three measurements on the VF and the palm: RHT and gradient of humidity

fridaya@uef.fi

could also come from the zero-drift effect, when the chamber doesn't come back to pre-measurement conditions between measurements. It is usually observed for RH, and guidelines [3] recommend shaking the instruments to remove condensation. The measurement started after a seven second delay. The application of the skin on the chamber is clearly visible at 7, 67 and 127 and 5, 65, 125 s for the VF and the palm respectively: temperature and RH increase. The slope is more important for the VF than the palm. For the palm the humidity reached saturation (80% RH).

On Fig. 6, the gradient is calculated with the previous measurements and a measurement is performed every 0.02 s (which accounts for the high variability observed). Sampling the gradient every second leads to less noise. However, this leads to data loss. The larger the number of occurrences, the smoother the gradient (see Fig. 7).

4.3 Optimisation with Noise Filter

We use moving average to have stabilised measurements to remove background noise. The larger the window, the better the gradient (see Figs. 8 and 9). Using windows of 1, 2 or 3 s, results are less noisy.

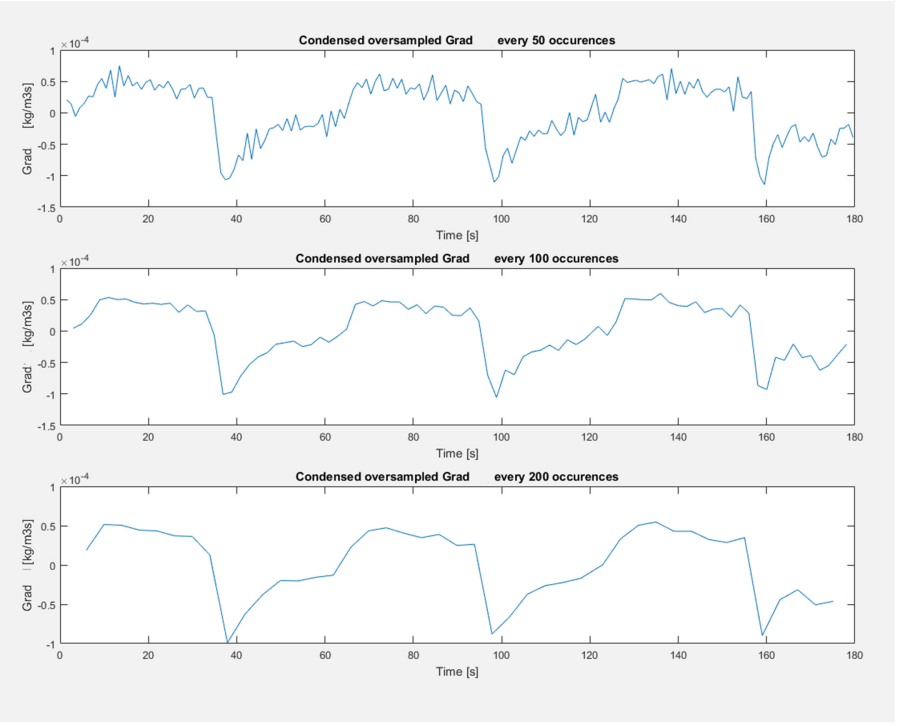


Fig. 7. Gradient of humidity sampled at different frequencies

fridaya@uef.fi

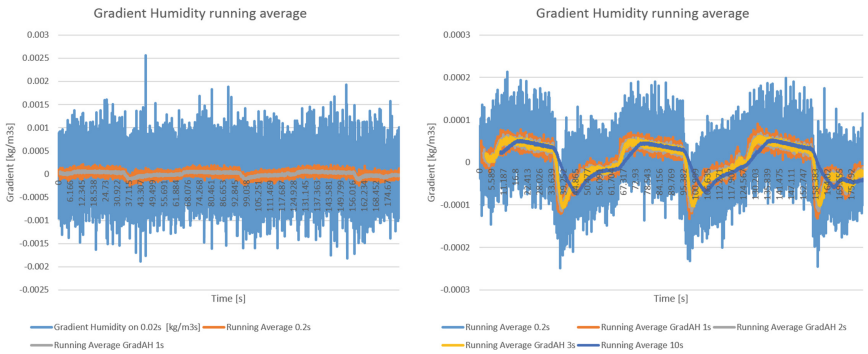


Fig. 8. Comparison gradient humidity and running average on 0.2 s, 1 s, 2 s, 3 s, 10 s for measurements of the volar forearm

For Savitzky-Golay filter, order three and frame length 11 were chosen first, then order 4 and length 45, and lastly order 5 frame length 99. This last result provided the best results as it allowed the observation of variations of gradient. An increase was clearly observed at 7 s, 67 s and 127 s corresponding to the

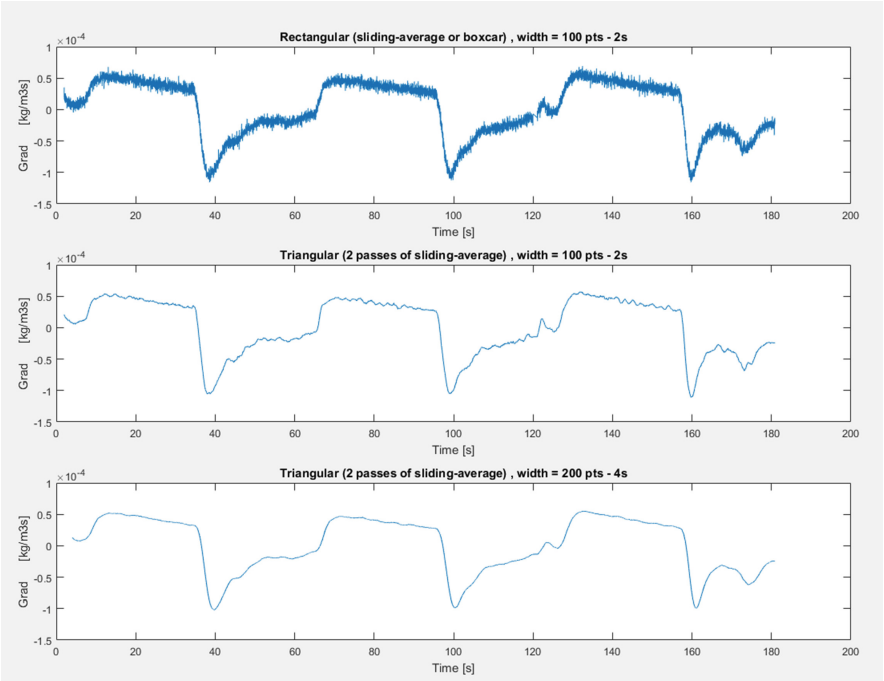


Fig. 9. Sliding average: rectangular box vs triangular (weighted)

application of the chamber on the skin (see Fig. 10). A decrease of gradient corresponding to the rest phase, and removal of the influx from the skin, was also observed.

We chose the same window to compare these approaches, but further study is needed to find the computational complexity for each filter depending on their implementation. The juxtaposition of the smoothing methods can be seen on Fig. 11. Rectangular and triangular averages create a delay to the size of the window. Sampling was not as smooth as the other filters. If the information is needed in real-time, Savitzky Golay provides the necessary smoothing without delay.

Filtering is optimal for visualisation but doesn't improve the gradient calculation.

4.4 Gradient Calculation Optimisation

It is recommended to use the sensor until 80% RH, otherwise the phenomenon of hysteresis is observed. A long exposure to higher humidity may temporarily offset the RH signal (+3% RH after 60 h kept at >80% RH) and decrease the life of the sensor. Therefore, the gradient calculation of the palm can be calculated with the slope only in this linear part at the beginning of the measurement until 15.5 s. This is the time region where (2) is applicable.

The humidity gradient is calculated by deriving humidity according to time. First, the script lines up the sequences of measurements and removes the beginning of the recording if there was a delay. Then the script finds the minimum of the curve and calculates the gradient by varying its length: delta. The results are visible on Fig. 12 with comparison to Aquaflux measurements.

It is obvious on Fig. 12 that for a high TEWL, e.g. on the palm, the instrument measures with significant errors, likely due to SSWL from e.g. sweat. For problematic sites near sweat glands like the palm, it is better to get a gradient calculated on a longer delta time to get a better sensitivity.

A factor affecting TEWL needing study is the effect of the timing of measurement. The microclimate of the chamber is disrupted at the beginning of a measurement because of the application of the chamber on the skin. However, the gradient estimation is possible only when the slope is in the linear phase, i.e. when $delta + starttime < 15.5$ s (see Fig. 6). The error on gradient computations compared to Aquaflux results has been plotted against time when the gradient calculation starts on the palm and the VF respectively (Fig. 13). The error is the ratio between the standard deviation and the mean values.

The larger the gradient (in terms of length delta), the less it is impacted by the time it starts (start time on Fig. 13). A good solution is to get a gradient calculated on a longer delta time on non problematic sites like the VF, or to delay the start time of measurements. Overall, for the palm and the VF, the smaller the delta, the later the measurement should be commenced.

To get an appropriate TEWL measurement, a calibration study is required. This work is currently ongoing.

fridaya@uef.fi

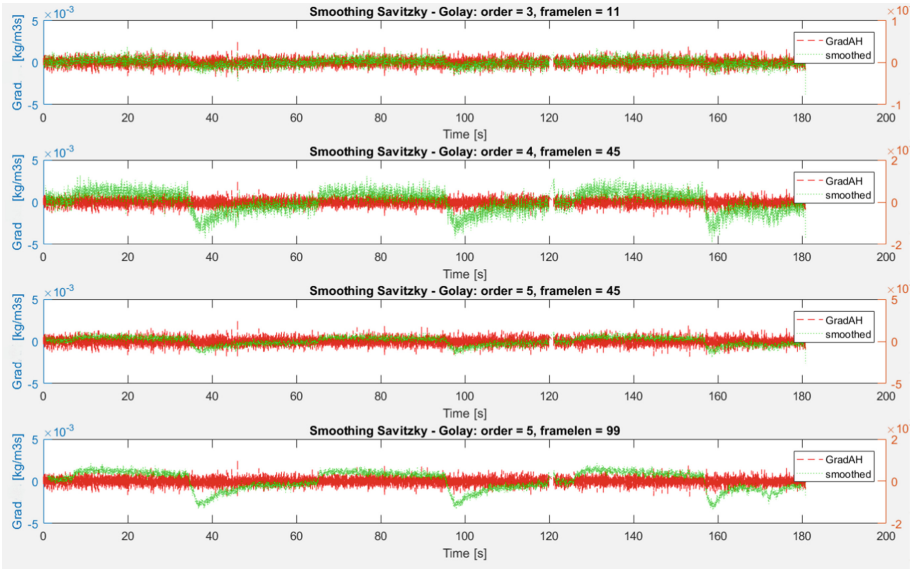


Fig. 10. Using Savitzky Golay Filter and various order and frame length

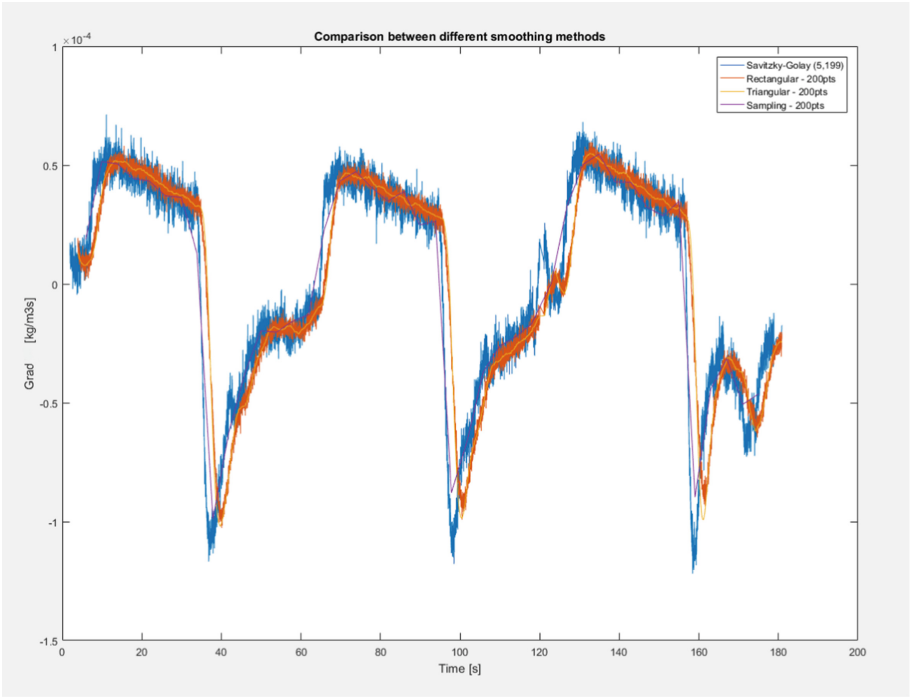


Fig. 11. Comparison of methods: Savitsky Golay (order 5 and window 199), rectangular and triangular window (width 200), and sampling every 200 points

4.5 Comparison with Simulations

The simulations have been done with Ansys Workbench. In Ansys Fluent, a transient model was set with gravity, laminar flow, energy equation and species transport. The boundary condition was mass flow inlet for the skin with water mass fraction (calculated from RH initial in the experiment) and temperature. Succeeding the temperature, the water mass fraction, the RH, the velocity were simulated in the chamber until 30 s, i.e. the end of the measurement.

A gradient of temperature was observed coming from the skin (on top) (see Fig. 14 top left). Temperature was initially 291.65 K and increased slightly to 291.8 K after 30 s. In Fluent, water mass fraction is set up instead of RH. This is similar to RH but not quite as intuitive, see Fig. 14 top right. The relative humidity gradient at 30 s showed a slight gradient of humidity, higher near the skin (see Fig. 14 bottom right). According to the simulations (Fig. 14 bottom left), the sensor should be closer to the skin to get higher flux quickly, but the

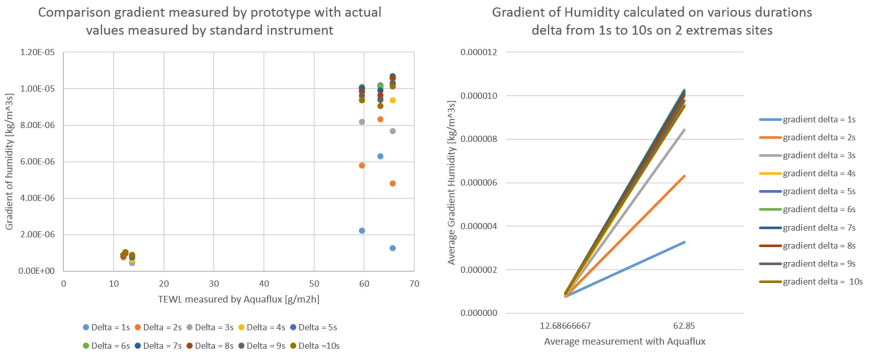


Fig. 12. Gradient of humidity compared with Aquaflux for various delta, gradient's length

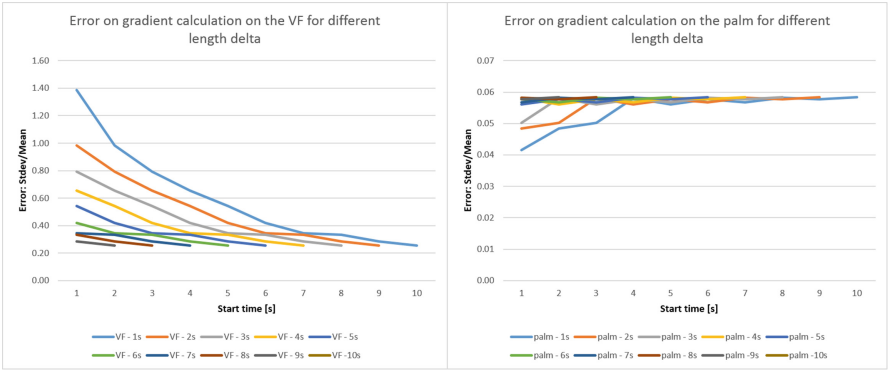


Fig. 13. Error on the gradient of humidity on the VF (left) and palm (right) for various length (delta) and start time of gradient computation

fridaya@uef.fi

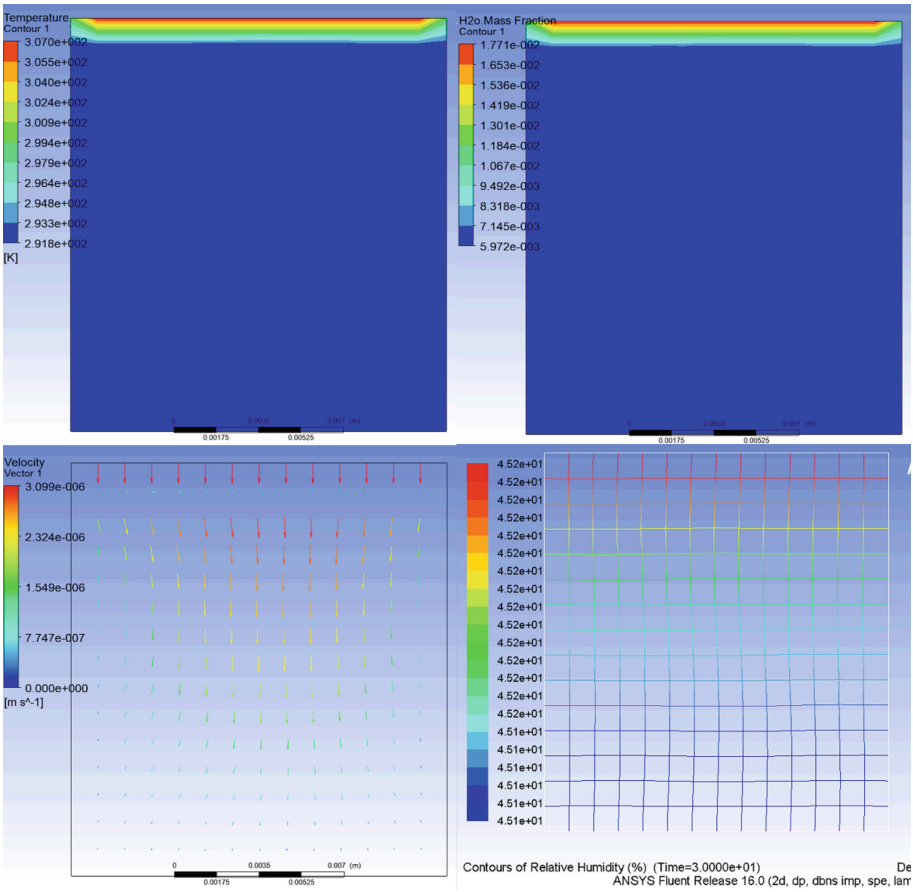


Fig. 14. Simulations results after 30 s in Fluent and in CFD-Post: temperature (top left), water mass fraction (top right), RH (bottom right) contours and velocity vector (bottom left)

microclimate would be disrupted when the chamber is applied. One limitation of this model is that there is no User Defined Function input at the initialisation of Fluent to describe this behavior.

The measurements being compared with are from the VF because it's a site known for easy and reproducible results. As seen on Fig. 12, the gradient measured on the palm is highly variable.

For quick results, the simulation time step is 1 s with 85 iterations per time step. The results from the practical experiments are sampled every 50 items for comparison with the simulation (see Fig. 15 for comparison of RHT). The RH and temperature measured increase more than expected from the simulations, but the difference was negligible. In the practical experiments RH increase is 2%, which is equal to the error accuracy of the sensor. Temperature increased by 1 °C,

fridaya@uef.fi

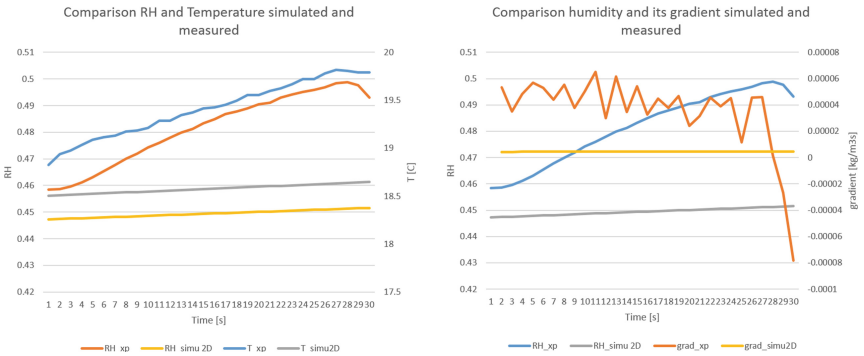


Fig. 15. Comparison measurements and simulations: RH and temperature compared to measured values (left) and humidity and its gradient (right)

which is larger than the 0.3 °C accuracy error. This discrepancy correlates with the increase of temperature during the measurement session.

For the practical experiment, the gradient was recalculated from the sampled results without any filters (see Fig. 15 on the right). The gradient was perfectly constant in the simulation. Without any smoothing the measured flux oscillates, demonstrating the importance of gradient optimisation. The mean value of the measured gradient was slightly higher than the expected.

5 Conclusions

Gradient calculation optimisation is essential. It is better to get a gradient calculated on a longer delta time for problematic sites near sweat glands such as the palm, but also for sites with good reproducibility such as the VF. Other important factors include the error of the SHT31 sensor and the need for having a running average. The simulations help in prototype design and demonstrate similarity to practical TEWL measurements. During any skin measurement, it is important to control the environment conditions. However, whilst 21 °C is the ideal temperature for TEWL measurement, trying to artificially maintain such a temperature can cause more variation through the use of cyclical air conditioning, varying air flows, and resultant temperature changes.

5.1 Other Optimisations: Future Work

The results presented are raw measurements and as such are not calibrated. A calibration study is ongoing. Further modelling will be beneficial to determine the optimal geometry.

fridaya@uef.fi

6 Acronyms

BLE Bluetooth Low Energy
GATT Bluetooth Generic ATTributes
LiPo Lithium-ion Polymer
PCB Printed Circuit Board
RH Relative Humidity
RHT Couple Relative Humidity and Temperature
SC Stratum Corneum
SD Secure Digital
SSWL Skin Surface Water Loss
TEWL Trans Epidermal Water Loss
VF Volar Forearm

Acknowledgments. We would like to thank London South Bank University and Biox Systems Ltd. for providing the joint PhD studentship for H.A.L. Ibrahim.

References

1. Darlenski, R., Sassning, S., Tsankov, N., Fluhr, J.W.: Non-invasive in vivo methods for investigation of the skin barrier physical properties. Eur. J. Pharm. Biopharm. **72**(2), 295–303 (2009)
2. Imhof, B., McFeat, G.: Evaluation of the barrier function of skin using transepidermal water loss (TEWL). In: Barel, A.O., Paye, M., Maibach, H.I. (eds.) Handbook of Cosmetic Science and Technology, pp. 131–136. CRC Press (2014)
3. Rogiers, V.: EEMCO guidance for the assessment of transepidermal water loss in cosmetic sciences. Skin Pharmacol. Physiol. **14**(2), 117–128 (2001). https://doi.org/10.1159/000056341
4. Seidenari, S., Giusti, F.: Transepidermal water loss and allergic contact dermatitis. In: Fluhr, J., Elsner, P., Berardesca, E., Maibach, H.I. (eds.) Bioengineering of the Skin: Water and Stratum Corneum, 2nd edn., pp. 105–112. CRC Press (2005)
5. Miteva, M., Richter, S., Elsner, P., Fluhr, J.W.: Approaches for optimizing the calibration standard of Tewameter TM 300. Exp. Dermatol. **15**(11), 904–912 (2006)
6. Nuutinen, J., Alanen, E., Autio, P., Lahtinen, M.-R., Harvima, I., Lahtinen, T.: A closed unventilated chamber for the measurement of transepidermal water loss. Skin Res. Technol. **9**(2), 85–89 (2013)
7. Cui, Y., Xiao, P., Ciortea, L.I., De Jesus, M.E.P., Berg, E.P., Imhof, R.E.: Mathematical modelling for the condenser method of trans-epidermal water loss measurements. Nondestr. Test. Eval. **22**(4), 229–238 (2007)
8. Elkeeb, R., Hui, X., Chan, H., Tian, L., Maibach, H.I.: A closed unventilated chamber for the measurement of transepidermal water loss. Skin Res. Technol. **16**(1), 9–15 (2010)
9. Farahmand, S., Tien, L., Hui, X., Maibach, H.I.: Measuring transepidermal water loss: a comparative in vivo study of condenser-chamber, unventilated-chamber and open-chamber systems. Skin Res. Technol. **15**(4), 392–398 (2009)
10. Mayrovitz, H.N., Bernal, M., Brlit, F., Desfor, R.: Biophysical measures of skin tissue water: variations within and among anatomical sites and correlations between measures. Skin Res. Technol. **19**(1), 47–54 (2013)
11. Imhof, R.E.: In-vivo and in-vitro applications of closed-chamber TEWL measurements. In: Lecture Notes: Intensive Course in Dermato-Cosmetic Science (2007)

fridaya@uef.fi

Technical Analysis Toolkit for Neural Networks in Finance and Investing

Chun-Yu Liu, Shu-Nung Yao(✉), and Ying-Jen Chen

National Taipei University, San Shia District, 23741 New Taipei City, Taiwan
snyao@gm.ntpu.edu.tw

Abstract. Stock market forecasting is an important issue in financial engineering. With the rapid growth of stock market data, artificial intelligence in finance has the potential to provide investors with accurate strategy analysis and effective investment decisions. This study combines the proposed technical analysis toolkit with back propagation neural networks. Many studies discuss how to modify the algorithm of machine learning, but few studies investigate how to achieve good data pre-processing. The proposed technical analysis toolkit provides several useful functions such as stock price analysis, forecasting, and data pre-processing. We use analyzed data rather than raw data as the input variables for the neural networks and verify whether the data pre-processing can provide more accurate stock price predictions. The technical analysis indicator package was written in the R language. The sample data include several major indexes in the U.S. stock market, such as the Dow Jones Industrial Average, Philadelphia Semiconductor Index, Standard & Poor's 500 Index, and NASDAQ Composite Index. The experiment results indicate that the proposed toolkit helped the neural networks achieve accurate predictions. The proposed toolkit passed a comprehensive R archive network (CRAN) check and contributes to the field of stock data analysis.

Keywords: Artificial intelligence · Data pre-processing · Stock market forecasting

1 Purpose

As the technological era advances, investors can rely on an increasing number of analytical methods and tools in their investment decision making. However, stock prices often fluctuate due to various political, human, and economic factors, as well as unexpected events and other elements of uncertainty. This makes stock price prediction rather challenging. As such, most modern stock investment methods rely on investors' personal experiences and subjectivity during the decision-making process. Furthermore, the current methods for analyzing and predicting stock prices are neither accurate nor stable enough.

This study uses technical indicators to analyze historical stock price data and obtain information on the trends of price bands, over-buying, and over-selling, among others. The developed toolkit can automatically find the technical indicators in the stock markets. Such information can serve as a reference for investors and become an input

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1170–1174, 2019.
https://doi.org/10.1007/978-3-030-22868-2_79

fridaya@uef.fi

for a neural network. In this study, we verify whether the technical analysis toolkit can aid with stock price predictions in a neural network.

2 Background and Significance of the Proposed Toolkit

This section introduces the background of technical indicators and shows several examples of technical analysis by using our proposed toolkit [1].

2.1 Trend Lines

Trends on a price chart indicate the direction in which the market is moving and can provide investors with ideal profit opportunities. For an upward trend, the troughs can often be connected with a straight line, as in Fig. 1(a), and the same applies to the peaks in a downward trend, as in Fig. 1(b). These straight lines are the trend lines. Forming trend lines is a fundamental technique and crucial tool for graph analysis: they connect either the highest points or the lowest points of each curve. Here, we can identify trends such as an "uptrend" or a "downtrend," and can categorize trend cycles as long-term (from six to 12 months), medium-term (from one to three months), or short-term (a week). The longer the cycle covered by a trend line, the greater the impact of its support or resistance.

Fig. 1. Examples of trend lines: (a) uptrend and (b) downtrend.

Fig. 2. Examples of V-type reversal patterns: (a) V-type pattern and (b) inverted V-type pattern.

fridaya@uef.fi

2.2 V-Type Reversal Patterns

The V-type reversal pattern represents extremely strong stock market signals. They appear frequently during market fluctuations, where a one-off reversal peaks or troughs occur at the bottom or top price ranges. This indicates that the stock price will undergo extreme changes trending in the opposite direction. As Fig. 2(a) shows, the stock price soared after a V-type pattern, whereas the opposite occurred after an inverted V-type pattern, as Fig. 2(b) shows.

2.3 W-Type Reversal Patterns

Due to the internal and external variables that affect prices, each type of commodity encounters bottlenecks or stagnates at some point. In this case, investors are more likely to change their views on commodity prices. The prices will reflect this change in attitudes, thereby forming a variety of patterns, each of which is a more complicated reversal pattern than the V-type pattern. W-type reversal patterns form during reversing stock price trends. In other words, they represent signals from the shifts in trend lines, such as from an uptrend to a downtrend, or vice versa. As such, investors can use trend graphs to predict future price trends. The M-type pattern in Fig. 3(a) tends to appear before a depression, whilst the W-type patterns in Fig. 3(b) are more likely to occur before a boom.

(a) (b)

Fig. 3. Examples of (a) an M-type pattern and (b) W-type patterns.

(a) (b)

Fig. 4. Examples of (a) bull power and (b) bear power.

fridaya@uef.fi

2.4 Bull Power and Bear Power

In the ever-changing stock market, the orientation of daily stock market movements fluctuates according to factors such as changes in the financial market, capital flows, political factors, or even periodic changes in the economic cycles. A stock market in a long-term upward trend due to long-term economic growth is called "bull power" or a "bull market." Economies and stock prices showing a long-term decline are called "bear power" or a "bear market." The bull and bear market indicators proposed in this study use the relative strength index (RSI) to identify the strengths or weaknesses of stock prices. We screen and analyze the data, and use the RSI to determine buying and selling levels. Figure 4(a) and (b) show examples of bull and bear power, respectively.

3 Method

The toolkit for both the technical analysis and the corresponding user guide are available for download from [1]. Therefore, we discuss only the application of processed data here. We designed several methods for pre-processing to check if the proposed toolkit can actually improve predictions. All scenarios have the same neural network architecture. The input layer contains three neurons to collect prices three days prior, the hidden layer contains 21 neurons, and the output layer contains one neuron to predict the future price.

We downloaded the stock prices from YAHOO FINANCE for January 4, 2007 to April 6, 2016. The data for January 4, 2007 to February 23, 2016 serve as the training data. After pre-processing, we fed the data into the neural network for training. The test data are from February 24, 2016 to April 6, 2016. The input features are the stock prices three days prior, and then the neural network predicts the future price.

4 Results

We designed and compared four pre-processing methods. In method A, the input data were the raw data. In method B, the raw data were low-pass filtered for smoothing purposes. We used a third-order Butterworth filter with a cut-off frequency of 1.885 rad/s for the digital filter. In method C, we analyzed the raw data using our proposed toolkit and then fed these data into the neural network. Method D combines methods B and C. The raw data were the treated data from method C that we then filtered using the low-pass filter in method B. We evaluated the accuracy using the mean square error (MSE):

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(a(i) - p(i))^2 \tag{1}$$

where N is the total number of days in the test data, $a(i)$ is the actual stock price on the i^{th} day, and $p(i)$ is the predicted price. In the pilot experiment, we examined several major markets: the Dow Jones Industrial Average (^DJI), PHLX Semiconductor (^SOX), NASDAQ Composite (^IXIC), and S&P 500 (^GSPC).

fridaya@uef.fi

Table 1 reports the MSEs for the different pre-processing models. The results indicate that method D tended to obtain the best performance. If we feed only raw stock prices into the neural network (method A), the prediction is less likely to be accurate. The results also indicate that the smoothing in method B can somewhat filter out the noise, thereby providing the tendency. Although method C performs similarly to method D, method C is not as good as method D.

Table 1. MSEs for each pre-processing model.

Pre-processing	^DJI	^SOX	^IXIC	^GSPC
Method A	15453.2468	49.7185	1835.0309	229.6357
Method B	15136.9070	48.1788	1819.0279	211.8701
Method C	13989.0779	46.6690	1840.9047	221.0585
Method D	13418.1936	43.5215	1796.3154	209.8973

5 Conclusion

We proposed a technical analysis toolkit for stock markets and applied the toolkit to a neural network to forecast financial data. The proposed technical analysis toolkit contains four technical indicators: trend lines, V-type reversal patterns, W-type reversal patterns, and bull and bear powers. The toolkit automatically finds the technical indicators after the financial data are fed into the toolkit. Therefore, the technical analysis is not based on an individual's personal impressions, feelings, or opinions. In the pilot experiment, we showed that data pre-processing is an important step in the data mining process. Both smoothing and technical analysis improved the accuracy of the predictions.

Acknowledgments. This work was supported in part by the Ministry of Science and Technology in Taiwan under Grant MOST 107-2221-E-305-010-MY2.

Reference

1. Liu, C.-Y., Yao, S.-N.: lcyanalysis: Stock Data Analysis Function, R Package Version 1.0.3, March 2018. https://CRAN.R-project.org/package=lcyanalysis

fridaya@uef.fi

Computer Graphics Based Analysis of Loading Patterns in the Anterior Cruciate Ligament of the Human Knee

Ahmed Imran(✉)

Department of Biomedical Engineering, Ajman University,
PO Box 346, Ajman, UAE
a.imran@ajman.ac.ae, ai_imran@yahoo.com

Abstract. Injuries of anterior cruciate ligament of the knee are common, particularly in young athletes. Though surgical reconstruction of the ligament attempts to restore the joint function, a significant percentage of the patients report unsatisfactory outcome and joint complications that may require repeated surgery. The present study used a computer-based model of the knee with intact ligaments and anatomical articular surfaces to visualize contributions of different fiber bundles in the ligament with distinct areas of attachment on the femoral bone that are searched for appropriate positions of femoral tunnel during single or double bundle ligament reconstruction. Knee motion during flexion and an anterior drawer test at different joint positions were simulated in the sagittal plane. The model analysis showed that the ligament fibers attached anteriorly on the femoral bone contributed significantly throughout the knee motion and resisted anterior loads on the tibia at all flexion positions, while the fibers attached posteriorly on the femoral bone contributed during 0–45° and above 90° flexion. The results agreed with experimental observations and have clinical relevance.

Keywords: Anterior cruciate ligament · ACL reconstruction · Double bundle reconstruction of ACL · Femoral tunnel position in ACL reconstruction

1 Introduction

Injuries of Anterior Cruciate Ligament (ACL) of the knee are common especially among young athletes. The ligament is frequently damaged while performing strenuous activities such as those in sports. Such injuries are often treated with surgical reconstruction of the ligament with an aim to restore joint stability and kinematics as well as to contain subsequent long-term risk of complications like osteoarthritic changes, chondral or meniscal damage [1, 2]. Though long-term satisfactory outcomes have been reported, about one-third patients remain unsatisfied with difficulty in returning to previous level of activity [1, 2]. One clinical study reported that less than 50% of athletes with ACL reconstruction were able to return to their pre-injury level activity [2]. Interestingly, another clinical study showed that 94% of patients from ligament surgery continued to have knee instability even after a five-year follow-up [3].

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1175–1180, 2019.
https://doi.org/10.1007/978-3-030-22868-2_80

fridaya@uef.fi

The anterior and posterior cruciate ligaments or ACL and PCL, respectively, are considered as the main stabilizers of the joint in the sagittal plane [4–7]. While the ACL restricts anterior translation, the PCL restricts posterior translation of the lower bone, or tibia, relative to the upper bone, or femur. Common understanding of the ACL through anatomic and histologic studies suggests two functional bundles of ligament fibers with distinct attachments on the femoral and tibial bones [8–11]. Accordingly, recent surgical techniques of ACL reconstruction emphasize importance of double bundle approach to restore original anatomy of the ligament. However, studies of the reconstruction outcome have shown that sometimes such joints produced abnormal instability and kinematics [2]. These observations suggest that more understanding of the knee ligaments and reconstruction characteristics are needed in order to improve outcome.

The aim of the present study is to develop a computer graphics based analysis of loading patterns in the femoral attachment of the ACL during a simulated anterior drawer test conducted at different flexion positions of the joint. The outcomes of the test in terms of loading of the ACL fibers could be visualized using graphics as determined by the mathematical model.

2 Methods

A planar mathematical model of the knee was used that allowed simulation of the knee kinematics and anterior drawer test in the sagittal plane. Anatomical shapes of tibial and femoral bones, shapes and positions of articular surfaces, points of attachment of the joint ligaments and material properties of ligaments utilized in the model were based on previous anatomical studies taken from literature [12–15].

As shown in Fig. 1, four major ligaments of the knee, namely, anterior and posterior cruciate ligaments and lateral and medial collateral ligaments were represent as bundles of elastic fibers showing resistance when stretched and buckling when slack [16–19].

The knee joint motion was simulated in the sagittal plane during 0–120° flexion [11–15]. It is shown that in the absence of muscle forces or external loads, the bones rotate and slide relative to each other such that selected fibers in the cruciate ligaments maintain nearly fixed length, that is, neither stretched nor slack [14]. This feature was used in the model to determine the joint kinematics in the unloaded state [15]. As a consequence of relative rotations and translations of the bones, the attachments of the ligaments also changed their relative positions accordingly. Ligaments when slack in any flexion position were represented graphically using buckled fiber [16].

At selected flexion positions, additional relative translations of the bones with all attachment points were superimposed in order to simulate an anterior drawer test that requires translation of the tibia anterior to the femur [16, 17]. Such tests are conducted to estimate integrity of the ligaments. For example, an increased displacement in the anterior direction, normally compared with similar measurement on the contralateral knee, may indicate damage to the ACL [16, 17].

During a simulated anterior drawer test that translates the lower bone in front of the upper bone, the model allowed application of a fixed magnitude of anterior force on

fridaya@uef.fi

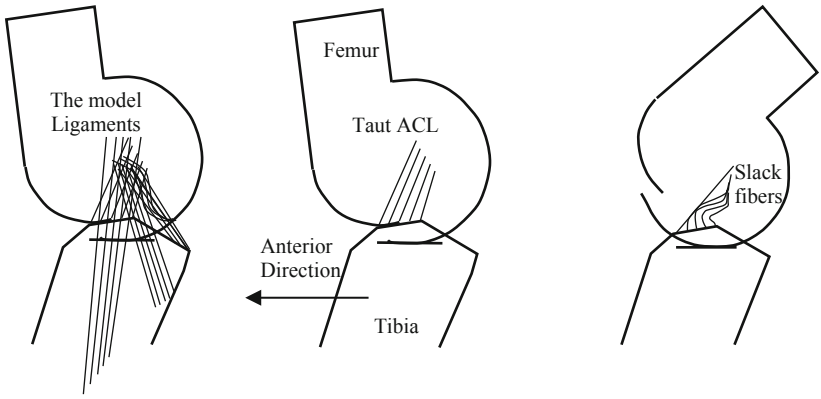


Fig. 1. Graphical representation showing outlines of the model knee and the ligaments represented as straight lines when taut or as curves when slack [16, 18].

the tibia, for example, 90° N and determine the resulting anterior tibial translation (ATT) with resisting forces developed in different fibers of the ACL. Alternatively, the model also allowed application of a pre-determined ATT like 6 mm used in this study with calculations of resulting forces in different fibers of the ACL.

The model ACL was divided into two distinct bundles of fibers representing anteromedial and posterolateral bundles as reported in the literature [4–8]. Each bundle composed of fibers that followed sequential stretching or slackening during motion or due to ATT. Material properties for each bundle were taken from literature [13].

(a) The model knee at 0° flexion with the ligaments represented as bundles of elastic fibers.
(b) The model knee at 0° flexion showing fibers of the ACL, ready to stretch as the tibia is translated anterior (to the left) of the femur.
(c) As the bones rotate and slide during flexion, the ligament fibers may slacken as shown for the ACL at 60° flexion.

3 Results and Analysis

Table 1 gives values of ATT resulting from 90° N anterior force applied on the tibia for different flexion positions of the joint. Model calculations are compared with experimental measurements on 14 cadaver knees shown as mean values with standard deviation as reported by Kondo et al [9]. The model calculations show trends similar to those from experiment. ATT was minimum at full extension or 0° flexion and high at 30° and 60°. The values of ATT decreased in high flexion at 90° and 110°.

Figure 2 shows posterior forces developed in the anterior and posterior bundles of the ACL fibers given as a percentage of total external force applied to translate the tibia 6 mm anterior to the femur at different flexion positions during 0–120° range. At extension, or 0° degree flexion, both the fiber bundles contributed significantly in

fridaya@uef.fi

Table 1. Comparison between model calculations and experimental measurements [9].

Flexion angle	ATT (mm) Model	ATT (mm) Experiment [9] mean (std. dev.)
0°	2.6	1.8 (1.3)
30°	5	4.0 (1.3)
60°	5	3.8 (2.2)
90°	4	3.0 (1.4)
110°	3.7	2.9 (1.7)

resisting the external force. Contribution of the posterior bundle decreased sharply with flexion till about 45°. During mid flexion the external force was fully resisted by the anterior bundle of the ACL. Again, for angles greater than 90°, contribution of the posterior bundle increased.

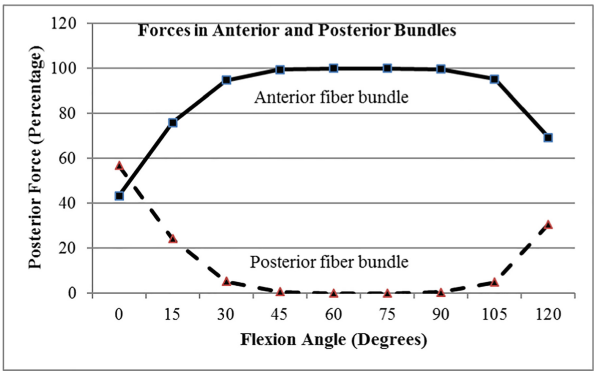


Fig. 2. Contributions of the anterior and posterior bundles of the ACL are shown as percentage of total force resulting from 6 mm ATT during the simulated test over 0–120° flexion.

These patterns of contributions due to anterior and posterior fiber bundles are supported by experimental observations from literature. Kawagachi et al [8] studied 8 cadaver knees to analyze load bearing function of fibers in the femoral attachment of the ACL in resisting tibial displacement. Femoral attachment of the ACL was divided into several areas and the fibers were resected sequentially and resulting changes in translating force were measured corresponding to 6 mm ATT. From measurements, the antero-medial bundle provided 66% to 84% of the total resistance during 0° to 90° flexion, while the postero-lateral bundle contributed nearly 16% to 9% resistance.

Computer graphics for analysis of changing patterns of contributions attributed to different fiber bundles that are presented in this work provide an opportunity to visualize what is, otherwise, difficult to observe through real or physically simulated systems.

fridaya@uef.fi

4 Conclusions

Calculations based on the joint model corroborate the experimental observations in the literature and emphasize the clinical relevance of appropriate tunnel positioning in ACL reconstruction such that the selected femoral areas have significance on load bearing. Further, as anteromedial bundle of the ACL is also associated with maintaining near isometry during flexion [4–8], this position is affected less due stretching or slackening effects during the joint motion. Though limited to 2-dimensions, the model allowed visualization of changes in the ligament fibers at different attachments on the bones during motion or during simulations of the test. Future development of the model will include analysis of partial tears by selective sacrifice of the ligament fibers.

References

1. Jarvela, S., Kiekara, T., Suomalainen, P., Jarvela, T.: Double-bundle versus single-bundle anterior cruciate ligament reconstruction: a prospective randomized study with 10-year results. Am. J. Sports Med. **45**(11), 2578–2585 (2017)
2. Arden, C., Taylor, N., Feller, J., Webster, E.: Return to sport outcome at 2 to 7 years after anterior cruciate ligament reconstruction surgery. Am. J. Sports Med. **40**, 41–48 (2012)
3. Murray, M., Martin, S., Martin, T., Spector, M.: Histological changes in the human anterior cruciate ligament after rupture. J. Bone Joint Surg. **82-A**(10), 1387–1397 (2000)
4. Amiri, S., Derek, T., Cooke, V., Wyss, U.: A multiple-bundle model to characterize the mechanical behavior of the cruciate ligaments. Knee **18**(1), 34–41 (2011)
5. Mommersteeg, M., Blankevoort, L., Huiskes, R., Kooloos, J., Kauer, J.: Characterisation of the mechanical behavior of human knee ligaments: a numerical-experimental approach. J. Biomech. **29**(2), 151–160 (1996)
6. Amis, A., Dawkins, G.: Functional anatomy of the ACL–fibre bundle actions related to ligament replacement and injuries. J. Bone Jt. Surg. (Br) **73-B**, 260–267 (1991)
7. Butler, D., Kay, M., Stouffer, D.: Comparison of material properties in fascicle-bone units from human patellar tendon and knee ligaments. J. Biomech. **19**(6), 425–432 (1986)
8. Kawaguchi, Y., Kondo, E., Takeda, R., Akita, K., Yasuda, K., Amis, A.: The role of fibers in the femoral attachment of the anterior cruciate ligament in resisting tibial displacement. J. Arthroscopic Relat. Surg. **31**(3), 435–444 (2015)
9. Kondo, E., Merican, A., Yasuda, K., Amis, A.: Biomechanical analysis of knee laxity with isolated anteromedial or posterolateral bundle deficient anterior cruciate ligament. J. Arthroscopic Relat. Surg. **30**(3), 335–343 (2014)
10. Imran, A.: Relating knee laxity with strain in the anterior cruciate ligament. In: Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering, pp. 1037–1042, London (2017)
11. Imran, A.: Analyzing anterior knee laxity with isolated fiber bundles of anterior cruciate ligament. In: Proceedings of The World Congress on Engineering, pp. 869–872, London (2016)
12. Imran, A., O'Connor, J.: Control of knee stability after ACL injury or repair: interaction between hamstrings contraction and tibial translation. Clin. Biomech. **13**(3), 153–162 (1998)
13. Zavatsky, A.: The functional architecture of human knee ligaments. PhD thesis, University of Oxford (1993)

fridaya@uef.fi

14. O'Connor, J., Shercliff, T., Biden, E., Goodfellow, J.: The geometry of the knee in the sagittal plane. Proc. Inst. Mech. Eng. (Part H), J. Eng. Med. **203**, 223–233 (1989)
15. Imran, A.: Modelling and simulation in orthopedic biomechanics-applications and limitations. In: Tavares, J.M., Jorge, R.M. (eds.) Computational and Experimental Biomedical Sciences: Methods and Applications. Springer, Cham (2015)
16. Lu, T.W., O'Connor, J.: Fiber recruitment and shape changes of knee ligaments during motion as revealed by a computer graphics based model. Proc. Inst. Mech. Eng. (Part H), J. Eng. Med. **210**, 71–79 (1996)
17. Imran, A.: Influence of flexing load position on the loading of cruciate ligaments at the knee–a graphics-based analysis. In: Tavares, J.M., Jorge, R.M. (eds.) Computational and Experimental Biomedical Sciences: Methods and Applications. Springer, Cham (2015)
18. Imran, A.: Sagittal plane knee laxity after ligament retaining unconstrained arthroplasty: a mathematical analysis. J. Mech. Med. Biol. **12**(2), 1–11 (2012)
19. Imran, A.: Computer graphics based approach as an aid to analyze mechanics of the replaced knee. In: Arai, K., Kapoor, S., Bhatia, R. (eds.) Advances in Intelligent Systems and Computing, vol. 857. Springer, Cham (2019)

fridaya@uef.fi

Gesture Recognition Using an EEG Sensor and an ANN Classifier for Control of a Robotic Manipulator

Rocio Alba-Flores(✉), Fernando Rios, Stephanie Triplett, and Antonio Casas

Georgia Southern University, 1100 IT Drive, Statesboro, GA 30458, USA
ralba@georgiasouthern.edu

Abstract. In recent years, electroencephalography (EEG) has gained popularity in the field of brain-computer interface (BCI). Current applications of BCI include control of prosthetics and robotic systems. In this project, the goal is to acquire and record EEG signals generated by human subjects performing specific facial gestures, and use them to control a robotic hand. Six facial gestures have been selected for this project: smile, raise eyebrows, look right, look left, hard blink, and blink. Once the signals were collected, a classification system based on artificial neural network (ANN) was designed. The classification system was able to recognize and differentiate each gesture with an accuracy of 98% for signals from a single person, and 75% for signals from multiple persons. The EEG signals were acquired using an Emotiv EPOC headset that has 14 sensors. This headset was selected mainly because its portability, affordable cost compared to similar products in the market, and it is easy to place on the subject's head. The ultimate purpose of this research is to use the classification system output to send control signals to a robotic hand that has been designed and built in our research lab. In this paper, the data collection, data conditioning, design, testing and results of the classification system are provided in detail.

Keywords: Electroencephalography (EEG) · Neural network classifier · Brain computer interface (BCI) · Robotic arm · Face gesture detection

1 Introduction

All human activities in form of communication and control require nervous and muscle coordination. The activity process starts from the intent of the subject by triggering several parts of the brain that will send signals to the part of the body that will perform the activity. The human brain consists of millions of neurons which play an important role for controlling behavior of human body with respect to internal and external motor and sensory stimuli. These neurons will act as information carriers between human body and the brain. Electroencephalography (EEG) is an efficient method that helps to acquire brain signals from the scalp surface area that corresponds to various states or activities of a subject. These signals are generally categorized as delta, theta, alpha, beta

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1181–1186, 2019.
https://doi.org/10.1007/978-3-030-22868-2_81

fridaya@uef.fi

and gamma based on signal frequencies ranges from 0.1 Hz to more than 100 Hz. In recent years, advancement of technology has expanded the application of EEG from pure medical applications to a wider area in brain-computer interface (BCI) applications.

Examples of early research on EEG signals for BCI applications include [1], where the authors indicated that in order to translate human's intent, the brain-computer interface (BCI) should be intensively trained for each user. Even though the BCI is properly trained, BCI usage still requires user's special skills that need training. In [2] the authors investigated the possibility that human's brain activity detected by EEG signals, could actually move a mouse cursor on a computer. This study showed that BCI is able to translate human's intent into action into a computer. Recent research areas where BCI has being used, include game controller [3], wheelchair controller [4], hearing aid devices [5], and prosthetic robotic arm controller [6, 7]. With these wide range of possible applications, the area for BCI has vast potential for further research.

The objective of this research was to acquire and classify EEG signals from subjects when they perform facial gestures. These gestures will be used to control a robotic hand. To acquire the EEG signals, the Emotiv EPOC [8] neuro headset system was used. The subjects perform six facial gestures, and an intelligent system (based on artificial neural networks, ANN) classifies the gestures. Based on the particular gesture, the intelligent system sends command signals to a robotic hand to perform a specific task such as close the hand. The sections in this paper are as follows: in Sect. 2 the neuro headset used in this research is described; in Sect. 3 the data collection process is explained; in Sect. 4 the type of ANN is described; in Sect. 5 the results that were obtained are discussed; and in Sect. 6 the conclusions and further work are presented.

2 EEG Neuro Headset System

The EEG neuro headset system used in this research is the Emotiv EPOC headset system that contains 16 biosensors attached to the headset as shown in Figs. 1 and 2. Fourteen of these sensors are signal sensors, and two are reference sensors. The two reference sensors are the Common Mode Sense (CMS), that is and active electrode and serves as the absolute voltage reference; and the Driven Right Leg (DRL) passive electrode, its function is to provide feedback cancellation and float the reference level on the common mode body potential. These two references electrodes are located at the P3 and P4 locations in the International 10–20 EEG placement system [9]. The 16 sensors are attached to flexible plastic fingers that are designed to easily be placed on the scalp of the subject. This EEG system has a sampling frequency of 128 samples per second, a resolution of 14 bits, and bandwidth 0.16–43 Hz. This headset was chosen due to: it is compact, inexpensive (compared to similar products in the market), easy to place on the subject's head, has been tested in similar research work, and it is reliable in gathering wave pattern data of motor movements for classification through neural networks.

fridaya@uef.fi

3 Subject Preparation and Data Collection

Before any activity was performed on human subjects, an Institutional Review Board (IRB) approval was received, and each subject signed a consent form giving permission to acquire her/his EEG signals and use them for this research. Currently the EEG database has data from 20 subjects (male and female, ages 19–27). The EEG data collection process was performed as follows: (i) preparation of the sensors (small felt pads of about 0.5 inch), all sensors and cleaned with an antiseptic solution; (ii) the sensors are placed on the corresponding sockets on the headset, and the headset is placed on the subject's head; (iii) subjects were asked to perform facial gestures, first the subject was instructed to carry out a relaxing position and after a few seconds in the relaxing position, the subject was instructed to perform the facial gestures, at this moment the recorded of the signal started; (iv) six facial gestures were performed by the subject, and each gesture was repeated three times. The six facial gestures that were selected for this study were: smile, raise eyebrows, look right, look left, hard blink, and blink (Table 1). These gestures were selected because they are easy to perform by any subject and are also the ones (that we detected) give the signals that have less correlation to each other.

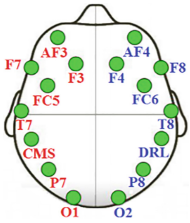
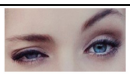
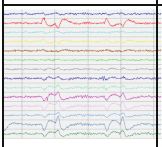
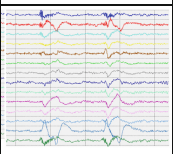
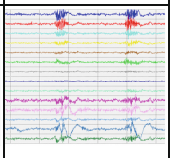
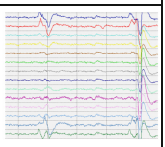


Fig. 1. The 16 sensors on the headset system

Fig. 2. Emotiv EPOC headset system [8]

Table 1. The six facial gestures, and samples of collected EEG signals.

Class	1	2	3	4	5	6
Motion	Hard Blink	Look Left	Look Right	Raise Eyebrows	Smile	Blink
						
						

fridaya@uef.fi

4 Artificial Neural Network Development

For this research, the development of the classifier was based on training an ANN for each individual subject. Once the ANN was trained, the subject was able to perform any of the 6 gestures and the ANN was able to classify the new data with an accuracy of 95%. If a subject perform a gesture not recognized the NN does not provide any output.

The architecture of the ANN implemented in this project was of the feedforward backpropagation type, with multi-layer perceptron. It was developed using the scaled conjugate gradient training function with 12 neurons in the input layer, each input correspond to signals from a specific sensor. The output layer contains 6 neurons, each corresponds to one of the facial gestures to be classified. The output layer uses the tan-sigmoid (tan-sig) function. The ANN has one hidden layer that contains 15 neurons, and uses the tan-sig function also. To select the number of neurons in the hidden layer, the starting point was to test between 2 and 20 neurons, and then adjusted for accuracy based on training results [10]. The best accuracy was found with 15 neurons in the hidden layer. 70% of the recording data were used for training, 15% was used for validation, and 15% for testing. The Neural Network Toolbox from Matlab was used in the development of the ANN.

5 Results

Figure 3 shows the overall confusion matrix (training, validation, and testing). The numbers of points entered for each gesture were between 9,500 and 11,300 depending on how much time took the subject to perform each gesture 3 times (this is 14 channels * 128 SPS * 3 trials * (1.5 to 2.5) seconds/trial). In the confusion matrix the rows correspond to the predicted output class and the columns correspond to the true target class. The number and percentage of correct observations is shown along the diagonal. The last box in the diagonal shows the overall accuracy, 98%. The boxes, in red show the number and percentage of errors. The largest error found in this study was 0.3% that was obtained between the hard blink and soft blink gestures. The training and testing samples were randomly selected in order to avoid correlation among the samples. Given that the model accuracy degrade from 98% for one subject to 75% with 20 subjects, it is expected that the accuracy will drop when more subjects are added to the training set. So further research is needed to find a more suitable neural network architecture that will be able to increase the accuracy when data of more than one subject is used to train the neural network.

fridaya@uef.fi

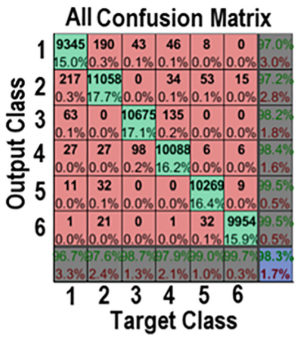


Fig. 3. Resulting of overall confusion matrix of the neural network classifier, showing an overall accuracy of 98.3%

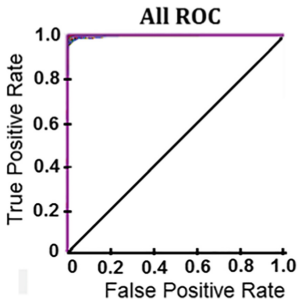


Fig. 4. Resulting Receiver Operating Characteristic demonstrating near perfect classification of all classes.

Figure 4 represents the results of the Receiver Operating Characteristic (ROC) analysis of the ANN. An ROC curve is a plot of operating points showing the possible tradeoff between a classifier's true positive (TP) rate versus its false positive (FP) rate. The TP rate is commonly referred to as "sensitivity," and (1-FP rate) is called "specificity" [11]. The interpretation of a ROC curve can be summarized as: (i) the closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test; (ii) the closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test. Therefore, the ROC curves in Fig. 4 indicate high accuracy of the designed classifier.

6 Conclusions

The ultimate goal of this research is to have an intelligent system that is able to control a robotic hand using facial gestures. The classification system was able to recognize and differentiate each gesture with an accuracy of 98% for signals from a single person.

Currently our research team has designed an intelligent classifier that uses the data from the 20 subjects, and preliminary results have shown an accuracy of 75%. Future work in this project includes to test the ANN classifier with more subjects, and then create the hardware interfaces to control the robotic hand. In the context of the proposed application of the system, this level of accuracy is acceptable because this is the starting state of the project and further improvements need to be made in order to make the system more reliable when data from more subjects is added to the training and testing data sets.

References

1. Guger, C., Schlogl, A., Neuper, C., Walterspacher, D., Strein, T., Pfurtscheller, G.: Rapid prototyping of an EEG-based brain-computer interfaces (BCI). IEEE Trans. Neural Syst. Rehabil. Eng. **9**(1), 49–58 (2001)

fridaya@uef.fi

2. Fabiani, G.E., McFarland, D.J., Wolpaw, J.R., Pfurtscheller, G.: Conversion of EEG activity into cursor movement by a brain computer interface (BCI). IEEE Trans. Neural Syst. Rehabil. Eng. **12**(3), 331–338 (2004)
3. Liao, L., Chen, C., Wang, I., Chen, S., Li, S., Chen, B., Chang, J., Lin, C.: Gaming control using a wearable and wireless EEG-based brain-computer interface device with novel dry foam-based sensors. J. Neuro Eng. Rehabil. **9**(5), 5 (2012). https://doi.org/10.1186/1743-0003-9-5
4. Long, J., Li, Y., Wang, H., Yu, T., Pan, J., Li, F.: A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair. IEEE Trans. Neural Syst. Rehabil. Eng. **20**(5), 720–729 (2012)
5. Yoon, J., Nam, K.W., Lee, J.C., Jang, D.P., Kim, I.Y.: An eye-blinking-based beamforming control protocol for hearing aid users with neurological motor disease or limb amputation. Artif. Organs **41**(2), 162–168 (2017). https://doi.org/10.1111/aor.12755
6. Bhattacharyya, S., Konar, A., Tibarewala, D.N.: Motor imagery, P300 and error–related EEG–based robot arm movement control for rehabilitation purpose. Med. Biol. Eng. Comput. **52**, 1007–1017 (2014). https://doi.org/10.1007/s11517-014-1204-4
7. Sunny, T.D., Aparna, T., Neethu, P., Venkateswaran, J., Vishnupriya, V., Vyas, P.S.: Robotic arm with brain-computer interfacing. In: International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST), Procedia Technology, vol. 24, pp. 1089–1096 (2016)
8. Emotiv EPOC14 Channel Wireless EEG Headset: https://www.emotiv.com/epoc/
9. Acharya, J.N., Hani, A., Cheek, J., Thirumala, P., Tsuchida, T.N.: American clinical neurophysiology society guideline 2: guidelines for standard electrode position nomenclature. J. Clin. Neurophysiol. **33**(4), 308–311 (2016). https://doi.org/10.1097/WNP.0000000000000316
10. Heaton, J.: Feedforward backpropagation neural networks, Ch. 5, Sec. 5. In: Introduction to Neural Networks for Java, 2nd edn. Heaton Research Inc., Chesterfield (2008)
11. Woods, K., Bowyer, K.W.: Generating ROC curves for artificial neural networks. IEEE Trans. Med. Imag. **6**(3), 329–337 (1997)

fridaya@uef.fi

An Eye Tracking Study: What Influences Our Visual Attention on Screens?

Arwa Mashat(✉)

King Abdulaziz University, Rabigh, Saudi Arabia
aasmashat@kau.edu.sa

Abstract. This study investigated the influence of language orientation on visual attention and recalled information using eye tracking technology. A total of 37 bilingual participants completed the study. All participants' native language orientation was from right to left such as Arabic, Urdu, Farsi and Hebrew, but all resided in the US. Participants speaking Arabic as their native language were assigned to Group1 (N = 20), otherwise they were assigned to Group2 (N = 17). Group1 received the instructional material in Arabic and Group2 in English. The study consisted of three parts: Word screen, Image screen and Sequence screen. An eye tracking devise was used to measure first visual fixation indicating the region (right or left) where the user's attention was on the screen for all three parts. Afterwards, the participants were asked to recall three words, three images and explain the sequence process of the images presented. One-way ANOVA was used to assess the relationship between the visual fixations with the recalled answered orientation for both groups. Significance was found in visualizing words, images and the sequence process. Results indicate initial visual fixation is influenced by language orientation when reading the instruction of questions as the case in Group1, or the influence of language orientation of the country of residence as for Group2. This can help designers, developers and instructors when designing to get the international users' attention.

Keywords: Eye tracking · Language · Visual fixation

1 Introduction

Usually people when looking at a page, screen or environment around them, they tend to organize what they see. Some features can influence visualization when designing on screens [5]. Visualization is important to consider when designing an advertisement, or instructional material for learners. Users' attention can make a difference in the outcomes of learning or even marketing. Knowing the target audience when designing is important. The audiences' culture, traditions and language can have an impact. Designers need to avoid English as the only language attitude [3]. English learner's reading pattern from left to right is strong. Reversing it goes against their logic and interferes with learning [8]. Other languages such as Arabic, Farsi, Hebrew and Urdu have a right to left orientation.

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1187–1192, 2019.
https://doi.org/10.1007/978-3-030-22868-2_82

fridaya@uef.fi

Investigating eye movements can help designers determine the user's visual attention. Visual fixation is the eye's period of stability when focused on a point of interest [2, 4]. Eye movements can be considered as the blueprint for how users process the visual information [7]. Eye tracking can be used to determine if the language orientation has an impact on visual attention. The number of years the individual's eyes and hands move in a particular direction when reading and writing can have some effects outside the domain of reading and writing [6].

Image recall, visual attention and sequential arrangements of English, Chinese and Taiwanese speakers were studied by Chan and Bergen [1]. They found that writing direction affects cognition by the way learners visualize, remember and arrange items.

This study was intended to investigate if the Arabic and English language orientation influences how users visualize material (text and images) on the screen. In addition, recalling was investigated to find out if recalled information orientation was aligned with visualization. This will help designers, instructors and developers when designing for audiences of different languages.

The study covered the following research questions:

1. To what extent does language orientation influence user's first visual fixation on the material presented on screen?
2. To what extent does language orientation influence user's recalling the material presented on screen?

2 Methodology

2.1 Participants

A total of 37 bilingual participants completed this study. All participants' native language orientation is from right to left such as (Arabic, Farsi, Urdu, Hebrew) and second language is English. Participants were assigned to Group1 if their native language was Arabic, otherwise they were assigned to Group2. Group1 (N = 20) received the instructional material in Arabic (language orientation right to left) and Group2 (N = 17) received the instructional material in English (language orientation left to right). They all reside in the US either for studying or working, so they are in an English language environment. However, 90% of Group1 speak Arabic at home, and 82% of Group2 speak Farsi or Urdu.

2.2 Procedure

All experimental material was delivered on a single PC with 20-inch LCD, color monitor. Material was designed by the researcher and presented to participants through online survey software. The material consisted of four parts: demographic questions, Word screen (see Fig. 1), Image screen (see Fig. 2) and Sequence screen (see Fig. 3). The Text screen consisted of 16 three letter words written in black as a 4×4 matrix.

fridaya@uef.fi

This arrangement allows for placing words equally into two regions: right and left. The Image screen consisted of black and white individual images distributed equally into two regions: right and left. However, the Sequence process screen consisted of three images logically sequenced in both orientations.

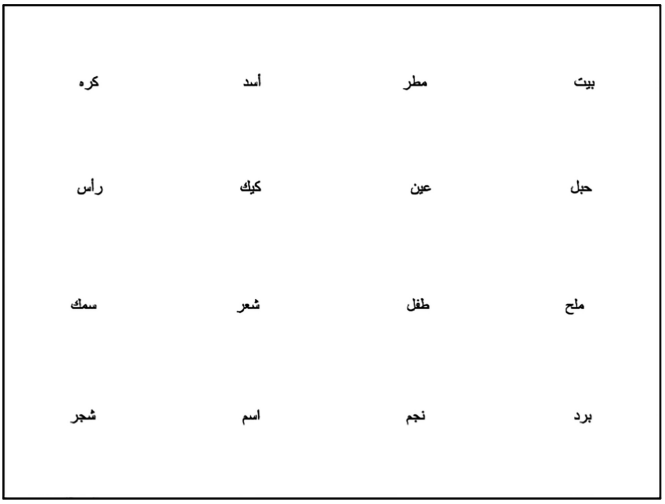


Fig. 1. The word screen

Fig. 2. The image screen

fridaya@uef.fi

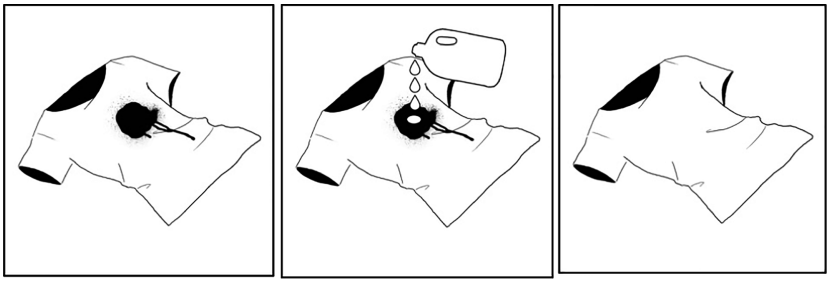


Fig. 3. The sequence process

To answer the first research question, an eye tracking devise was used to measure first visual fixation orientation indicating the region of attention either right or left. Eye tracking measures were collected using a SmartEye Pro, unteathered eye tracking system configured with three mini CCD camera, infra-red (IR) lighting, and eye tracking software to capture head and eye movements. Data were collected using MAPPS accessory software to facilitate recording and overlap of both instructional software screens and eye movements, identification of regions associated with text and images, and real-time video replay of each participant's visual attention to on-screen multi-media elements for the purpose of verifying accuracy of real-time data and summative measures of eye fixations.

For the second research question, participants were asked to list three words they recall from the Word screen, and three images from the Image screen. The words and images recalled were counted to indicate the region of orientation. They were also asked to write a sentence explaining the process of the sequenced images they comprehended. Their explanation start point determined the orientation right or left.

3 Results

For each group, we compared their first visual fixation orientation with their recalled written responses' orientation. Data were analyzed using one-way ANOVA to assess for the independent variable groups (Group1, Group2) and dependent variables (VisualWords, RecallWords) to evaluate the relationship between users recalled responses orientation and their first visual fixation orientation on screen. Significance was found in visualizing the Word screen F (1, 35) = 4.359, p = 0.04, $\eta^2 = 0.111$ between the two groups, Group1 (M = 1.75) and Group2 (M = 1.35).

Another one-way ANOVA assessed the independent variable groups (Group1, Group2) and dependent variables (VisualImage) and (RecallImage) to evaluate the relationship between users recalled responses orientation and their first visual fixation orientation on screen. Significance was found visualizing the Image screen F (1, 35) = 35.461, p = 0.00, $\eta^2 = 0.503$ between Group1 (M = 1.85) and Group2 (M = 0.94).

Finally, one-way ANOVA assessed the independent variable groups (Group1, Group2) and dependent variables (VisualSequence) and (WriteSequence) to evaluate the relationship between users written response orientation and first visual fixation

fridaya@uef.fi

orientation on screen. Significance was found $F(1, 35) = 14.821$, $p = 0.00$, $\eta^2 = 0.297$ between groups Group1 ($M = 2.7$) and Group2 ($M = 1.8$) (see Fig. 4).

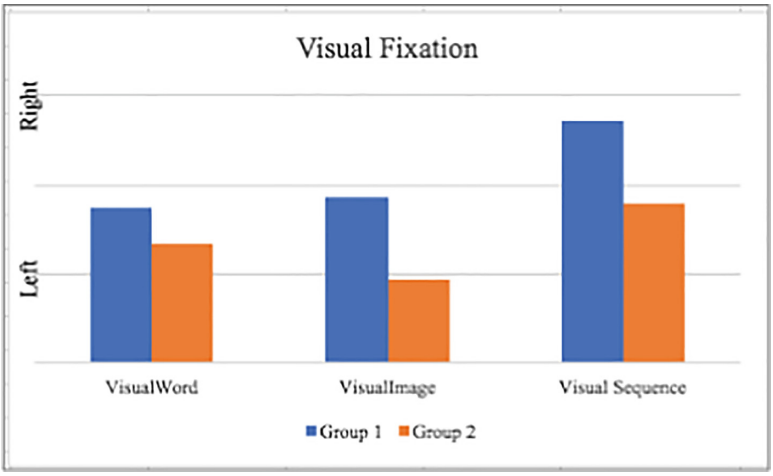


Fig. 4. Eye tracking measuring visual fixation orientation

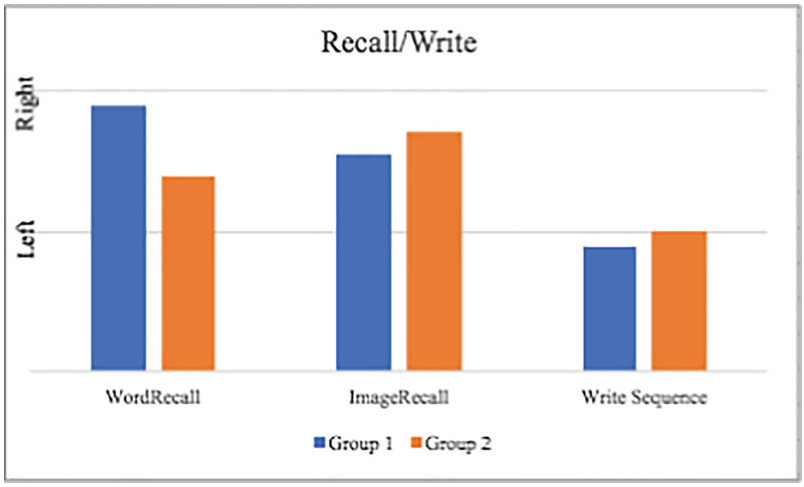


Fig. 5. Recall and written response orientation

4 Conclusion

Measuring learners' attention on screens is difficult, however it is possible by using eye tracking technology [2]. Results indicate initial visual fixation is influenced by language orientation when reading the instruction of questions as the case in Group1, or

fridaya@uef.fi

the influence of language orientation of the country of residence as for Group2. No significance was found with recalled and written responses orientation (see Fig. 5).

Designers, developers and instructors need to be familiar with international users' languages and cultures from the beginning to get their attention and avoid any misunderstanding.

References

1. Chan, T., Bergen, B.: Writing direction influences spatial cognition. In: Bara, B., Barsalou, L. W., Bucciarelli, M. (eds.) Proceedings of the 27th Annual Conference of the Cognitive Science Society, pp. 412–417. Erlbaum, Mahwah (2005)
2. Chuang, H., Liu, H.: Effects of different multimedia presentations on viewers' information-processing activities measured by eye-tracking technology. J. Sci. Educ. Technol. **21**(2), 276–286 (2012)
3. Lu, M.: Meeting the challenge of designing multimedia for an international audience. Educ. Media Int. **35**(4), 272–274 (1998)
4. Slykhuis, D., Wiebe, E., Annetta, L.: Eye-tracking students' attention to PowerPoint photographs in a science education setting. J. Sci. Educ. Technol. **14**(5–6), 509–520 (2005)
5. Renshaw, J.A., Finlay, J.E., Tyfa, D., Ward, R.D.: Designing for visual influence: an eye tracking study of the usability of graphical management information. Hum. Comput. Interact. **1**, 144–151 (2003)
6. Treiman, R., Allaith, Z.: Do reading habits influence aesthetic preferences? Read. Writ. **26**(8), 1381–1386 (2013)
7. Underwood, G., Radach, R.: Eye guidance and visual information processing: reading, visual search, picture perception and driving. In: Underwood, G. (ed.) Eye Guidance in Reading and Scene Perception, pp. 1–28. Elsevier, Oxford (1998)
8. Winn, W., Holliday, W.: Design principles for diagrams and charts. Technol. Text **1**, 277–299 (1982)

fridaya@uef.fi

Sports Vision Based Tennis Player Training

Kohei Arai(✉)

Saga University, Saga 8408502, Japan
arai@is.saga-u.ac.jp

Abstract. Sport vision based tennis player training is proposed to accelerate players' skill-up in tennis play with instructions provided by the proposed sport vision based system. In sports, gaze, dynamic visual acuity, eye movement and viewing place are important. In sports vision, Static eyesight, Dynamic visual acuity, Contrast sensitivity, Eye movement, Deep vision, Instant vision, Cooperative action of eye, hand and foot, and Peripheral field are have to be treated. In particular for the tennis, all of the items are very important. Therefore, sports vision based tennis player training system is proposed. Through experiment, it is found that the proposed system does work well for improvement of tennis players' skills.

Keywords: Sport vision · Static eyesight · Dynamic visual acuity · Contrast sensitivity · Eye movement · Deep vision · Instant vision · Cooperative action of eye · Hand and foot, and peripheral field

1 Introduction

Sports vision based tennis player training system is proposed. There is Sport Vision Association in Japan. In sports, gaze, dynamic visual acuity, eye movement and viewing place are important. In sports vision, Static eyesight, Dynamic visual acuity, Contrast sensitivity, Eye movement, Deep vision, Instant vision, Cooperative action of eye, hand and foot, and Peripheral field are have to be treated. In particular for the tennis, all of the items are very important. The proposed sport vision based system would help beginners of tennis player by providing useful information such as (1) predicted position of tennis ball at an appropriate seconds after hitting of the tennis ball, (2) which direction tennis players should look at, etc. These instructions are provided through wireless sensor network from the server to the clients (beginner of tennis player).

To identify the line of sight vector of the tennis player, it requires face detection and gaze estimation. Face detection from the acquired images is common and popular technology. Computer input system based on viewing vector estimation with iris center detection from face image acquired with web camera allowing users' movement is proposed [1]. Also, method for face identification with Facial Action Coding System: FACS based on eigen value decomposition is proposed [2].

Using the line-of-sight camera, acquire eye-gaze video and analyze the difference between the beginner and the experienced person or the gaze at good and bad times. Also, the proposed system would help beginners of tennis players for their technical

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1193–1201, 2019.
https://doi.org/10.1007/978-3-030-22868-2_83

fridaya@uef.fi

improvement on tennis play. Using head set of work camera, players' gaze is estimated together with ball trajectory for prediction. Then, some instructions can be provided to the player in concern.

In the next section, the proposed system for improving tennis players' skills is described followed by the preliminary experiment. After that, conclusion is described together with some discussions and future research works.

2 Proposed System

2.1 Design Concept

The proposed system provides expected position of tennis ball immediately after the ball is hit by the offensive player, as well as instructions and week points to tennis players. Figure 1 shows design concept of the proposed system. There are two cameras, one is observe the tennis court and the other one is attached to the tennis player's head in the system.

From the video data acquired #1 camera, tennis ball is looked at and always calculate the trajectory and predict the expected tennis ball position touch down on the court. Therefore, instruction can be provided to the tennis player. Thus, the tennis player can take a next action so quickly. On the other hand, #2 camera is always looking forward forehead direction, head pose direction. Therefore, it is understand which direction is tennis player is looking at. Thus, week points can be provided when the tennis player is not looking at the appropriate direction after the play. That will help to improve their skills.

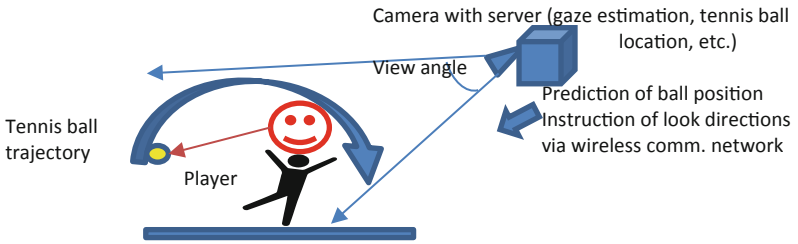


Fig. 1. Design concept

fridaya@uef.fi

2.2 Software and Hardware

Tennis ball can be extracted from the video and analyze where the players are looking together with gaze direction. There are some actions in tennis plays, Serve, Stroke, Volley, Smash, Match (singles). These actions are to be identified from the acquired videos. The identified action types can be provided to the tennis player through wireless communication network.

Major hardware and software used are as follows:

- GoPro HERO6 (Camera)
- OpenCV 3.1.0 for image analysis
- Python 2.7 or Visual Studio2017 C++

Major specification of GoPro HERO6 is as follows:

- Weight: 117 g
- Action camera
- Image stabilization
- Hi-Vision
- Waterproof function
- Field of view: “Wide angle”, “fisheye”
- Frame rate 4 K/60 fps, 1080p/240 fps

Figure 2 shows outlook of the GoPro HERO6.

(a)Back view (b)Front view

Fig. 2. Outlook of the GoPro HERO6.

Major characteristics of tennis players have to be assumed from data by SONY smart tennis sensor (SONYSmartTennisSensor). Major characteristics of tennis player are assumed as follows:

Forehand Stroke Speed: 90 km/h (25 m/s) within 1.0 s
Serve Speed: 115 km/h (31 m/s) with in about 0.8 s

On the other hand, the vertical length of the tennis court (the distance between the players): 23.77 m (about 25 m). Figure 3 shows the tennis court dimension.

fridaya@uef.fi

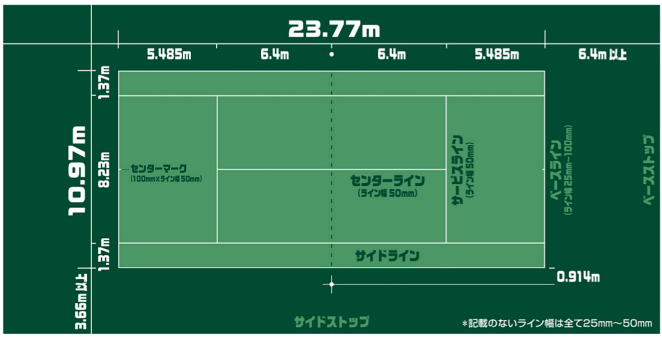


Fig. 3. Tennis court dimension.

Meanwhile, OpenCV is used to analyze the acquired imagery data. The main functions are Filtering, Matrix operation, Object tracking, Area segmentation, Camera calibration, Feature point extraction, Object recognition, GUI, Machine learning, Panorama synthesis, Computational photography. Using these functions, tennis ball location and gaze direction can be estimated.

2.3 Proposed Ball Detecting Method

There are two conventional methods for ball detection, shape based and color based methods. Figure 4 shows example of detected circle using shape based detection method. On the other hand, Fig. 5 shows the detected tennis ball based on color based method.

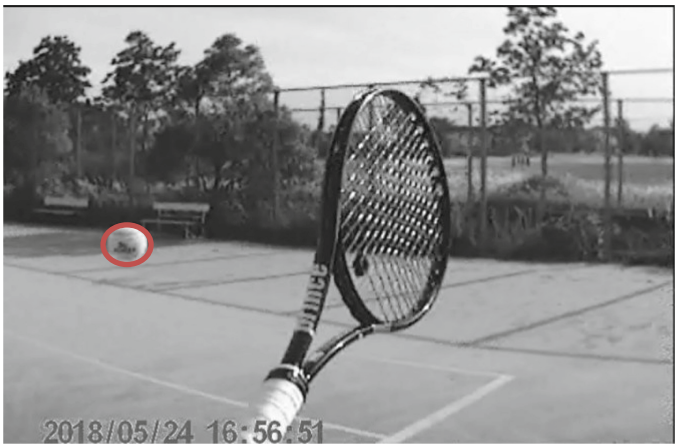


Fig. 4. Examples of circle detection

fridaya@uef.fi

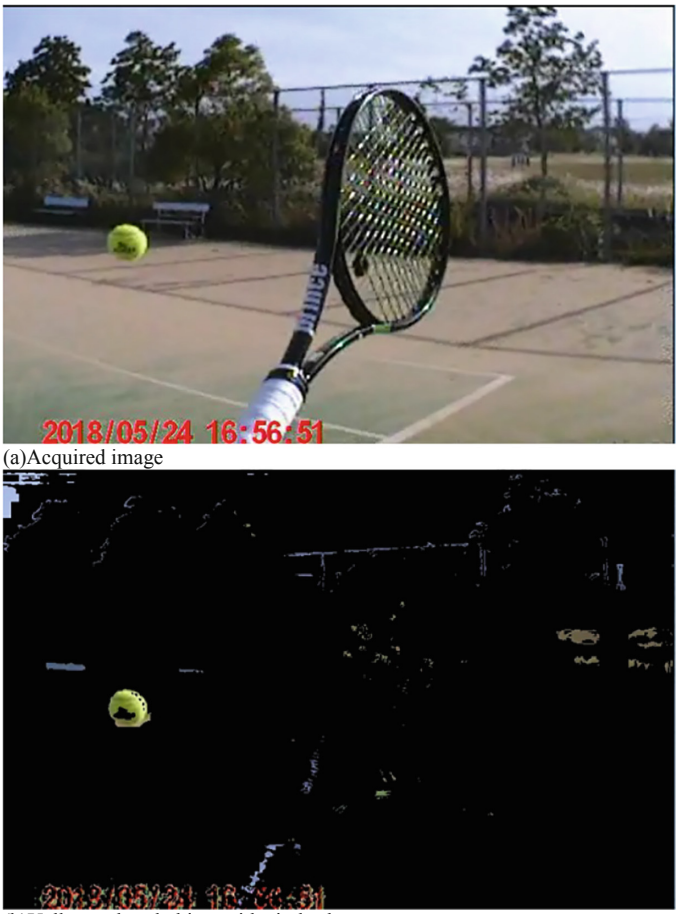
Fig. 5. Example of yellow colored object detection.

The proposed tennis ball detection method is based on the combination between shape based and color based methods. Therefore, tennis ball can be detected even if the tennis ball is not circle shape and even if the tennis ball is not yellow color.

3 Preliminary Experiments

Figure 6a shows the acquired camera image while Fig. 6b shows the resultant image of yellow colored object with circle shape feature detection. It is not always true that the shape of tennis ball is circle. As shown in Fig. 7, the closest tennis ball can be detected even if the extracted shape of ball is not circle. Figure 7a shows yellow colored portion of the acquired image while Fig. 7b shows the detected tennis ball which is the closest tennis ball among the candidates of tennis ball without circle shape feature. As shown in Fig. 7, even if the aspect ratio of the acquired image is changed due to image acquisition condition is changed, tennis ball in concern can be detected even if the some portions of the tennis ball is shaded as well as includes shadows. In this situation, maximum contour of object detection is one of the key issues here.

fridaya@uef.fi

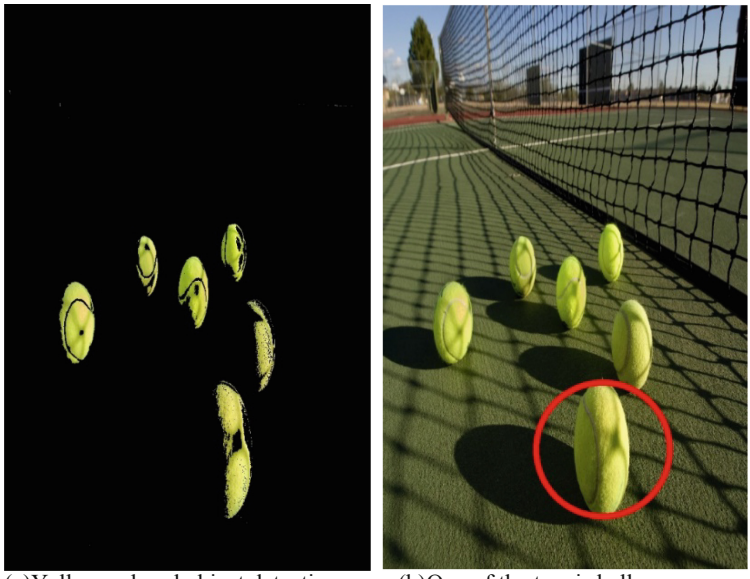


(a)Acquired image

(b)Yellow colored object with circle shape

Fig. 6. Example of yellow colored object with circle shape feature detection.

fridaya@uef.fi



(a)Yellow colored object detection (b)One of the tennis ball

Fig. 7. Another example of tennis ball detection

Meanwhile, yellow color feature can be adjustable for changing threshold of hue information. Figure 8a shows and the resultant image of the yellow colored object detection with relatively large threshold of hue information while Fig. 8b shows the resultant image of the yellow colored object detection with comparatively small threshold of hue information. Eventually, the most appropriate tennis ball can be detected by adjusting threshold for hue information as shown in Fig. 8c.

By adjusting the threshold for the hue information of the color based method, Fig. 8 of the resultant image can be derived. In this case, skin color and yellow colored objects are detected.

fridaya@uef.fi

(a)Yellow colored objects detected image with large threshold of hue information

(b)Yellow colored objects detected image with small threshold of hue information

(c)Detected tennis ball

Fig. 8. Result image obtained by adjusting the threshold of hue information

fridaya@uef.fi

4 Next Step

Although this paper deals with just tennis ball and face detections, estimation of line of sight of tennis players would be the next step. Beginners' look directions are different from those points looking at for expertized tennis player. Instructions for most appropriate look direction will be provided by the system. Then, action classification from the acquired moving pictures will be conducted. If it is possible to identify just after the initial action, then next recommendable action can be recognized. After that some preparatory action can be done with the recognized actions.

5 Conclusion

Major conclusion is that the proposed sport vision based system allows detection of tennis ball by adjusting threshold level of hue information of acquired images even if the shape of object is not circle, and even if the color of tennis ball is close to players' skin color.

The followings are future research works:

(1) Gaze analysis from acquired images,
(2) Some valuable instructions can be delivered to the players after the estimation and prediction of the tennis ball trajectory as well as players intensive look direction derived from the gaze analysis of the players. Beginners' look directions are different from those look directions for expertized tennis player.
(3) Then, action classification from the acquired moving pictures will be conducted. If it is possible to identify just after the initial action, then next recommendable action can be provided to tennis player.

Acknowledgment. The author would like to thank Mr. Toshiki Nishimura of former student in the Information Science Department of Saga University for his experimental efforts.

References

1. Arai, K., Uwataki, H.: Computer input system based on viewing vector estimation with iris center detection from face image acquired with web camera allowing users' movement. Electron. Commun. Jpn. **92**(5), 31–40 (2009)
2. Arai, K.: Method for face identification with facial action coding system: FACS based on eigen value decomposition. Int. J. Adv. Res. Artif. Intell. **1**(9), 34–38 (2012)

fridaya@uef.fi

No-Reference Image Quality Assessment Based on Multi-scale Convolutional Neural Networks

Peikun Chen[1,2], Yuzhen Niu[1,2(✉)], and Dong Huang[1,2]

[1] College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350116, China
{chenpk20,yuzhenniu}@gmail.com, huangdongfzu@163.com
[2] Fujian Key Laboratory of Network Computing and Intelligent Information Processing, Fuzhou University, Fuzhou 350116, China

Abstract. Big data brings more challenges to image quality assessment (IQA). First of all, reference images are not always available, so full-reference IQA (FR-IQA) and reduced-reference IQA (RR-IQA) can be problematic for big data. Because of no requirement for any pristine knowledge, no-reference IQA (NR-IQA) has recently received a great deal of attention. In recent years, deep convolution neural network (CNN) has achieved many successes in visual recognition tasks. Researches in literature have shown that NR-IQA metrics can be learned from features extracted using deep CNN. But existing CNN-based IQA models focus on local spatial information but ignore global spatial structure. In this paper, we propose a NR-IQA model learns from multi-scale CNN (MSCNN). Multi-scale mechanism has boosted the performances of many singe-scale methods in a variety of applications. MSCNN is a multi-scale CNN model which uses fully connected layers on top of multiple single-scale CNN models which responsible for feature extraction at different scales. In this way, the features extracted at multiple scales work together to assess image quality. Extensive experiments are conducted on the LIVE database and the experimental results validate the effectiveness of our MSCNN index compared to typical FR-IQA and state-of-the-art NR-IQA metrics. Furthermore, the results of cross database experiments show that MSCNN has good generalization ability.

Keywords: Image quality assessment · Convolutional neural networks · Multi-scale

1 Introduction

People are increasingly fond of using images to share information. Images are usually distorted during image acquisition, compression, storage, transmission, and display. The massive amount of images in the internet makes it impossible to assess the qualities of all images using a subjective method. So, it is necessary to develop objective image quality assessment (IQA) metrics for image big data. Objective IQA becomes an important research focus in the fields of big data and image processing.

In general, there are three types of objective IQA: full-reference IQA (FR-IQA) [1–3], reduced-reference IQA (RR-IQA) [4–6], and no-reference IQA (NR-IQA) [7–9]. With full information of reference image, FR-IQA metrics can achieve a high

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1202–1216, 2019.
https://doi.org/10.1007/978-3-030-22868-2_84

fridaya@uef.fi

correlation with human perception. However, it is very expensive or even impossible to get full or partial information of the reference image in many practical applications, especially in big data applications. RR-IQA only uses partial information of the original reference image. Without relying on any pristine knowledge of the reference image, NR-IQA has more practical application prospect.

NR-IQA algorithms have three common trends: distortion-specific approaches [10, 11], natural scene statistics (NSS) approaches [12–14] and training-based approaches [15–17]. The distortion-specific approaches assume that the distorted image is afflicted by a particular type of distortion and design a specific distortion model to assess its quality. The NSS-based algorithms are designed to assess the quality of distorted image according to a distance between the NSS properties of the distorted image and those of distortion-free images. The training-based approaches use machine learning algorithms to train a model to assess the quality of a distorted image from the features extracted from the distorted image.

Recently, the convolutional neural networks (CNN) has been successfully introduced into the fields of big data and compute vision, and has become a new solution to IQA problem [18, 19, 26, 27]. Fu et al. proposed a CNN-based NR-IQA model to predict a quality score for multiply distorted images [18]. Zhang et al. proposed a NR-IQA model to predict a quality score for stereoscopic images [19]. Existing CNN-based NR-IQA models uses single-scale architectures which focus on local spatial information of distorted images but ignore global spatial structure. Multi-scale mechanism has boosted the performances of many singe-scale methods in a variety of applications, including visual saliency detection [38, 39], image segmentation [40, 41], and IQA [20, 21]. Previous studies have shown that the IQA based on multiple image scale can effectively improve performance on single image scale. However, multi-scale mechanism has not yet been used in CNN-based NR-IQA.

In this paper, we propose a NR-IQA model learns from multi-scale CNN (MSCNN). MSCNN is a multi-scale CNN model which uses fully connected layers on top of multiple single-scale CNN (SSCNN) models which responsible for feature extraction at different image scales. In this way, the features extracted at multiple image scales work together to assess image quality. The experimental results demonstrate that extracting and learning from multi-scale features can effectively improve the performance of CNN- based NR-IQA models.

2 Related Work

Recently, NR-IQA has become an advanced research hotspot in IQA because of no need any pristine knowledge of reference images. DIIVINE [7] is a 2-stage framework. DIIVINE first identifies the distortion type of the distorted image and then estimates its quality with respect to the distortion type using a number of NSS features. BLIINDS-II [8] is an efficient single-stage framework. BLIINDS-II relies on a statistical model of DCT coefficients. The DCT features are then fed to the Bayesian probabilistic inference model to assess the quality of the distorted image. BRISQUE [9] uses a statistical model of locally normalized luminance coefficients in the spatial domain to assess image quality.

fridaya@uef.fi

On the other hand, CNN have achieved great successful on various computer vision tasks, especially in visual recognition [22–25]. Instead of using traditional hand-crafted features, such deep networks are capable of extracting features from raw image automatically. In particular, Kang et al. first proposed a NR-IQA framework based on CNN (IQA-CNN) [26]. Without using traditional hand-crafted features, IQA-CNN takes image patches as input to evaluate each patch's score and averages patch scores to obtain the distorted image's quality estimation. Kim et al. proposed a BIQA based on CNN (BIECON) [27]. BIECON is a 2-stages framework. BIECON first uses FR-IQA metrics to compute local quality maps to serve as the regression targets for CNN and then incorporates pooling into the final regression stage.

What more, many existing IQA methods had demonstrated that using multi-scale mechanism can significant improve the performance of visual quality recognition. Wang et al. proposed a multi-scale structural similarity (MS-SSIM) [20] method for IQA. The extension from single-scale to multi-scale methods can provide more viewing conditions. Gu et al. used a multi-scale model to perceive distorted image quality at different scales [21]. The overall quality score is calculated by weighting the scores on different scales.

3 MSCNN Model for NR-IQA

MSCNN model combines multiple SSCNN models which responsible for feature extraction at different image scales. In this way, local and global features extracted at multiple image scales work together to assess image quality. The features extracted from the multiple SSCNN are fed into the three fully-connected layers to learn the quality scores.

We use local normalization as a pre-processing step, and then split the normalized image into patches with overlapping. The MSCNN model is trained over each image patch to generate patch score predictions. Then the test image's quality score is computed by averaging all the predicted scores of the patches belong to the test image. By combining the advantages of features learned by multiple SCCNN models on different image scales, the proposed MSCNN model can effectively improve the performance of CNN-based NR-IQA.

The remainder of the section is organized as follows. Sub-section A presents the architecture of SSCNN model. Sub-section B extends the architecture of SSCNN to that of MSCNN, which uses multi-scale mechanism to boost the performance of NR-IQA.

3.1 Architecture of SSCNN Model

The proposed architecture of SSCNN has eight layers, as shown in Fig. 1. The input is locally normalized n $\times$ n image patches. The size of 32 $\times$ 32 has been widely used in CNN-based IQA [18, 19, 26, 27]. There are two convolutional layers. The first convolutional layer filters the input with 32 3 $\times$ 3 kernels, followed by 32 3 $\times$ 3 max pooling layers. The second convolutional layer contains 64 3 $\times$ 3 kernels, followed by 64 3 $\times$ 3 max pooling and average pooling layers. Then, four fully connected layers

fridaya@uef.fi

are employed to learn quality score of the input image patch. Then, we average all the scores of the patches from a same image to obtain the final quality estimation of the image.

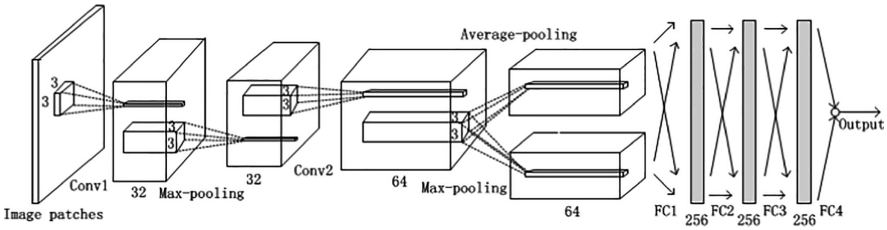


Fig. 1. The architecture of the proposed SSCNN model

In the following subsections, we give the details of image normalization, image patch generation, convolutional and poling layers, fully connected layers, and training method.

Local Normalization. The work in [28] shows that a decorrelating effect for the log contrast can be achieved by first removing local mean displacements and then normalizing the local variance. Mittal et al. utilized this non-linear transformation as a preprocessing step in their computational IQA index. It has shown that the statistical properties that changed by image distortion can be validly captured by local normalization [29]. Local normalization has shown its effectiveness in a variety of IQA metrics [7, 26, 27, 29]. So local normalization method is also adopted in our MSCNN model. Local normalized value is computed as follows:

$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + \mathrm{C}}, \tag{1}$$

$$\mu(i,j) = \sum_{k=-K}^{K} \sum_{l=-L}^{L} \omega(k,l) \cdot I(i+k, j+l), \tag{2}$$

$$\sigma(i,j) = \sqrt{\sum_{k=-K}^{K} \sum_{l=-L}^{L} \omega(k,l) \cdot [I(i+k, j+l) - \mu(i,j)]^2}, \tag{3}$$

where $I(i, j)$ is the intensity value of a pixel at location (i, j) of image I, $\omega = \{\omega(k,l) | k = -K, \ldots, K, l = -L, \ldots, L\}$ indicates a 2D circularly-symmetric Gaussian weighting function. C is a positive constant that prevents dividing by zero. K and L are the height and width of the normalization window. In our implementation, C = 0.1, K = 3 and L = 3.

fridaya@uef.fi

Image Patch Generation. The locally normalized image is divided into many n × n image patches with overlapping for two reasons. First, typical IQA databases only contain limited numbers of images, which may be insufficient to train a CNN-based IQA model directly using the images in the databases. The images of these databases have homogeneous distortions, so we divide the image into patches and use the subjective quality score for each image as the scores for the patches from it as in [26]. Image patch generation plays an important role of data augmentation which can effectively increase the sizes of IQA databases and make it possible to learn a CNN-based IQA model. In addition, the image patches are augmented by horizontal reflection to further increase the numbers of patches. Second, the databases may contain images of various sizes, which prohibit using fully-connected layers in the CNN model. Since the fully-connected layers are required in our CNN model to learn the quality score of the image, we need to divide the images into a number of image patches of the same size.

Convolutional and Pooling Layers. As shown in Fig. 2, the proposed architecture of SSCNN model has two convolutional layers followed by pooling layers. In the convolutional layer, the locally normalized image patches are convolved with a number of filters and each filter generates a feather map. Pooling operation is a sample-based discretization process. After convolution operation, pooling operation is applied to each feature map to reduce its dimensionality. Pooling is done to in part to help over-fitting by providing an abstracted form of the representation. As well, pooling reduces the computational cost by reducing the number of parameters to learn and introduces basic invariances (e.g. translation, rotation, scale) so that irrelevant variations in the input affect the output smoothly.

Let R_{ij} denotes the local pooling region around position (i, j), u_{ij}^{k} and v_{ij}^{k} are the output of the max and average pooling operators related to the k-th feature map at position (i, j), respectively. The max and average values u_{ij}^{k} and v_{ij}^{k} are computed by:

$$u_{ij}^{k} = \max_{(p,q)\in R_{ij}} (x_{pq}^{k}), \tag{4}$$

$$v_{ij}^{k} = \operatorname{average}_{(p,q)\in R_{ij}} (x_{pq}^{k}), \tag{5}$$

where k = 1,2,…, K and K is the number of feature maps, x_{pq}^{k} is the element at (p, q) within the pooling region R_{ij}.

The first convolutional layer is followed by a 3 × 3 max pooling layer and the second convolutional layer is followed by a 3 × 3 max pooling and a 3 × 3 average pooling layers. Max and average pooling operations decrease the size of each input feature map by outputting only the maximum and average values for a local pooling region of 3 × 3 of the input feature map, respectively. Generally speaking, the max pooling operation can retain more texture information, while the average pooling operation can retain more background information. So we introduce average pooling

fridaya@uef.fi

layer into our SSCNN model and found it can effectively boost the performance of our model.

Fully Connected Layers. There are four fully connected layers in the SSCNN architecture. Except for the last fully connected layer which is used to compute the score for each patch, the other three fully connected layers are activated through activation function. Activation function performs a non-linearity operation. There are some common activation functions, such as Sigmoid, the Rectified Linear Units (ReLU), Leaky ReLU, Tanh, etc. Compared to Sigmoid and Tanh, ReLU achieve nonlinearity by simply thresholding a matrix of activations at zero. ReLU can effectively reduce the likelihood of the gradient to vanish and greatly accelerate the process of convergence when using stochastic gradient descent (SGD) method [22]. In our experiments, ReLU is used in the first three fully connected layers.

Training Method. We train our SSCNN network using n × n image patches generated from the distorted images in IQA databases. Since the distorted images of these databases have homogeneous distortions, so use the subjective quality score for each distorted image as the scores for the patches from it. The quality scores for the image patches from each test distorted image are first predicted and then averaged as the final quality score for each test image.

During the training stage, we define the loss function as mean square error (MSE):

$$\mathrm{L} = \frac{1}{N}\sum_{n=1}^{N} \|f(x_n;\omega) - S_n\|^2, \tag{6}$$

where N is the number of image patches during the training stage, x_n is the n-th image patch, S_n is its corresponding subjective score, $f(x_n;\omega)$ is the predicted quality score of x_n via the SSCNN model with weights ω. By minimizing the loss function, the optimal weights of SSCNN model can be learned via back-propagation.

During the training stage, we put 128 random chosen image patches into one mini-batch, which can enhance the generalization ability of CNN. In addition, we accelerate training by employing the adaptive moment estimation optimizer (Adam) [31]. The momentum constant is set to be 0.9. The initial learning rate is set to be 3×10^{-5}.

3.2 Architecture of MSCNN Model

The design of MSCNN model follows the practical intuition that visual information should be processed at various scales and then aggregated so that the model can abstract features from different scales simultaneously. Figure 2 shows the architecture of the proposed MSCNN model which is a multi-scale network. The architecture for each scale in the multi-scale network is the same as the SSCNN model as shown in Fig. 1. MSCNN uses multiple CNN with identical structure to learn the information from image patches in different scales.

fridaya@uef.fi

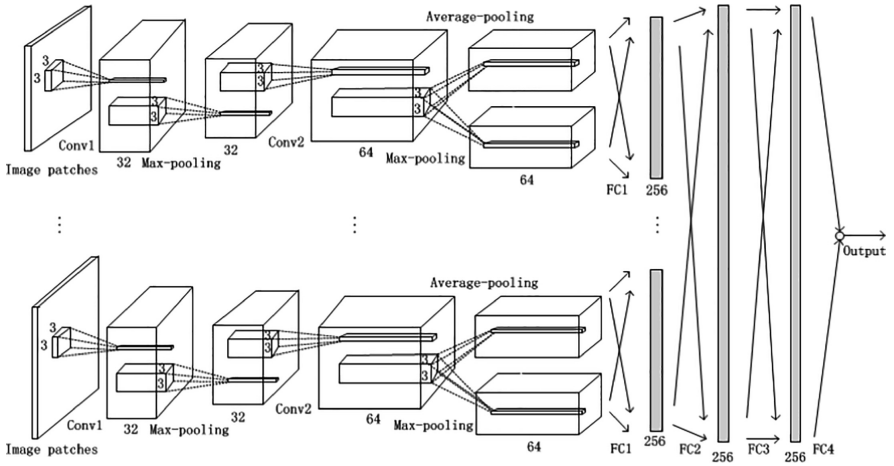


Fig. 2. The architecture of the proposed MSCNN model

A small image patch, such as a patch of 32 × 32, keeps more local information, but may not able to represent more global spatial structure. On the contrary, a big image patch, such as a patch of 80 × 80, has more global spatial structure information, but may mask some local information. The first fully connected layer of each SSCNN generates a 256 dimensional vector. Across different scales, multiple vectors are concentrated into a 256 × s (s is the number of different scales) dimensional vector and feed into three fully connected layers to learn the quality scores. Therefore, the proposed multi-scale architecture can jointly consider the image patches in different image scales. In the experiments, the proposed MSCNN model learns from three image scales, including 32 × 32, 48 × 48, and 80 × 80.

4 Experiments

In this section, we evaluate and compare the proposed MSCNN model in terms of both performance and generalization ability.

4.1 Experimental Settings

The following three databases are used in our experiments:

LIVE Database [30]**.** This IQA database contains 779 distorted images. The number of the reference images is 29. There are five different distortion types: JPEG compression (JPEG), Fast Fading (FF) JPEG2000 compression (JPEG2K), Gaussian blur (BLUR) and White Gaussian (WN) at different distortion levels. Differential Mean Opinion Score (DMOS), roughly from 0 to 100, is provided as the subjective quality score for each image. A higher DMOS value indicates a lower visual quality of the image.

fridaya@uef.fi

LIVE Multiply Distorted (LIVEMD) Database [32]**.** This IQA database contains two subsets of multiply distorted images. The first subset includes images with blur followed by JPEG (B&J). The second subset includes images with blur followed by noise (B&N). There are totally 15 reference images and 225 distorted images. Each distorted image in LIVEMD has a DMOS in the range from 0 to 100.

TID2008 Database [33]**.** There are 25 reference images and 1700 distorted images. The numbers of distortion types and levels are 17 and 4, respectively. Each image has a Mean Opinion Score (MOS) which ranges from 0 to 9. A higher MOS value indicates a higher visual quality of the image.

As suggested in [34], Spearman's Rank Order Correlation Coefficient (SROCC) and Pearson's Linear Correlation Coefficient (LCC) are adopted to evaluate the performance of IQA metrics. Before computing LCC, the predicted scores were fitted using an acknowledged logistic nonlinearity function as follows:

$$\mathrm{f}(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2 (\mathrm{x} - \beta_3)}} \right) - \beta_4 x + \beta_5, \tag{7}$$

where x indicates the predicted scores, f(x) indicates the mapped scores, and β_{i} (i = 1, 2,…, 5) are five parameters to be estimated during the fitting process. Given two quantities, the linear dependence and the monotonic are measured by LCC and SROCC, respectively. For SROCC and LCC, a large value, especially when closing to 1, indicates good performance.

We report the median results obtained from 20 train-test iterations. In each iteration, the images in the database are divided into two randomly chosen subsets, 80% of reference images together with their distorted images are the training set and the remaining images are the testing set, such that there is no overlap between training and testing sets. A symbol "—" means that the values are not provided in the papers.

4.2 Performance Evaluation

We compare the proposed MSCNN model to seven classical FR-IQA metrics (PSNR, FSIM [2], SSIM [3], ADD-GSIM [21], PSIM [35], MS-SSIM [20], and ParaBoost [36]) and six state-of-the-art NR-IQA metrics (BRISQUE [9], SESANIA [37], BLIINDS-II [8], IQA-CNN [26], BIECON [27] and DIIVINE [7]).

We first conduct both distortion-specific experiments and non-distortion-specific experiments on LIVE database and given the experimental results in Tables 1 and 2. As shown in Tables 1 and 2, the proposed SSCNN using image patches of 80 × 80 outperforms the compared NR-IQA metrics and is comparable to the classical FR-IQA metrics in the non-distortion-specific experiments. SSCNN outperforms the compared NR-IQA metrics on all the distortion types except JPEG in the distortion-specific experiments. The experimental results validate the effectiveness of the proposed SSCNN architecture. As shown in the last rows of Tables 1 and 2, the proposed MSCNN model can utilize both local and global information and further improve the performance. Specifically, the MSCNN model outperforms all the compared NR-IQA

fridaya@uef.fi

Table 1. SROCC of IQA metrics on LIVE database. The best three values are formatted in boldface.

Method	FF	WN	JPEG	JP2K	BLUR	ALL
PSNR	0.891	**0.985**	0.881	0.895	0.782	0.876
MS-SSIM	0.939	0.977	0.979	0.963	0.954	0.954
FSIMc	**0.950**	0.965	**0.983**	**0.972**	**0.971**	**0.965**
ADD-GSIM	**0.944**	**0.979**	**0.980**	**0.968**	**0.965**	**0.968**
SSIM	**0.956**	0.969	0.976	0.961	0.952	0.948
PSIM	0.942	**0.986**	**0.979**	**0.968**	**0.964**	0.962
ParaBoost	—	—	—	—	—	**0.984**
BRISQUE	0.877	0.979	0.965	0.914	0.951	0.940
BLIINDS-II	0.889	0.969	0.942	0.929	0.923	0.931
IQA-CNN	0.908	0.978	**0.977**	**0.952**	**0.962**	0.956
BIECON	**0.923**	0.980	**0.974**	**0.952**	0.958	**0.961**
SESANIA	0.881	0.931	0.929	0.886	0.941	0.934
DIIVINE	0.863	**0.984**	0.910	0.913	0.921	0.916
SSCNN	**0.957**	**0.988**	0.963	**0.967**	**0.979**	**0.972**
MSCNN	**0.961**	**0.988**	**0.980**	**0.973**	**0.969**	**0.974**

Table 2. LCC of IQA metrics on LIVE database. The best three values are highlighted formatted in boldface.

Method	FF	WN	JPEG	JP2K	BLUR	ALL
SSIM	**0.951**	0.982	0.946	**0.964**	0.900	0.945
PSIM	**0.971**	**0.987**	**0.977**	0.954	**0.958**	0.958
FSIMc	0.912	0.976	**0.985**	0.910	**0.978**	**0.961**
ADD-GSIM	**0.957**	**0.976**	0.953	**0.960**	0.958	**0.966**
MS-SSIM	0.949	**0.988**	**0.979**	**0.975**	**0.965**	0.951
ParaBoost	—	—	—	—	—	**0.984**
PSNR	0.880	0.917	0.903	0.876	0.780	0.872
BLIINDS-II	0.896	0.980	0.968	0.935	0.938	0.930
BIECON	**0.931**	0.970	**0.987**	**0.965**	0.945	**0.962**
IQA-CNN	0.917	0.984	0.981	0.953	0.968	0.953
SESANIA	0.920	0.981	0.973	0.954	**0.975**	0.948
DIIVINE	0.888	0.988	0.921	0.922	0.923	0.917
BRISQUE	0.903	**0.985**	0.973	0.923	0.951	0.942
SSCNN	**0.957**	**0.990**	**0.992**	**0.982**	**0.981**	**0.976**
MSCNN	**0.964**	**0.995**	**0.993**	**0.984**	**0.976**	**0.978**

fridaya@uef.fi

metrics in both the distortion-specific and the non-distortion-specific experiments. MSCNN model outperforms SSCNN except for distortion BLUR.

For TID2008 database, we only use the first 13 global distortions since the last 4 distortions (mean shift, local block-wise distorted, contrast change, non-eccentricity pattern noise) are not very heterogeneous to intensity change and the computer-generated image was excluded. As shown in Table 3, the proposed MSCNN outperforms compared NR-IQA metrics on AGN, ANMC, SCN, MN, HFN, IMN, QN, and J2TE in the distortion-specific experiments and has best performance on non-distortion-specific experiments.

Table 3. SROCC of IQA metrics on TID2008 database. The best three values are formatted in boldface.

Method	ANMC	QN	HFN	SCN	MN	AGN	IMN
PSNR	**0.916**	**0.900**	**0.938**	**0.931**	**0.871**	**0.923**	**0.918**
SSIM	0.829	**0.910**	0.868	0.858	0.821	0.843	**0.766**
MS-SSIM	0.823	0.896	0.879	0.847	0.821	0.837	0.706
DIIVINE	0.844	0.874	0.922	0.854	0.713	0.812	**0.915**
ADD-GSIM	**0.873**	0.871	**0.909**	**0.898**	**0.862**	**0.900**	0.719
FSIMc	0.859	0.887	0.909	0.854	0.836	0.867	**0.747**
PSIM	**0.910**	**0.912**	**0.921**	**0.937**	**0.860**	**0.924**	0.735
BIECON	0.835	0.893	**0.931**	0.903	0.835	**0.913**	0.913
BRISQUE	**0.861**	0.881	**0.931**	0.885	0.810	0.853	**0.927**
BLIINDS-II	0.807	0.851	0.917	0.887	0.691	0.779	0.908
SESANIA	0.832	**0.895**	0.893	**0.905**	**0.837**	**0.913**	0.910
SSCNN	**0.876**	**0.916**	**0.946**	**0.909**	**0.861**	**0.916**	0.911
MSCNN	**0.901**	**0.934**	**0.947**	**0.933**	**0.874**	**0.923**	**0.938**

Method	J2TE	DEN	GB	JP2K	JPEG	JGTE	ALL
PSNR	0.783	0.937	0.934	0.882	0.922	0.801	0.677
SSIM	0.855	0.948	**0.943**	0.867	0.922	0.830	0.734
MS-SSIM	**0.873**	0.970	**0.963**	0.969	**0.941**	0.873	0.890
DIIVINE	0.852	0.912	0.943	0.938	0.930	0.873	0.901
ADD-GSIM	**0.877**	**0.973**	0.934	**0.982**	**0.947**	**0.883**	**0.926**
FSIMc	0.860	**0.971**	**0.955**	**0.977**	0.927	**0.874**	**0.923**
PSIM	**0.885**	**0.981**	0.920	**0.986**	**0.933**	**0.894**	**0.935**
BIECON	0.827	**0.917**	**0.953**	**0.956**	**0.943**	0.862	**0.923**
BRISQUE	0.836	**0.924**	0.933	**0.944**	**0.934**	**0.891**	0.908
BLIINDS-II	**0.855**	0.908	0.952	0.940	0.928	0.865	0.898
SESANIA	0.872	**0.959**	**0.963**	**0.944**	0.931	**0.884**	0.872
SSCNN	**0.895**	0.899	0.941	0.931	0.927	0.874	**0.923**
MSCNN	**0.900**	0.885	**0.955**	**0.952**	**0.935**	**0.883**	**0.927**

fridaya@uef.fi

A method which works well for a distorted image with a single distortion type is not necessarily suitable for multiply distorted images. To validate the performance of the proposed MSCNN model for multiply distorted IQA, we conduct experiments on LIVEMD database. Since the experiments result on LIVEMD database are not provided in the corresponding reference paper of ParaBoost, IQA-CNN, and BIECON, and we also can't find the corresponding source code, these metrics are not compared in our experiments on LIVEMD database. As shown in Table 4, the MSCNN model outperforms all the compared classical FR-IQA metrics and state-of-art NR-IQA metrics in terms of both SROCC and LCC. The results show that MSCNN model also works well for multiply distorted images.

Table 4. SROCC and LCC of IQA metrics on LIVEMD database.

Method	SROCC			LCC		
	B&N	B&J	ALL	B&N	B&J	ALL
PSIM	0.846	0.850	0.851	0.872	**0.899**	0.885
ADD-GSIM	0.776	0.847	0.802	0.833	0.889	0.854
MS-SSIM	**0.863**	0.837	0.837	**0.892**	0.885	0.875
SSIM	0.695	0.738	0.642	0.770	0.792	0.734
FSIMc	0.862	**0.854**	**0.862**	0.876	0.905	**0.890**
PSNR	0.669	0.643	0.626	0.748	0.706	0.693
BRISQUE	0.890	0.895	0.895	0.885	0.904	0.888
SESANIA	0.871	0.842	0.836	0.864	0.863	0.838
DIIVINE	0.612	0.726	0.669	0.684	0.798	0.731
BLIINDS-II	0.868	0.854	0.835	0.859	0.859	0.820
MSCNN	**0.907**	**0.916**	**0.908**	**0.928**	**0.936**	**0.929**

4.3 Cross Database Test

To evaluate the generalization ability of our proposed MSCNN model, we train MSCNN model on LIVE database and test it on TID2008 database. The TID2008 database consists of 25 reference images – 24 natural images and one computer-generated image. Since our model is trained on the natural images, we excluded the computer-generated image from the TID2008 database. Since there are four types of distortions shared by LIVE and TID2008, i.e., White Gaussian (WN), JPEG compression (JPEG), JPEG2000 compression (JPEG2 K), and Gaussian blur (BLUR), we test MSCNN model only on the images distorted by these four distortion types. Experimental results of the cross database test are shown in Table 5. It shows that MSCNN model outperforms previous state-of-the-art methods and the performance does not depend on the database.

fridaya@uef.fi

Table 5. SROCC of IQA metrics obtained by training on LIVE database and testing on images distorted by 4 shared distortion types of TID2008 database. The best three values are formatted in boldface.

Method	JPEG	BLUR	JP2K	WN	ALL
PSIM	**0.933**	0.920	**0.986**	**0.924**	**0.957**
SSIM	0.922	**0.943**	0.867	0.843	0.767
MS-SSIM	**0.941**	**0.963**	0.969	0.837	0.910
FSIMc	0.928	**0.953**	**0.977**	0.877	**0.952**
ADD-GSIM	**0.947**	0.934	**0.982**	**0.900**	**0.961**
PSNR	0.922	0.934	0.882	**0.923**	0.824
IQA-CNN	—	—	—	—	**0.920**
BRISQUE	**0.924**	**0.881**	0.832	0.829	0.896
BIECON	**0.941**	**0.913**	0.878	**0.842**	**0.923**
BLIINDS-II	0.890	0.850	**0.916**	0.660	0.844
DIIVINE	0.866	0.862	**0.924**	**0.851**	0.889
MSCNN	**0.944**	**0.878**	**0.890**	**0.923**	**0.928**

4.4 Effects of Patch Size and the Proposed Multi-scale Architecture

Since all image patches sampled from an image share the same score during the training stage, we investigate how the patch size affects the performance of single-scale network architecture. As shown in Table 6, five different patch sizes (32, 48, 64, 80, and 96) were compared on LIVE database and the best performance values of NR-IQA metrics are formatted in bold. We can find that in general a bigger patch size results in a better performance. Because that the small image patches focus on local information which cannot suitable present the whole image's distorted information. The network training from big image patches has more global spatial structure information than the small patches and will get a better result. And the performance increases slightly as the patch size increases from 80 to 96. Moreover, a bigger patch size has larger number of parameters to learn and will increase the computational cost exponentially. Therefore we chose the patch size of 80 in the experiments of single-scale network.

Table 6. SROCC, LCC, KROCC, RMSE and parameters on different patch size.

SIZE	32	48	64	80	96
SROCC	0.9676	0.9702	0.9726	**0.9734**	0.9731
LCC	0.9695	0.9734	0.9758	0.9764	**0.9767**
KRCC	0.8571	0.8662	0.8718	**0.8742**	0.8730
RMSE	7.5473	7.0499	6.6940	6.6213	**6.6171**
Parameters (million)	1.4	3.5	6.7	10.9	16.1

fridaya@uef.fi

To find the best architecture of different scale image patches, we conduct contrast experiments of two, three and four different scales combination. As shown in Table 7, the performance of two different scales architecture is better than the single model of small scale but a little worse than the single model of big patch. It's reasonable that the bigger patch, which contains more information of global spatial structure, has a larger proportion in IQA and the small patch of two scale architecture limit the overall performance. And three different scales architecture contains more information of different scales and can significantly improve the performance. Especially, the combination of 32×32, 48×48, and 80×80, which has more local information and less parameter than other three different scales architecture, get the best result. And, the four different scales architecture fails to improve the performance since it is hard to train a good model with too many parameters to optimize.

Table 7. SROCC, LCC, KROCC, RMSE and parameters on different multi-scale architecture.

SIZE	32 + 80	48 + 80	64 + 80	32 + 48 + 80	32 + 64 + 80	48 + 64 + 80
SROCC	0.9729	0.9726	0.9729	**0.9740**	0.9739	0.9738
LCC	0.9764	0.9762	0.9761	**0.9778**	0.9770	0.9773
KROCC	0.8721	0.8718	0.8731	0.8743	0.8748	**0.8761**
RMSE	6.6292	6.6304	6.6201	**6.3977**	6.5316	6.5217
Parameters (million)	12.1	14.2	17.3	15.3	18.5	20.6

5 Conclusion

In this paper, we proposed a multi-scale CNN architecture for NR-IQA named MSCNN. Most existing CNN-based NR-IQA algorithms only takes raw image patches as input to evaluate each patch's score and averages patch scores to obtain the distorted image's quality estimation, which pay more attention on local information but ignore obvious global spatial structure. Aiming to solve this issue, a new multi-scale architecture with three columns CNN is designed to aggregate the features which extracted from a small scare image patch and a big scare image patch. By fully utilizing the features from bath local information and global spatial structure, MSCNN has stronger generality and can further improve the performance of CNN-based IQA. Extensive experiments on LIVE dataset validate the effectiveness of our MSCNN relative to classical FR and state-of-the-art NR IQA approaches. Moreover, experiments in cross dataset show that MSCNN has excellent generalization ability. Future work will focus on deeper CNN structure and more reasonable combination of different scale image patches.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant 61672158, Grant 61672159, Grant 61502105, and Grant 61300102, in part by the Fujian Natural Science Funds for Distinguished Young Scholar under Grant 2015J06014, in part by the Technology Guidance Project of Fujian Province under Grant No. 2017H0015, and in part by the Fujian Collaborative Innovation Center for Big Data Application in Governments.

fridaya@uef.fi

References

1. Wang, Z., Bovik, A.C., Sheikh, H.R., et al.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
2. Zhang, L., Zhang, L., Mou, X., et al.: FSIM: a feature similarity index for image quality assessment. IEEE Trans. Image Process. **20**(8), 2378–2386 (2011)
3. Zhang, L., Shen, Y., Li, H.: VSI: a visual saliency-induced index for perceptual image quality assessment. IEEE Trans. Image Process. **23**(10), 4270–4281 (2014)
4. Li, Q., Wang, Z.: Reduced-reference image quality assessment using divisive normalization-based image representation. IEEE J. Sel. Top. Sig. Process. **3**(2), 202–211 (2009)
5. Soundararajan, R., Bovik, A.C.: RRED indices: reduced reference entropic differencing for image quality assessment. IEEE Trans. Image Process. **21**(2), 517–526 (2012)
6. Rehman, A., Wang, Z.: Reduced-reference image quality assessment by structural similarity estimation. IEEE Trans. Image Process. **21**(8), 3378–3389 (2012)
7. Moorthy, A.K., Bovik, A.C.: Blind image quality assessment: from natural scene statistics to perceptual quality. IEEE Trans. Image Process. **20**(12), 3350–3364 (2011)
8. Saad, M.A., Bovik, A.C., Charrier, C.: Blind image quality assessment: a natural scene statistics approach in the DCT domain. IEEE Trans. Image Process. **21**(8), 3339–3352 (2012)
9. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process. **21**(12), 4695–4708 (2012)
10. Wang, S., Deng, C., Zhao, B., et al.: Gradient-based no-reference image blur assessment using extreme learning machine. Neurocomputing **174**, 310–321 (2016)
11. Zhan, Y., Zhang, R.: No-reference JPEG image quality assessment based on blockiness and luminance change. IEEE Sig. Process. Lett. **24**(6), 760–764 (2017)
12. Gao, X., Gao, F., Tao, D., et al.: Universal blind image quality assessment metrics via natural scene statistics and multiple kernel learning. IEEE Trans. Neural Netw. Learn. Syst. **24**(12), 2013–2026 (2013)
13. Wang, Q., Chu, J., Xu, L., et al.: A new blind image quality framework based on natural color statistic. Neurocomputing **173**, 1798–1810 (2016)
14. Zhang, Y., Wu, J., Xie, X., et al.: Blind image quality assessment with improved natural scene statistics model. Digit. Sig. Process. **57**, 56–65 (2016)
15. Narwaria, M., Lin, W.: Objective image quality assessment based on support vector regression. IEEE Trans. Neural Netw. **21**(3), 515–519 (2010)
16. Oszust, M.: No-reference image quality assessment using image statistics and robust feature descriptors. IEEE Sig. Process. Lett. **24**(11), 1656–1660 (2017)
17. Zhou, W., Yu, L., Qiu, W., et al.: Local gradient patterns (LGP): an effective local-statistical-feature extraction scheme for no-reference image quality assessment. Inf. Sci. **397**, 1–14 (2017)
18. Fu, J., Wang, H., Zuo, L.: Blind image quality assessment for multiply distorted images via convolutional neural networks. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1075–1079. IEEE (2016)
19. Zhang, W., Qu, C., Ma, L., et al.: Learning structure of stereoscopic image for no-reference quality assessment with convolutional neural network. Pattern Recogn. **59**, 176–187 (2016)
20. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, vol. 2, pp. 1398–1402. IEEE (2003)
21. Gu, K., Wang, S., Zhai, G., et al.: Analysis of distortion distribution for pooling in image quality prediction. IEEE Trans. Broadcast. **62**(2), 446–456 (2016)

fridaya@uef.fi

22. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
23. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
24. Szegedy, C., Liu, W., Jia, Y., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
25. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
26. Kang, L., Ye, P., Li, Y., et al.: Convolutional neural networks for no-reference image quality assessment. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1733–1740 (2014)
27. Kim, J., Lee, S.: Fully deep blind image quality predictor. IEEE J. Sel. Top. Sig. Process. **11**(1), 206–220 (2017)
28. Ruderman, D.L.: The statistics of natural images. Netw. Comput. Neural Syst. **5**(4), 517–548 (1994)
29. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process. **21**(12), 4695–4708 (2012)
30. Sheikh, H.R., Wang, Z., Cormack, L., et al.: LIVE image quality assessment database release 2 (2005). 2007-07-17. http://live.ece.utexas.edu/research/quality
31. Hinton, G.E., Srivastava, N., Krizhevsky, A., et al.: Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580 (2012)
32. Jayaraman, D., Mittal, A., Moorthy, A.K., et al.: Objective quality assessment of multiply distorted images. In: 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), pp. 1693–1697. IEEE (2012)
33. Ponomarenko, N., Lukin, V., Zelensky, A., et al.: TID2008-a database for evaluation of full-reference visual quality assessment metrics. Adv. Modern Radioelectronics **10**(4), 30–45 (2009)
34. Video Quality Experts Group: Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, Phase II (FR_TV2). ftp://ftp.its.bldrdoc.gov/dist/ituvidq/Boulder_VQEG_jan_04/VQEG_PhaseII_FRTV_Final_Report_SG9060E.doc (2003)
35. Gu, K., Li, L., Lu, H., et al.: A fast reliable image quality predictor by fusing micro-and macro-structures. IEEE Trans. Ind. Electron. **64**(5), 3903–3912 (2017)
36. Liu, T.J., Liu, K.H., Lin, J.Y., et al.: A paraboost method to image quality assessment. IEEE Trans. Neural Netw. Learn. Syst. **28**(1), 107–121 (2017)
37. Li, Y., Po, L.M., Xu, X., et al.: No-reference image quality assessment with shearlet transform and deep neural networks. Neurocomputing **154**, 94–109 (2015)
38. Li, G., Yu, Y.: Visual saliency based on multiscale deep features. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5455–5463 (2015)
39. Li, G., Yu, Y.: Deep contrast learning for salient object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 478–487 (2016)
40. Li, H., Song, Y., Chen, C.L.P.: Hyperspectral image classification based on multiscale spatial information fusion. IEEE Trans. Geosci. Remote Sens. **55**(9), 5302–5312 (2017)
41. Zheng, S., Fang, B., Li, L., et al.: B-Spline based globally optimal segmentation combining low-level and high-level information. Pattern Recogn. **73**, 144–157 (2018)

fridaya@uef.fi

Motion Detection with IoT-Based Home Security System

Pei Kee Tiong, Nur Syazreen Ahmad(✉), and Patrick Goh

School of Electrical and Electronic Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia
syazreen@usm.my

Abstract. This paper presents a development of motion detection technique with an IoT-based Home Security System. Two low-cost motion detectors, namely, pyroelectric infrared radial (PIR) and microwave sensors are considered in this work, and a Raspberry Pi is selected as the main platform to receive the sensor signals and send a push notification to the user whenever an intruder is detected. The performance of each sensor was firstly analysed in terms of its detection accuracy and false activation. A series of experiments were designed in a car porch area with various walking trajectories to evaluate the accuracy of each sensor, and three environmental scenario with different wind speeds were considered for the evaluation on false triggers. We also introduce a simple algorithm for each sensor to minimize the effects from the sensors' noise. Based on the experimental results, it is shown that the microwave sensor outperforms the PIR in terms of both detection accuracy and false activation. It is also observed that, the implementation of the algorithm offers a great reduction in terms of false activation, but with a slight trade-off in motion detection accuracy. Nevertheless, the overall performance of the microwave sensor is enhanced as the false trigger can be totally eliminated at the cost of only 5% accuracy reduction, compared to that of the PIR sensor which still shows a non-zero false triggers with 6.25% accuracy reduction after the implementation of the algorithm.

Keywords: IoT · Motion detection · Home security system

1 Introduction

Motion sensing is one of the important functional blocks in many home security systems. Various motion sensing techniques have been employed in the existing security systems and they can be categorized into software-based and hardware-based. The software-based motion sensing technique commonly practiced in previous researches is usually embedded in microcomputers with integrated cameras. In order to detect motions, image processing is required such as background subtraction algorithm which is also known as frame differencing or differential comparison [1–3]. This approach can be further enhanced with additional security features including human face detection and tracking [4], making it more

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1217–1229, 2019.
https://doi.org/10.1007/978-3-030-22868-2_85

fridaya@uef.fi

efficient in terms of intruder recognition. Nevertheless, the software-based detection typically involves heavy computations, is relatively more expensive and can raise the concern of violating privacy of residents nearby due to the use of a camera that could bring adverse psychological impact to them [5,6].

Hardware-based motion sensing technique, as the name suggests, relies on hardware devices, namely motion detectors or motion sensors. They can be classified into two major types i.e. passive motion sensors and active motion sensors. Passive motion sensors such as pyroelectric infrared radial (PIR) are widely adopted in most of the existing security systems due to their affordable costs and lower power consumptions, making them a great choice particularly for systems with power-constrained designs [7]. A PIR sensor is made up of a crystalline material that generates a voltage change in response to change in infrared radiation, which occurs whenever a human body comes across the sensor's detection area. Within the sensor itself, a pair of sensitive elements are wired as opposite inouts to a differential amplifier to cancel off the changes in ambient infrared radiation. Its sensing coverage is comprised of an array of Fresnel lens where it allows the sensor to be more sensitive to motion within its coverage area by creating multiple cones of views. In some applications such as the work presented in [8], PIR sensors are installed at the building's entrance and a video call is sent to the user through the internet when motion is detected. The user can then check whether the motion presented is an intruder. A quite similar approach is also introduced in [9] but is targeted for indoor environment where the PIR sensor is placed in the room with an additional magnetic switch is positioned at the door to detect any openings. Whenever a motion is detected, a buzzer will be turned on and the image of the intruder is captured and directly sent to the owner via mail.

Active motion sensors have relatively higher power consumptions compared to passive sensors due to their operation principle which regularly emit waves or pulses. Some of the examples from this type are microwave sensors, Doppler radar [10] and ultrasonic sensors which are usually preferable for applications that require larger sensing areas such as collision avoidance strategies of wheeled mobile robots and automatic doors in malls. A microwave sensor utilizes a signal generator that actively generates microwave pulses and a detector that receives and calculates the reflection of pulses from objects. If the reflected patterns are changing continuously, then it is considered there is motion nearby. In the work proposed in [11], an insight of integration of a microwave sensor with a low-cost LED lamp is presented. The microwave sensor is selected over PIR as it can be integrated in the luminaire without additional costs needed on the changes of existing housing designs. In this method, a simple signal processing is applied to the amplified analogue signal, and it is found that the derivative of the captured signal is close to zero when no motion is present, and close to maximal value otherwise.

PIR and microwave sensors have a common feature of being independent of light illumination changes. They are widely employed in many IoT-based security systems due to their easy installation with microcomputers, small form factor,

fridaya@uef.fi

unobtrusive and privacy-preserving interactions. Despite the advantages, their performances are limited because of their defined coverage in terms of detection angle and effective distance [12,13]. Hence, false alarms can be easily triggered and may be inconvenient to the users.

In this work, a development of motion detection technique with an IoT-based Home Security System is presented. We consider both PIR and microwave sensors for performance evaluation in terms of detection accuracy and false activation, and a Raspberry Pi is selected as the main controller to receive the sensor signals and send a push notification to the user whenever an intruder is detected. A series of experiments in a car porch area were designed with different walking trajectories to evaluate the accuracy of each sensor, while for the evaluation on false triggers, three different environmental scenario were considered. We also introduce a simple algorithm for each sensor to minimize the effects from the sensors' noise. The main methodology and experimental setup are further explained in the next section.

2 Methodology and Experimental Setup

Throughout the studies, the experiments were conducted on a car porch with an area of 48 m^2 (8 m width $\times$ 6 m length). Figure 1 shows the top view of the house, marked with entrances of the car porch, the house and the roof coverage.

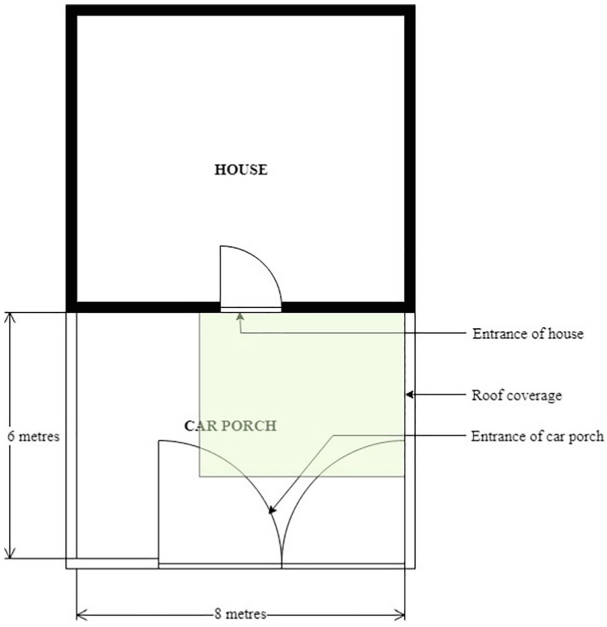
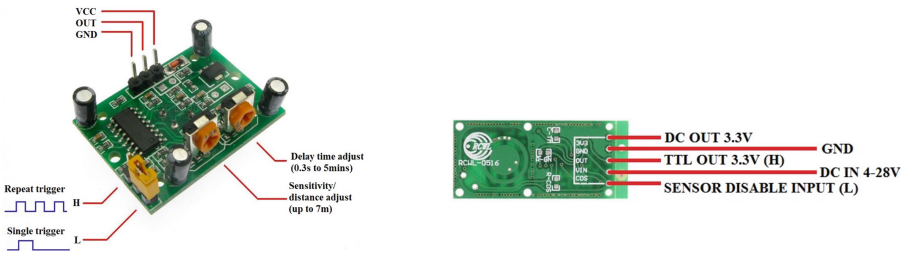


Fig. 1. Top view of the house with car porch as the experimental area

fridaya@uef.fi

In this work, a Raspberry Pi Zero W which runs on operating system of Raspbian Stretch was selected to serve as the main platform for the control algorithms, and we consider two specific types of motion sensors, which are HC-SR501, a PIR sensor (Fig. 2a) and RCWL-0516, a microwave sensor (Fig. 2b). The PIR sensor produces a high output voltage level when a motion is detected, and low otherwise. Two adjustable potentiometers are available on the PIR sensors; one is used to control its maximum sensing distance while the other is used to trigger a delay. The microwave sensor on the other hand does not have any potentiometer to adjust its sensing distance, and it is capable to sense motion from any direction up to 7 m based on its datasheet.



(a) HC-SR501 (b) RCWL-0516

Fig. 2. PIR and microwave sensors.

2.1 Performance Analyses

In order to analyse the detection accuracy for each sensor, four experiments (Exp. 1.1, Exp. 1.2, Exp. 1.3 and Exp. 1.4) with different walking trajectories as depicted in Fig. 3 were carried out. We also introduced an algorithm (Alg. M) as shown in Fig. 4 which acts as a filter to minimize the erratic readings from the sensor.

With regard to the analysis on the false activation rate, three different environmental scenarios with no motion (from any objects, humans or animals) were considered for the experiments: (S1) sunny weather with no wind; (S2) sunny weather with low speed wind (simulated with a standing fan at lowest speed setting); and (S3) sunny weather with high speed wind (simulated with a standing fan at highest speed setting). The standing fan was placed at $(\theta, r) = (90, 3)$, i.e. 3 m away from the sensor. The experiments for (S1), (S2) and (S3) above were named as Exp. 2.1, Exp. 2.2 and Exp. 2.3 respectively and each experiment was conducted for 10 minutes.

2.2 IoT-Based Home Security System

The overview of the IoT-based home security system architecture developed in this work is illustrated as in Fig. 5 which can be divided into two major blocks, the server and the client. Both blocks consist of a Raspberry Pi Zero W as

fridaya@uef.fi

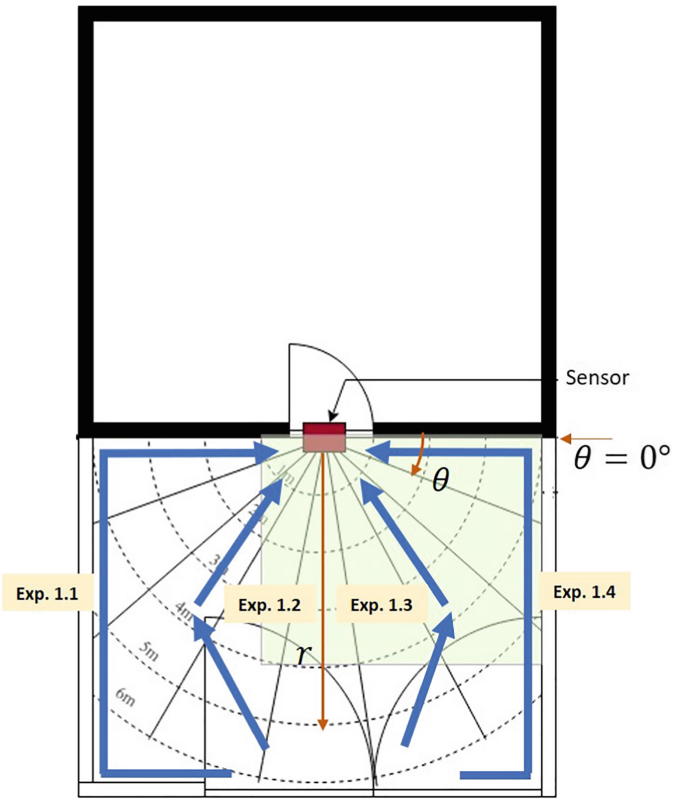


Fig. 3. Experimental procedure for the performance analysis of the motion sensors.

the main controller, together with a built-in bluetooth module and a motion detector sub-block formed by the sensor fusion and Algorithm M. Camera and Wi-Fi modules for user push notification alert are also included as additional sub-functional blocks in the server.

The flowchart for the server module is depicted in Fig. 6a. Upon the execution, system initialization will take place which includes the initialization of the camera module, configuration of GPIO pins, connection of bluetooth socket for server-client communication and settling down of the PIR sensor. Event handlers are then assigned to listen to the motion detection event as described in the previou section, and the server module will continuously check from bluetooth client socket for data receival. Either motion detection at server module or receival of data from client socket will activate the camera module to capture the image of the surrounding. Finally, a push notification alert with the captured image will be sent to the user via Telegram application. The system then continues until the next motion is detected.

fridaya@uef.fi

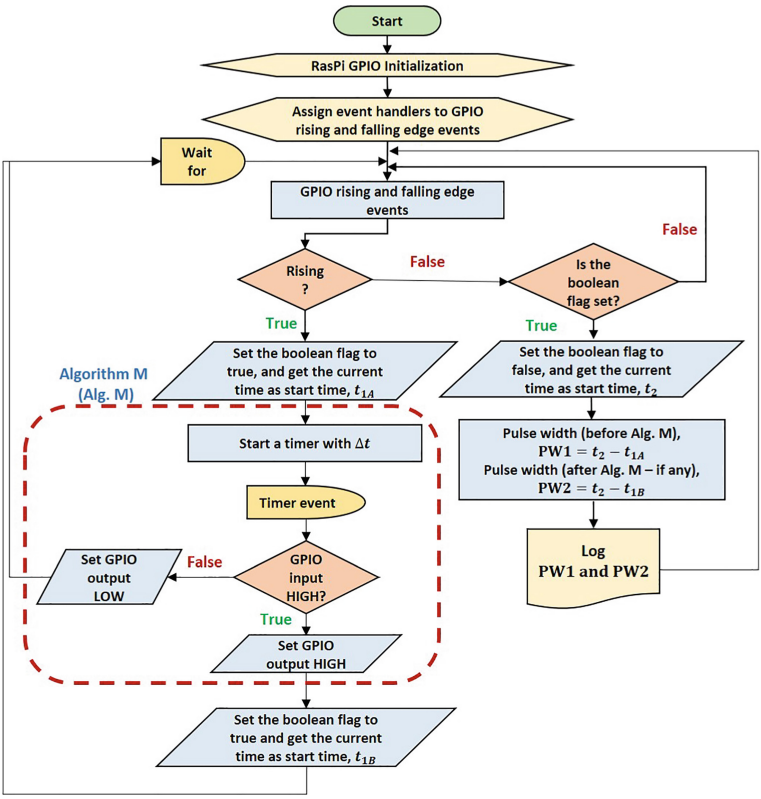


Fig. 4. Flowchart for the motion detection with Algorithm M (Alg. M).

The operation in the client module is comparably simpler. Referring to Fig. 6b, the initialization takes place at the client module side before it waits for any motion detection event. The actions during the initialization step is similar to those from the server module side except for the camera. Upon detection of motion, sensor data will be sent to the server module through the bluetooth and the system then will wait for the next motion detection event.

3 Experimental Results and Discussions

The experiments to evaluate the detection accuracy and false activation were carried out simultaneously for both sensors, with and without Alg. M. The timer Δt in Alg. M was set to 3 s as a human motion generally can generate an output with duration more than 3 s. The numerical results for accuracy test are recorded in Table 1 where we denote Sensor P for PIR sensor, and Sensor M to represent the microwave sensor. From the table, it is shown that the microwave sensor outperforms the PIR sensor for each experiment. The implementation of Alg. M

fridaya@uef.fi

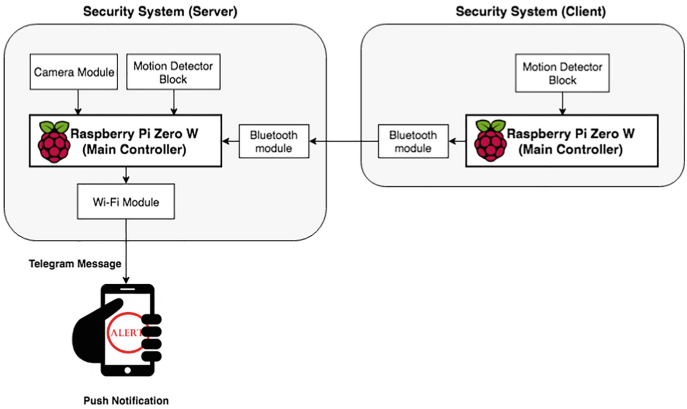


Fig. 5. Overview of the IoT-based home security system architecture.

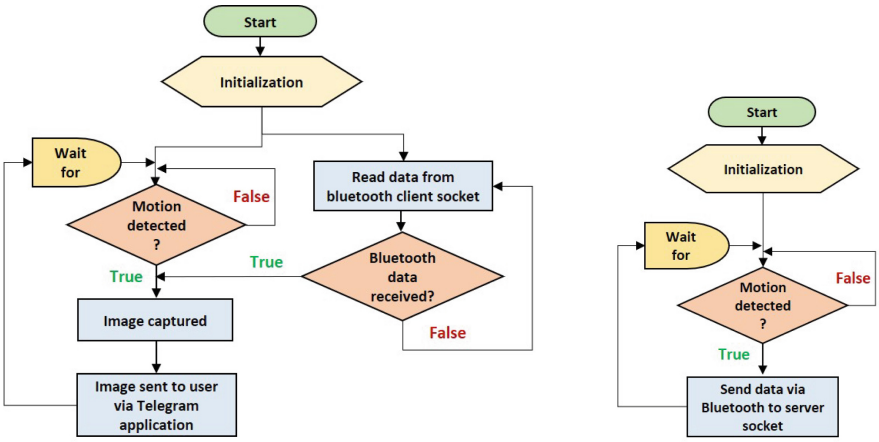
however reduces the accuracy of Sensor P by 6.25%, whereas for Sensor M, the reduction is by 5%.

The numerical results from false activation experiments are presented in Table 2 where it is observed that without Alg. M, Sensor P shows a significant number of false triggers for each scenario in 10 minutes, ranging from 7 to 11. Sensor M, on the other hand, only shows a small number of false triggers with a maximum of 4, which occurs when the wind speed is increased. It is also clear that the implementation of Alg. M is advantageous for this case as it can greatly reduce the false activation for both sensors where Sensor P's is reduced from 9 to 1.67, while Sensor M's false trigger can be totally eliminated.

Table 1. Numerical results for detection accuracy experiments.

Exp.	Sensor	No. successful detections		Accuracy (%)	
		No Alg. M	With Alg. M	No Alg. M	With Alg. M
1.1	P	15	11	75	55
	M	18	15	90	75
1.2	P	14	14	70	70
	M	18	18	90	90
1.3	P	14	13	70	65
	M	18	18	90	90
1.4	P	15	15	75	75
	M	18	17	90	85
Average	P	14.5	13.25	72.5	66.25
	M	18	17	90	85

fridaya@uef.fi



(a) Server module (b) Client module

Fig. 6. Flowcharts of the server and client modules

The home security system is implemented using two Python programs, ServerHomeSecurity.py and ClientHomeSecurity.py, for server module and client module respectively. In the server module, it waits for incoming Bluetooth connection request after initialization stage takes place. When the client requests to connect to the server, the server accepts the connection requested and thus communication between them is established. Figure 7 shows the message displayed on the PuTTY terminal window of server module when server accepted the connection requested by the client, in which 'B8:27:EB:A2:54:6C' refers to the Media Access Control (MAC) address of the client module while '1' following refers to the Bluetooth port. The value of Bluetooth port must be set identical at coding of both server and client programs. Figure 8 shows a simple message

Table 2. Numerical results for false activation experiments.

Exp.	Sensor	No. false triggers	
		No Alg. M	With Alg. M
2.1	P	9	2
	M	0	0
2.2	P	7	1
	M	1	0
2.3	P	11	2
	M	4	0
Average	P	9	1.67
	M	1.67	0

fridaya@uef.fi

displayed at PuTTY terminal window of the client module when the requested connection had been accepted at the server side.

Programs at both module sides will wait for the event of motion detection. When there is motion detected by the motion sensors, messages as highlighted in Figs. 9 and 10 are displayed. The text '[Microwave]', '[PIR]' and '[PIR AND Microwave]' are as the indicators for motion detected by microwave sensor, PIR sensor and both sensors respectively. The next few following lines are displayed when the motion has stopped and indicate the duration of the output signal from the sensors.

Fig. 7. Message displayed at server when it accepts connection requested by client.

Fig. 8. Message displayed at client when Bluetooth connection accepted by server.

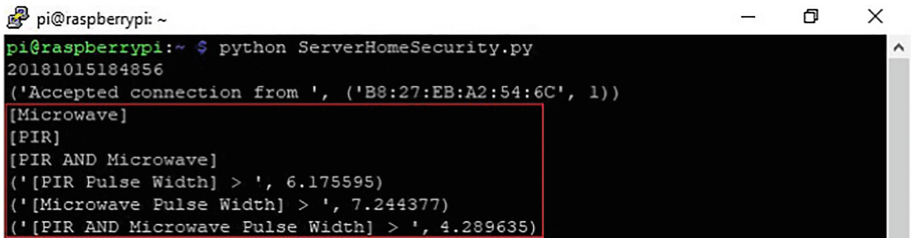


Fig. 9. Motion detected by server module.

Besides waits for the motion detection event, the server module also waits for the incoming Bluetooth data from client module when there is motion detected at client side. The highlighted text in Fig. 11 shows the sensor data from client is received by server.

Detection at either server or client side triggers the server module to take a photo using the camera module connected. The photo taken is saved at Raspberry Pi, with the current timestamp as the file name as shown in Fig. 12, which is more convenient for future reference.

fridaya@uef.fi

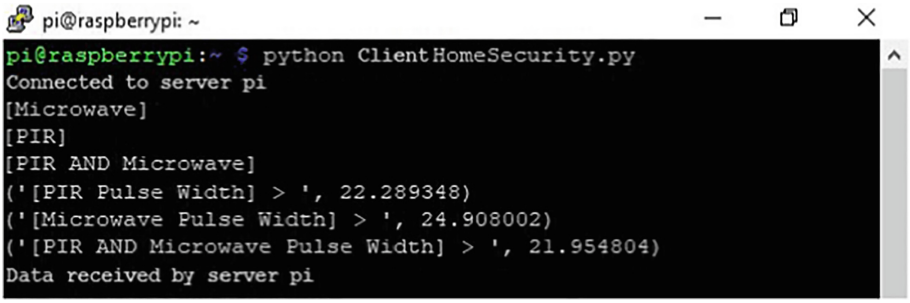


Fig. 10. Motion detected by client module.

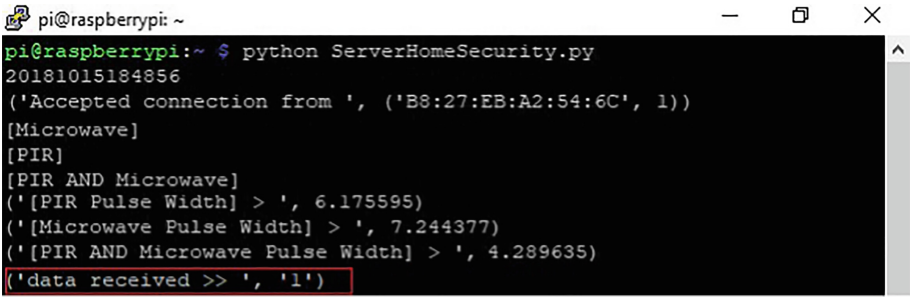


Fig. 11. Motion detection data from client module received by server module.

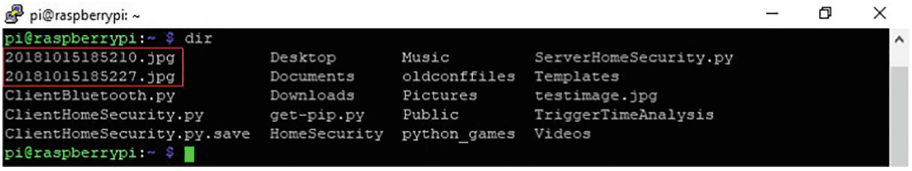


Fig. 12. Photo taken when motion detected.

The captured photo is then sent as a push notification via Telegram application, with the timestamp information to the users. Figure 13 shows a screenshot of the Telegram message received by the user.

fridaya@uef.fi

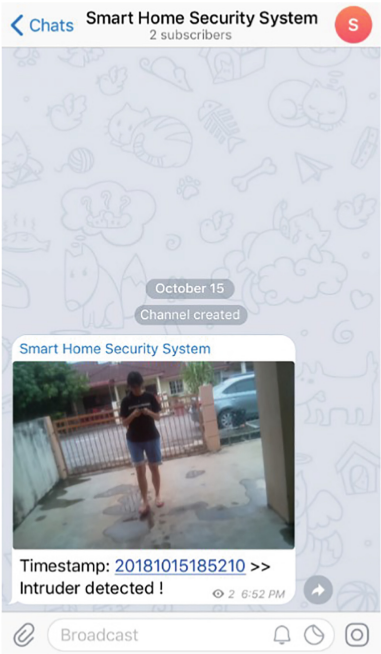


Fig. 13. Intruder detection push notification via Telegram.

4 Conclusion and Further Work

In this work, a motion detector with an IoT-based Home Security System is developed using a Raspberry Pi as the main platform for the control algorithm. Both PIR and microwave sensors were firstly evaluated in terms of their detection accuracy and false activation. A series of experiments in a car porch area were designed with various walking trajectories to evaluate the accuracy of each sensor, and three environmental scenario with different wind speed were considered for the evaluation on false triggers. We introduced a simple algorithm for each sensor to minimize the effects from the sensors' noise. Based on the experimental results, it is shown that the microwave sensor outperforms the PIR in terms of both detection accuracy and false activation. It is also observed that, the implementation of the algorithm offers a great reduction in terms of false activation, but with a slight trade-off in motion detection accuracy. Nevertheless, the overall performance of the microwave sensor is enhanced as the false trigger can be totally eliminated at the cost of only 5% accuracy reduction, compared to that of the PIR sensor which still shows a non-zero false triggers with 6.25% accuracy reduction after the implementation of the algorithm. The motion detection technique with microwave sensor was then integrated with the server-client modules using a simple Python program to test the functionality of the overall IoT-based home security system that has been developed in this work.

fridaya@uef.fi

For future work, sensor fusion technique may be considered and additional environmental scenario can be designed to increase the robustness of the motion detector.

Acknowledgments. The authors would like to thank Universiti Sains Malaysia for the financial support under the Research Universiti (RUI) Grant Scheme (1001/PELECT/8014055).

References

1. Ge, R., Shan, Z., Kou, H.: An intelligent surveillance system based on motion detection. In: 2011 4th IEEE International Conference on Broadband Network and Multimedia Technology, pp. 306–309, October 2011
2. Nguyen, H.-Q., Loan, T.T.K., Mao, B.D., Huh, E.-N.: Low cost real-time system monitoring using raspberry pi. In: 2015 Seventh International Conference on Ubiquitous and Future Networks, pp. 857–859, July 2015
3. Patil, Ambatkar, S., Kakde, S.: IoT based smart surveillance security system using Raspberry Pi. In: 2017 International Conference on Communication and Signal Processing (ICCSP), pp. 0344–0348, April 2017
4. Abaya, W.F., Basa, J., Sy, M., Abad, A.C., Dadios, E.P.: Low cost smart security camera with night vision capability using Raspberry Pi and OpenCV. In: 2014 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), pp. 1–6, November 2014
5. Kumar, K.N.K., Natraj, H., Jacob, T.P.: Motion activated security camera using Raspberry Pi. In: 2017 International Conference on Communication and Signal Processing (ICCSP), pp. 1598–1601, April 2017
6. Luo, X., Shen, B., Guo, X., Luo, G., Wang, G.: Human tracking using ceiling pyroelectric infrared sensors. In: 2009 IEEE International Conference on Control and Automation, pp. 1716–1721, December 2009
7. Yun, J., Song, M.: Detecting direction of movement using pyroelectric infrared sensors. IEEE Sens. J. **14**(5), 1482–1489 (2014)
8. Kodali, R.K., Jain, V., Bose, S., Boppana, L.: IoT based smart security and home automation system. In: 2016 International Conference on Computing, Communication and Automation (ICCCA), pp. 1286–1289, April 2016
9. Chandra, M.L.R., Kumar, B.V., Suresh Babu, B.: Iot enabled home with smart security. In: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 1193–1197, August 2017
10. Nuti, P., Yavari, E., Boric-Lubecke, O.: Doppler radar occupancy sensor for small-range motion detection. In: 2017 IEEE Asia Pacific Microwave Conference (APMC), pp. 192–195, November 2017
11. Tetervenoks, O., Suskis, P., Stegura, J.: Integration of microwave sensor into low cost indoor led lamp – element of smart lighting system. In: 2017 5th IEEE Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE), pp. 1–4, November 2017

fridaya@uef.fi

12. Kaushik, A.R., Lovell, N.H., Celler, B.G.: Evaluation of PIR detector characteristics for monitoring occupancy patterns of elderly people living alone at home. In: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3802–3805, August 2007
13. Akbas, S., Efe, M.A., Ozdemir, S.: Performance evaluation of PIR sensor deployment in critical area surveillance networks. In: 2014 IEEE International Conference on Distributed Computing in Sensor Systems, pp. 327–332, May 2014

fridaya@uef.fi

Check for updates

A Simple Parameter Estimation Method for Permanent Synchronous Magnet Motors

Dong-Myung Lee(✉)

School of Electronic and Electrical Engineering, Hongik University, Seoul, Korea
dmlee@hongik.ac.kr

Abstract. In order to improve control performance, the vector control technique has been widely used in household appliances, such as a washing machine. To determine the operation mode of the washing machine it is necessary to detect the amount of the laundry, and the information of the motor constants is required for these operations. For instance, the motor torque is estimated from torque calculation using motor parameters such as stator inductance and flux-linkage. For the exact detection of motor torque, online parameter estimation scheme is necessary. Hence, this paper proposes an online stator induction calculation method and flux-linkage estimation scheme based on the MRAS algorithm. Simulations have been carried out to show the validity of the proposed algorithm.

Keywords: Parameter estimation · SPM motor · MRAS (Model reference adaptive System)

1 Detection of Motor Parameters

This paper presents a new method for estimating stator inductance and flux-linkage of a surface-mounted permanent magnet (SPM) motor used for washing machines. It is necessary to know the motor constant for detecting the amount of the laundry for the laundry mode selection. Therefore, to improve the performance of washing machines through the precise estimation of the motor constant, this paper proposes the new algorithm of detecting motor parameters. The detection algorithm derives from stator voltage equations. The stator voltage equation of the SPM motor is represented by the following equations:

$$v_{ds} = R_s i_{ds} + L_s \frac{di_{ds}}{dt} - \omega L_s i_{qs},\ v_{qs} = R_s i_{qs} + L_s \frac{di_{qs}}{dt} + \omega(L_s i_{ds} + \lambda_m) \tag{1}$$

The state equation form for the above equations can be written as follows:

$$\frac{di_{ds}}{dt} = -\frac{R_s}{L_s} i_{ds} + \omega i_{qs} + \frac{v_{ds}}{L_s},\ \frac{di_{qs}}{dt} = -\frac{R_s}{L_s} i_{qs} - \omega i_{ds} + \frac{v_{qs}}{L_s} - \omega \frac{\lambda_m}{L_s} \tag{2}$$

As can be seen from Eq. (2), at least two or more of motor constants are included in the formula in each equation. Equation (2) has R_s, L_s, and λ_m, which vary depending on

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1230–1234, 2019.
https://doi.org/10.1007/978-3-030-22868-2_86

fridaya@uef.fi

the operating condition [1]. It is necessary to reduce the number of unknowns to one. Hence, in this paper, L_s is estimated at first as below.

1.1 Calculation of Stator Inductance

The stator inductance L_s is obtained by converting the stator voltage equations of (1) to those in an arbitrary rotation frame. In the calculation, the operating condition of $i_{ds} = 0A$ is applied as in general SPM operation [2]. Equation (1) is voltage equations in d-q rotational coordinate system based on the flux linkage (λ_m) in the stator. Those rewritten with the arbitrary $\delta - \gamma$ axis rotating at ω speed can be described as below:

$$v_{\delta s} = R_s i_{\delta s} + L_s \frac{di_{\delta s}}{dt} - \omega L_s i_{\gamma s} + e_\delta,\ v_{\gamma s} = R_s i_{\gamma s} + L_s \frac{di_{\gamma s}}{dt} - \omega L_s i_{\delta s} + e_\gamma \tag{3}$$

where, $e_\delta = e \cdot sin\varphi$, $e_\gamma = e \cdot cos\varphi$, and φ is the angle between γ-axis and δ-axis (see Fig. 1).

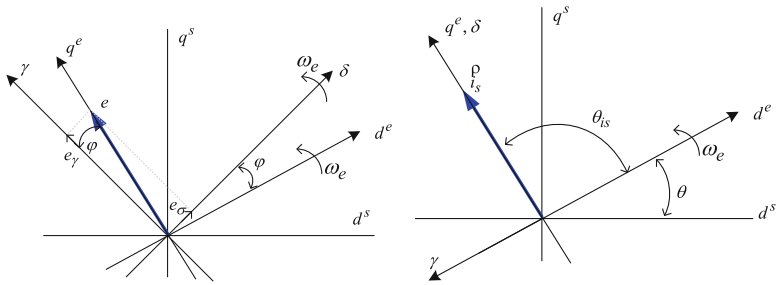


Fig. 1. (a) The relationship between the electromotive force e, d-q axis, and the arbitrary rotating axis $\delta - \gamma$, (b) The vector diagram of i_s when it coincides with δ axis.

where, $\overrightarrow{v_{\delta\gamma}} = \overrightarrow{v_{dqs}} e^{-j\varphi}.v_{\delta\gamma}$ represents the voltage vector in $\delta - \gamma$ axis, and $\overrightarrow{v_{dqs}}$ is stator vector in the d-q frame. The stator current can be expressed as (4) due to $i_{ds} = 0A$.

$$\left|\vec{i}_s\right| = i_s = \sqrt{i_{ds}^2 + i_{qs}^2} = \left|i_{qs}\right| \tag{4}$$

The differential term in Eq. (3) disappears because it is assumed the motor rotates at the constant speed during estimation. In order to reduce the unknowns in Eq. (3), the below relationship shown in Fig. 1 is introduced by aligning the δ axis with the q-axis. Where, θ_{is} is the angle between the d-axis and i_s. If i_{ds}= 0A, θ_{is} is 90°.

$$i_{\delta s} = i_{qs} = \vec{i}_s,\ i_{\gamma s} = 0,\ e_\delta = e \cdot sin\varphi = e,\ e_\gamma = e \cdot cos\varphi = 0,\ \theta_{is} = 90^\circ \tag{5}$$

fridaya@uef.fi

Therefore, Eq. (3) is simplified as follows:

$$v_{\delta s} = R_s i_s + e \cdot sin\theta_{si} = R_s i_s + e,\ v_{\gamma s} = \omega L_s i_s + e \cdot cos\theta_{si} = \omega L_s i_s \tag{6}$$

Thus, L_s is obtained from Eq. (6) as $L_s = v_{\gamma s}/\omega i_s$.

On the other hand, if the fact that $\varphi = \theta_{is} = 90^\circ$ is used from the relation of $\overrightarrow{v_{\delta\gamma}} = \overrightarrow{v_{dqs}} e^{-j\varphi}$, $v_{\gamma s} = -v_{ds}$ is known. Since, $i_s = i_{qs}$, L_S can be obtained in real-time by using the value shown on the d-q axis without needing coordinate conversion for the current, voltage, etc. expressed in $\delta - \gamma$ axis as $L_s = -v_{ds}/\omega i_{qs}$. Where, R_s can be obtained by various methods, and R_s detection is not subject of this paper, not mentioned in this paper.

1.2 Estimation of Flux-Linkage by Magnet

In Eq. (2), unknown is λ_m. The equation for the estimate in Eq. (2) can be written as (7). The flux-linkage(λ_m) by the magnet is estimated on the basis of the error between measured i_{qs} and the estimated one from Eq. (7). The error between Eq. (2) and (7) is described as (8).

$$\frac{d\hat{i}_{qs}}{dt} = -\frac{R_s}{L_s}\hat{i}_{qs} - \omega i_{ds} + \frac{v_{qs}}{L_s} - \omega\frac{\hat{\lambda}_m}{L_s} \tag{7}$$

$$\frac{de_q}{dt} = -\frac{R_s}{L_s}e_q - \frac{\omega}{L_s}\left(\lambda_m - \hat{\lambda}_m\right) = -\frac{R_s}{L_s}e_q - \frac{\omega}{L_s}\Delta\lambda_m = Ae_q - W \tag{8}$$

where, $e_q = i_{qs} - \hat{i}_{qs}$, and $\Delta\lambda_m = \lambda_m - \hat{\lambda}_m$. Therefore, in order to estimate λ_m through Eq. (8), it should satisfy the Popov's integral inequality [3]. Since $E^T W = \Delta\lambda_m \cdot e_q \cdot (\omega/L_s)$, to meet the condition of (8), the adaptive law's for λ_m estimation is selected as Eq. (9) with the form of PI controller. Figure 2 shows the block diagram for obtaining λ_m by MRAS scheme.

$$\hat{\lambda}_m = \hat{\lambda}_m(0) - \left(k_p + \frac{k_i}{s}\right)\left(e_q \cdot \frac{\omega}{L_s}\right) \tag{9}$$

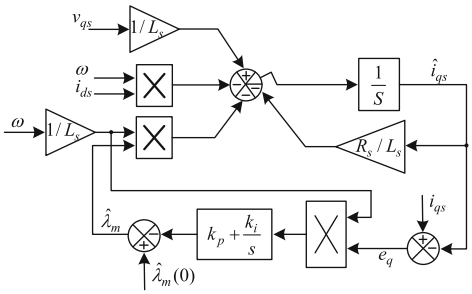


Fig. 2. Block diagram of λ_m estimation by MRAS scheme.

fridaya@uef.fi

2 Results

Figure 3 shows the estimated L_s waveform. The motor rotates at 150[rpm] and after t = 0.4[seconds] revolves at 250[rpm]. As shown in Fig. 3(b), the load torque has been varied during the simulation period. As can be seen in Fig. 3, the L_s is well estimated with 32[mH] even though the load condition and speed reference have been changed.

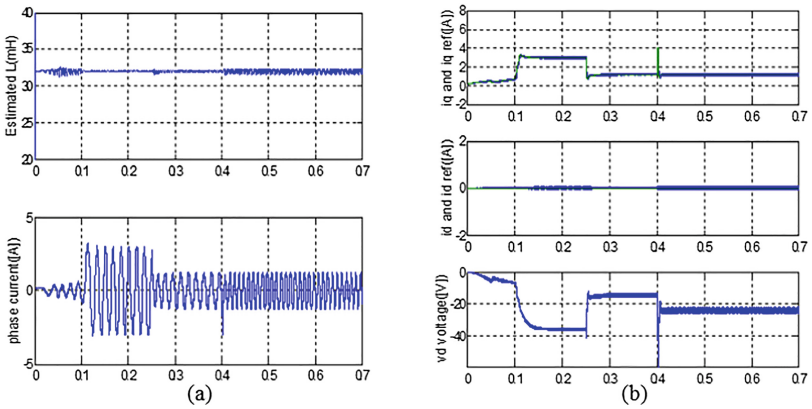


Fig. 3. (a) Estimated stator inductance and phase current; (b) q-axis current & reference, d-axis current & reference, and d-axis voltage(v_{ds}).

Figure 4 illustrates the simulation waveform of λ_m estimation. As seen in Fig. 4(a), estimated i_{qs} converges to the actual q-axis value by using the adaptive laws of Eq. (9) and implementation as Fig. 2, and λ_m converges to the actual value, 0.15[V-s].

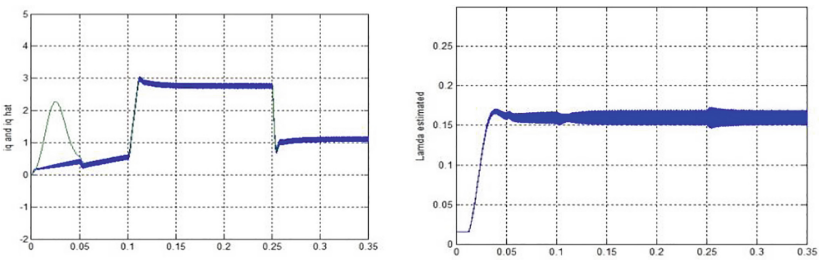


Fig. 4. (a) i_{qs} and estimated i_{qs}, (b) estimated λ_m.

3 Conclusion

In this paper, a simple motor constant operation method was proposed. L_s was obtained by calculation from the voltage command value and current, and the flux-linkage was obtained using the MRAS technique. Simulation showed the validity of the proposed method.

fridaya@uef.fi

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF- 2016R1A2B4011954).

References

1. Zhong, C., Lin, Y.: Model reference adaptive control (MRAC)-based parameter identification applied to surface-mounted permanent magnet synchronous motor. Int. J. Electron. **104**(11), 1854–1873 (2017)
2. Deng, W., Xia, C., Yan, Y., Geng, Q., Shi, T.: Online multiparameter identification of surface-mounted PMSM considering inverter disturbance voltage. IEEE Trans. Energy Convers. **32**, 202–212 (2017)
3. Kumar, M., Das, S.: Model reference adaptive system based sensorless speed estimation of brushless doubly-fed reluctance generator for wind power application. IET Power Electron. **11**(14), 2355–2366 (2018)

fridaya@uef.fi

A Bug-Inspired Algorithm for Obstacle Avoidance of a Nonholonomic Wheeled Mobile Robot with Constraints

Sing Yee Ng and Nur Syazreen Ahmad(✉)

School of Electrical and Electronic Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Pulau Pinang, Malaysia
syazreen@usm.my

Abstract. Many bug-inspired algorithms are well-developed and can be easily simulated if the robot is assumed to be a small particle with perfect movements and self-positioning capability. The sensor is also presumed to be able to obtain complete obstacle information in order for the algorithm to work perfectly in the simulation. In practice, however, applications of the algorithms on mobile robots are not straightforward as the robots generally have limited mobility due to disturbances such as frictions as well as constraints on the actuators. The obstacle detection capability also depends on the type of sensors used which usually has a trade-off between accuracy and price. In this work, we propose a simple approach to design and develop an obstacle avoidance strategy based on one of bug-inspired algorithms for a nonholonomic wheeled mobile robot with constraints on its velocities and sensing range. The proposed method is validated via several experiments with different goal coordinates and number of obstacles, and we show that the robot is able to avoid collisions and reach the destinations with very small steady-state position errors.

Keywords: Bug algorithm · Obstacle avoidance · Constraints

1 Introduction and Related Work

A Bug-inspired algorithm is a kind of autonomous navigation algorithm suitable for obstacle avoidance strategy of autonomous mobile robots in unknown environment. The method was first proposed by Lumelsky and Stepanov [1] which initially assumes only local knowledge of the environment and a global destination. The algorithm relies on obstacle detection mechanisms and it can in practice be implemented together with either a contact (tactile) sensor or a range sensor [2–5]. These two categories are using different types of sensors to collect local data such as the location of the obstacle.

The earlier versions of Bug algorithm with contact sensors are named Bug 0, Bug 1 and Bug 2. Figure 1 shows the Bug 0 algorithm where the grey blocks represent the obstacles, and the robot is assumed only a "point". The red and green points indicate the starting and ending (goal) points respectively. Initially,

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1235–1246, 2019.
https://doi.org/10.1007/978-3-030-22868-2_87

fridaya@uef.fi

the robot will compute a collision-free path and move towards the goal. If an obstacle is detected, the robot will follow the obstacle until it can head towards the goal again. Although the algorithm looks simple, the problem with Bug 0 is that it has no memory. This problem however is solved in Bug 1 and Bug 2 algorithms.

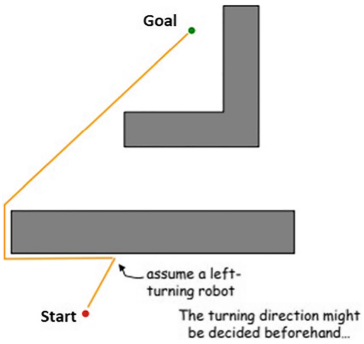


Fig. 1. Bug 0 algorithm.

Bug 1 algorithm is illustrated in Fig. 2. The first step is similar to that of Bug 0 algorithm. However, if an obstacle is encountered, the robot will circumnavigate it and remember the distance between the current point and the goal point. Then, the robot will return to the point with the shortest distance from the goal, and continue moving towards it.

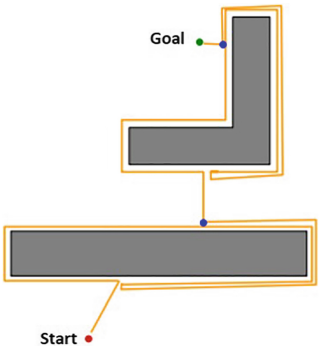


Fig. 2. Bug 1 algorithm.

Bug 2 algorithm is slightly different from Bug 0 and Bug 1 as it remembers the shortest collision-free path in the first step. The path is typically termed as the "m-line". Figure 3 shows the Bug 2 algorithm with the the corresponding

fridaya@uef.fi

m-line. If an obstacle is detected on its way to the goal, the robot will move along its boundary until the m-line is encountered again. The robot will then leave the obstacle and continue towards the goal. This is more preferable than Bug 1 if the shortest path is a priority in the design requirement.

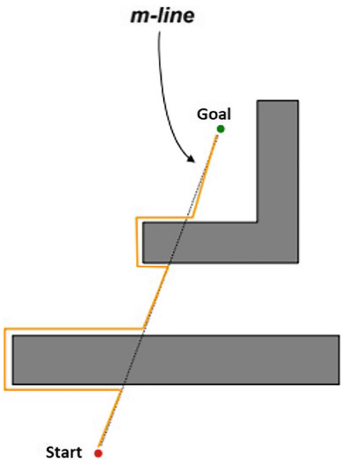


Fig. 3. Bug 2 algorithm.

The Bug algorithm with range sensor, also called the Tangent Bug, heavily relies on the sensor to compute finite continuous segment endpoints on boundaries of the obstacle. It continually updates the minimum distance to goal after clearing each obstacle. With a high accuracy range sensor and large detection area, this algorithm is best suited for unstructured environments with scattered obstacles.

The aforementioned algorithms are well-developed and can be easily simulated as the robot is only assumed to be a small particle with perfect movements and self-positioning capability [6,7]. The sensor is also presumed to be able to obtain complete obstacle information. In practice, however, applications of the algorithms on mobile robots are not straightforward as the robots generally have limited mobility due to external disturbances such as frictions as well as constraints on the actuators [8–10]. The obstacle detection capability also depends on the type of sensors used which has a trade-off between the accuracy and price. In this work, we propose a simple approach to design and develop an obstacle avoidance strategy based on Bug 2 algorithm for a nonholonomic wheeled mobile robot with constraints on its velocities and sensing range. The proposed method is validated via several experiments with different goal coordinates and number of obstacles, and we show that the robot is able to avoid collisions and reach the destinations with very small steady-state position errors.

fridaya@uef.fi

2 Methodology

In this work, we use a nonholonomic wheeled mobile robot with a forward-looking sensor as depicted in Fig. 4 where the detection range is limited as shown by the light yellow area. The variable ρ represents the distance between the center of the sensor and the obstacle, and θ_s denotes the angle from the negative y_R-axis to the center line of the sensing area.

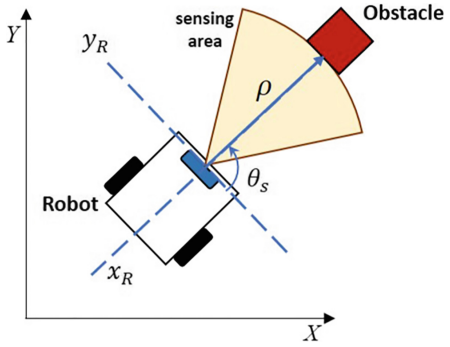


Fig. 4. Overview of the nonholonomic wheeled mobile robot with its sensing area.

The mobility of the robot is not only limited by its nonholonomic constraints, but is also bounded by a certain velocity range. In this regard, we write $\bar{\omega} = (\omega_L, \omega_R)$ as the left-right wheel velocity pair of the mobile robot, and f_D is the command velocity. The velocity $\bar{\omega}$ is constrained as follows:

$$\bar{\omega} = \begin{cases} (0,0) & \text{if } f_D = 1 \\ (\omega_L^*, \omega_R^*) & \text{if } f_D = 2 \\ (\omega_L^*, 0) & \text{if } f_D = 3 \\ (0, \omega_R^*) & \text{if } f_D = 4 \end{cases} \tag{1}$$

where ω_L^* and ω_R^* are pre-specified angular velocities of the left and right wheels, respectively.

As the movement of the sensor's position is also limited, another constraint is immediately imposed on the sensing area. To this end, we write f_s as the command for the sonar sensor's position. The angle θ_s is constrained as follows:

$$\theta_s = \begin{cases} \pi/2 & \text{if } f_s = 1 \\ \pi/4 & \text{if } f_s = 2 \\ 3\pi/4 & \text{if } f_s = 3 \end{cases} \tag{2}$$

As discussed in the previous section, Bug 2 algorithm is less conservative than Bug 1 and Bug 0 because the robot can leave earlier due to the m-line. Normally the m-line of Bug 2 algorithm is the shortest path between the starting point and goal point. In this work however, the path is slightly modified to suit the velocity

fridaya@uef.fi

constraint of the robot which can also reduce the systematic error. Hence, the task resembles a combination of path following and Bug 2 algorithm. The m-line in this case is predefined where it is divided into x-axis and y-axis based on the x and y coordinates of the starting point goal point. An example is shown in Fig. 5 where "**s**" is the starting point and "**G**" is the desired goal point with a coordinate (x_g, y_g). Without the obstacle, the robot will move along the x-axis first until it reaches the x coordinate of the desired goal point, then will only go to the y coordinate of the goal point. With the presence of the obstacle, the robot will operate similar to Bug 2 algorithm, and when it detects an obstacle (represented by the blue square box), it will turn to either left or right, follow the obstacle's boundaries until it reaches the m-line again.

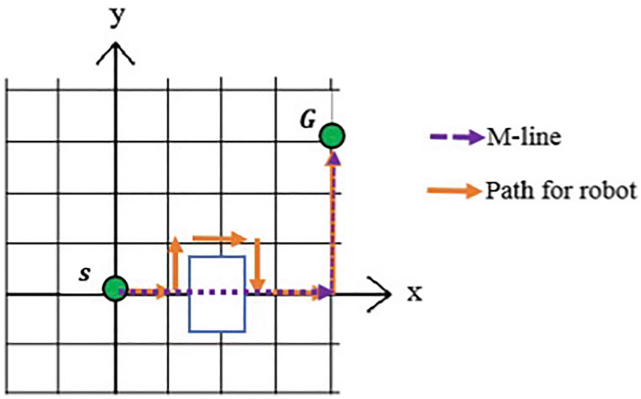


Fig. 5. Combination of path following and Bug 2 algorithm.

Based on the constraints described in (1) and (2), we propose a suitable algorithm to be implemented on the robot. This is described as shown in Figs. 6 and 7. We implement a left wall follower whenever the robot detects obstacles, and set ϵ as the minimum distance of ρ before it turns to the left.

From Fig. 6, the algorithm starts with predefined x_g and y_g of the goal point. Then, the x coordinate of the current position of the robot is compared with x_g. If there is no obstacle detected, the robot will move to x_g first, and then straight to y_g. If an obstacle is detected, the robot stops and the sensor turns to the right. If there is no obstacle, the right wheel stops and the left wheel moves, making it turn to the right. At the same time, the sensor is positioned to face the left side of the obstacle so that the robot is able to follow the boundaries. When the m-line is encountered again, the robot will follow it until x_g is met, and then move straight until it reaches y_g (the algorithm continues to Fig. 7). Otherwise, the same procedure follows in order for the robot to turn left until it reaches the goal.

fridaya@uef.fi

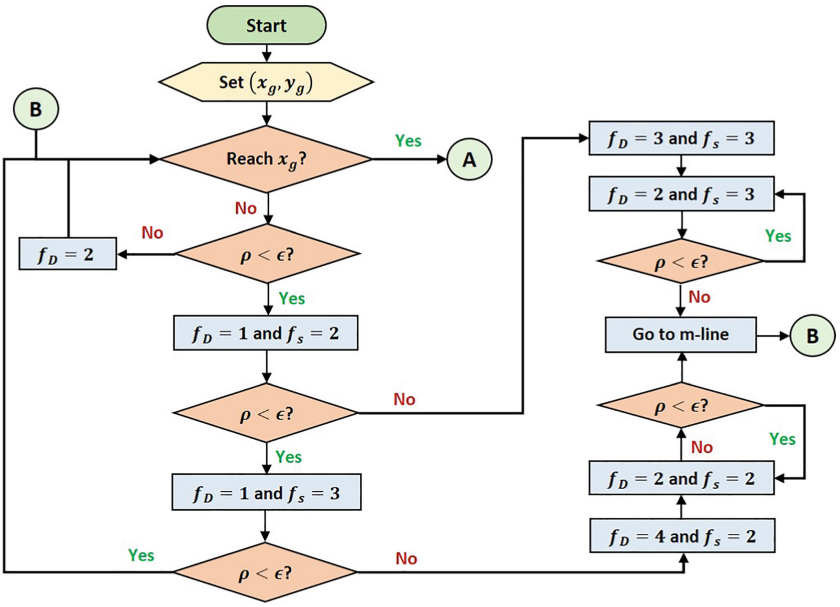


Fig. 6. Bug-inspired algorithm.

3 Experimental Results

The nonholonomic wheeled mobile robot built for the experiments is shown in Fig. 8 which is equipped with a forward-looking sonar sensor for obstacle detection. We use MATLAB to plot the robot's path and a bluetooth for wireless data acquisitions.

Two desired goal coordinates were considered for the experiments, i.e. $(x_g, y_g) = (120, 50)$ and $(x_g, y_g) = (60, -60)$. For each goal coordinate, three scenario were tested, i.e. when there was no obstacle, when there was one obstacle and when there were two obstacles near/on the m-line.

For Experiment 1 (i.e. $(x_g, y_g) = (120, 50)$), the path of the mobile robot for the three scenario are shown in Figs. 9, 10 and 11, whereas for Experiment 2 (i.e. $(x_g, y_g) = (60, -60)$), the path of the mobile robot for the three scenario are shown in Figs. 12, 13 and 14. We name "Targeted path" as the expected path of the robot by simulation when the constraints are taken into account, and "Real path" as the path of the robot from the experiments.

Let (x_s, y_s) be the final position of the mobile robot from the sensor data, and (x_m, y_m) be the final position of the mobile robot from the real measurement. For each experiment, three tests were carried out and the results including the average final positions are tabulated in Tables 1 and 2.

Define e_s and e_m as steady-state position errors from the sensor data and real measurements respectively, the results are tabulated in Tables 3 and 4.

fridaya@uef.fi

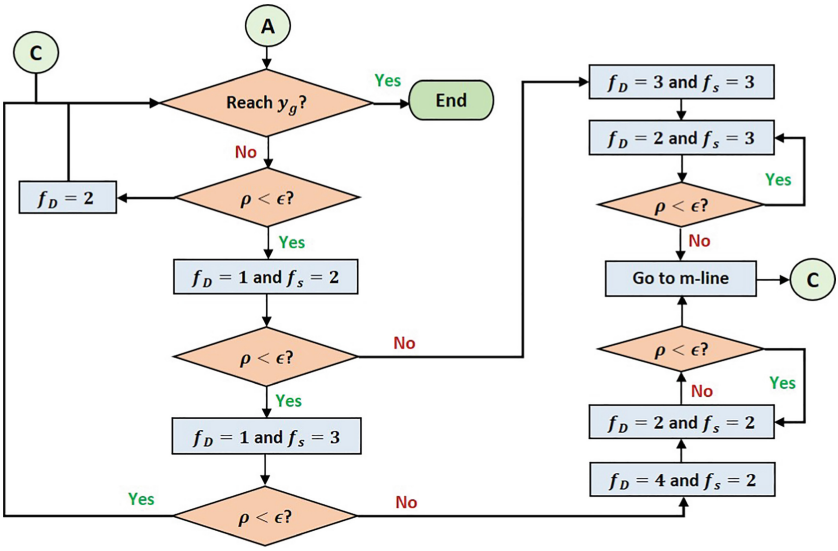


Fig. 7. Bug-inspired algorithm *(continued)*.

Table 1. Final positions of the mobile robot when $(x_g, y_g) = (120, 50)$ from the sensor data and real measurement.

Test	No obstacle		One obstacle		Two obstacles	
	(x_s, y_s)	(x_m, y_m)	(x_s, y_s)	(x_m, y_m)	(x_s, y_s)	(x_m, y_m)
1	$(120, 50)$	$(120, 50)$	$(121, 50)$	$(121, 49)$	$(121, 50)$	$(121, 50)$
2	$(120, 50)$	$(118, 50)$	$(121, 50)$	$(120, 49)$	$(121, 50)$	$(118, 52)$
3	$(120, 50)$	$(118, 51)$	$(121, 50)$	$(124, 47)$	$(121, 50)$	$(121, 47)$
Average	$(120, 50)$	$(118.7, 50.3)$	$(121, 50)$	$(121.7, 48.3)$	$(121, 50)$	$(120, 49.7)$

Table 2. Final positions of the mobile robot when $(x_g, y_g) = (60, -60)$ from the sensor data and real measurement.

Test	No obstacle		One obstacle		Two obstacles	
	(x_s, y_s)	(x_m, y_m)	(x_s, y_s)	(x_m, y_m)	(x_s, y_s)	(x_m, y_m)
1	$(59, -61)$	$(59, -60)$	$(62, -61)$	$(63, -61)$	$(62, -61)$	$(62, -60)$
2	$(59, -61)$	$(59, -60)$	$(62, -61)$	$(59, -61)$	$(62, -61)$	$(63, -60)$
3	$(59, -61)$	$(59, -60)$	$(59, -61)$	$(58, -61)$	$(62, -61)$	$(58, -60)$
Average	$(59, -61)$	$(59, -60)$	$(61, -61)$	$(61, -60)$	$(62, -61)$	$(60, -60)$

fridaya@uef.fi

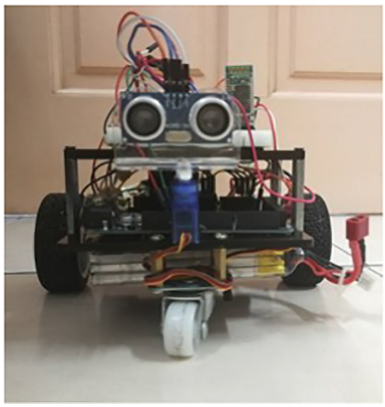


Fig. 8. The wheeled mobile robot built for the experiments.

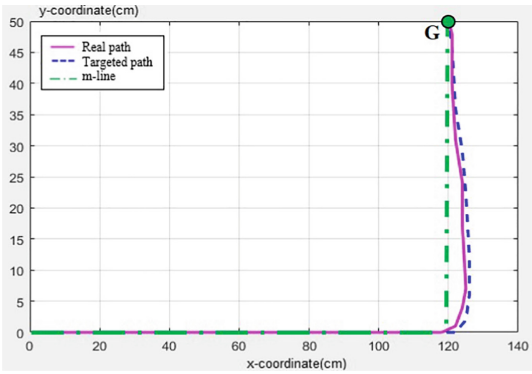


Fig. 9. Experiment 1: No obstacle

Table 3. Steady-state position errors when $(x_g, y_g) = (120, 50)$ from the sensor data and real measurement. All values are in "cm".

Test	No obstacle		One obstacle		Two obstacles	
	e_s	e_m	e_s	e_m	e_s	e_m
1	0	0	1.00	1.41	1.00	1.00
2	0	2.00	1.00	1.00	1.00	2.83
3	0	2.24	1.00	5.00	1.00	3.16
Average	0	1.41	1.00	2.47	1.00	2.33

fridaya@uef.fi

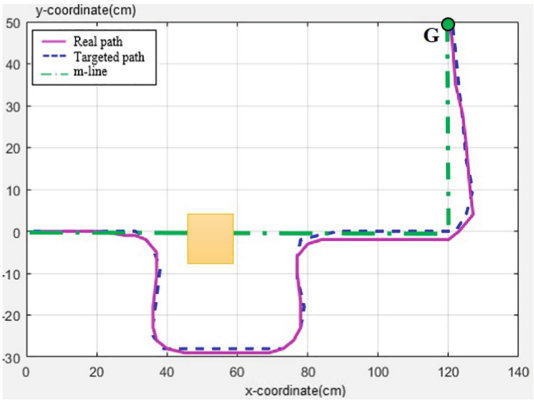


Fig. 10. Experiment 1: One obstacle

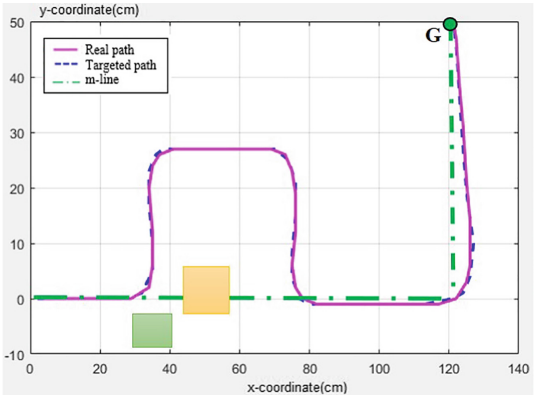


Fig. 11. Experiment 1: Two obstacles

Table 4. Steady-state position errors when $(x_g, y_g) = (60, -60)$ from the sensor data and real measurement. All values are in "cm".

Test	No obstacle		One obstacle		Two obstacles	
	e_s	e_m	e_s	e_m	e_s	e_m
1	1.41	1.00	2.24	3.16	2.24	2.00
2	1.41	1.00	2.24	1.41	2.24	3.00
3	1.41	1.00	1.41	2.24	2.24	2.00
Average	1.41	1.00	1.96	3.18	2.24	2.33

fridaya@uef.fi

From the results of both experiments, it can be observed that the algorithm proposed is able to drive the robot to the goal points without colliding with the obstacles. The resulting steady-state position errors via sensor data and real measurements are considered quite small as they are no greater than 2.5 cm for Experiment 1, and 3.2 cm for Experiment 2.

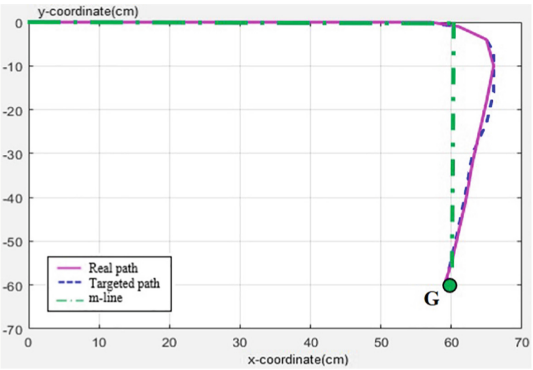


Fig. 12. Experiment 2: No obstacle

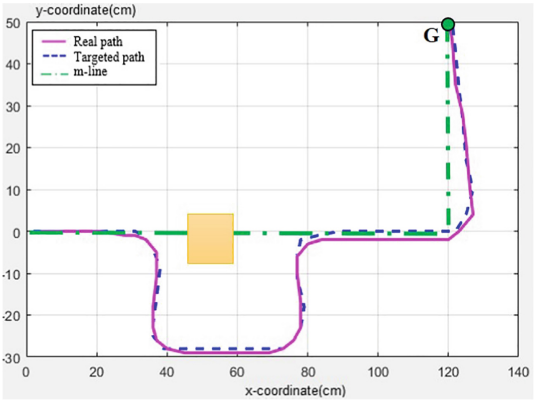


Fig. 13. Experiment 2: One obstacle

fridaya@uef.fi

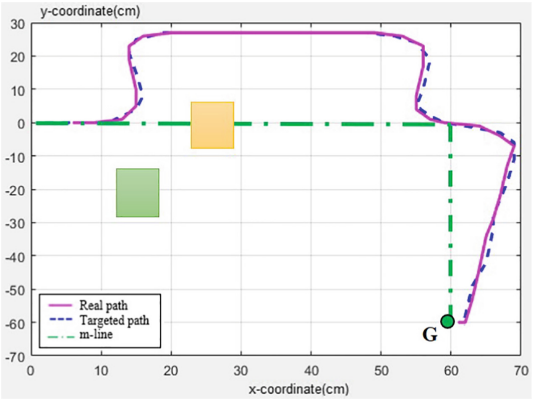


Fig. 14. Experiment 2: Two obstacles

4 Conclusion and Future Works

In this work, a bug-inspired algorithm has been proposed which is based on combination of path following method and Bug 2 algorithm. The technique is proposed to accommodate for the velocity and sensing range constraints of the mobile robot. Experiments with different goal points and number of obstacles were conducted to test the performance of the proposed algorithm, and it has been shown that the robot is able to avoid collisions and reach the destinations with very small steady-state position errors.

For future works, different kind of range sensors may be considered to enhance the performance of the obstacle detection, and the algorithm may need some modifications to suit other environments such as those with cluttered obstacles.

Acknowledgments. The authors would like to thank Universiti Sains Malaysia for the financial support under the Research Universiti (RUI) Grant Scheme (1001/PELECT/8014055).

References

1. Lumelsky, V.J., Stepanov, A.A.: Path-planning strategies for a point mobile automation moving amidst unknown obstacles of arbitrary shape. Algorithmica **2**(1), 403–430 (1987)
2. Lumelsky, J., Skewis, T.: Incorporating range sensing in the robot navigation function. IEEE Trans. Syst. Man Cybern. **20**, 1058–1069 (1990)
3. Alpaslan, Y., Osman, P.: A new range-sensor based globally convergent navigation algorithm for mobile robots. In: Proceedings of the IEEE International Conference on Robotics and Automation (1996)
4. Lozano-Pérez, T., Wesley, M.A.: An algorithm for planning collision-free paths among polyhedral obstacles. Commun. ACM **22**(10), 560–570 (1979)

fridaya@uef.fi

5. Kandathil, J.J., Mathew, R., Hiremath, S.S.: Modified bug-1 algorithm based strategy for obstacle avoidance in multi robot system. In: MATEC Web Conference, vol. 144, p. 01012 (2018)
6. Alpaslan, Y., Osman, P.: Performance comparison of bug algorithms for mobile robots. In: 5th International Advanced Technologies Symposium (IATS 2009), 13–15 May 2009, pp. 4944–4949 (2009)
7. Zohaib, M., Javaid, N., Salaam, A., Iqbal, J.: An improved algorithm for collision avoidance in environments having U and H shaped obstacles. Stud. Inform. Control **23**(1), 97–106 (2014). COMSATS Institute of Information Technology
8. Zhu, Y., Zhang, T., Song, J., Li, X.: A new bug-type navigation algorithm considering practical implementation issues for mobile robots. In: IEEE International Conference on Robotics and Biomimetics, pp. 531–536 (2010)
9. Xu, Q., Yu, T., Bai, J.: The mobile robot path planning with motion constraints based on bug algorithm. In: 2017 Chinese Automation Congress (CAC), pp. 2348–2352, October 2017
10. Sezer, V., Gokasan, M.: A novel obstacle avoidance algorithm: "Follow the Gap Method". Robot. Auton. Syst. **60**(9), 1123–1134 (2012)

fridaya@uef.fi

Obstacle Avoidance Strategy for Wheeled Mobile Robots with a Simplified Artificial Potential Field

Sing Yee Ng and Nur Syazreen Ahmad(✉)

School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Pulau Pinang, Malaysia
syazreen@usm.my

Abstract. Obstacle avoidance is one of the most basic path planning strategies particularly for mobile robot navigation in unknown environments. In this work, the focus is on design and development of an Artificial Potential Field (APF) based obstacle avoidance strategy of a nonholonomic wheeled mobile robot. A simplified APF-based algorithm is proposed to accommodate for the constraints of the robot due to its limited mobility. The method was validated via a series of experiments with different goal coordinates and number of obstacles. From the numerical and experimental results, we showed the effectiveness of the proposed strategy as it was able to drive the mobile robot to the goal coordinates with very small steady-state position errors without collisions with enroute obstacles.

Keywords: Artificial Potential Field · Obstacle avoidance · Mobile robot

1 Introduction

Autonomous mobile robots have been considered for many service applications such as intelligent household cleaners, search and rescue and unmanned delivery systems [1,2]. For some specific tasks particularly in unknown or cluttered environments, obstacle avoidance strategy is crucial to ensure the robots can safely reach their destinations [3–5]. In a series of papers devoted to this work, many different techniques have been introduced including Dynamic Window Approach (DWA) [6–8], Virtual Force Field (VFF), Vector Field Histogram (VFH) [9], Bug algorithms [10], and Artificial Potential Field (APF) [11–14]. The DWA is a search method in a velocity space which consists of three main stages. The first stage eliminates unreachable velocities coming from the acceleration limits of the robot. At the second stage, all velocity pairs which are not able to stop the robots before colliding with oncoming obstacles are eliminated, and at the final stage, the velocity pair which has the maximum value from the objective function optimization will be executed [6,7]. The success rate via this method however will decrease as the

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1247–1258, 2019.
https://doi.org/10.1007/978-3-030-22868-2_88

fridaya@uef.fi

obstacle density increases and the space becomes more crowded. In such a situation, a long trajectory will be generated to distant the robot from the crowd, which consequently leads to a longer tracking time [8].

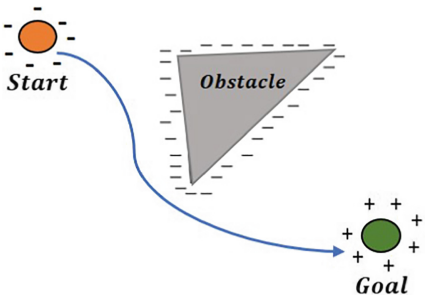


Fig. 1. Artificial Potential Field (APF) concept

VFF and VFH on the other hand are techniques specifically introduced to account for sensor uncertainties including crosstalk and spurious readings, as well as modeling errors. These collision avoidance methods are accomplished by creating certainty or histogram grids for obstacle representation which then allows the robots to continuously move without having to stop when obstacles are detected [9]. In environments with relatively bigger or wall-like obstacles, Bug algorithm is usually more preferred as it has the capabilities to follow boundaries and store memories. The basic idea of this insect-inspired algorithm is to follow a straight line towards the target and circumvent any obstacle if it is blocking the way. Despite their systematic way to reach the goal, these methods are only limited to two-dimensional scenes [10,15]. To account for multidimensional configuration spaces, other methods were developed such as the APF. This method is consistent with the principle that any objects will attract or repel when interacting with the robots during their navigations. In [12], the APF-based algorithm is applied using ultrasonic sensor where the travelling times of the transmitted pulses are used to derive the APF. In the case of point to point tracking as illustrated as in Fig. 1, the attractive force is generated by the goal coordinate and the repulsive force is exerted by the obstacle. Various mathematical descriptions of APF have been proposed in the literature, and one of the most commonly adopted is presented in [16]. If $q = (x, y)$ is the point representing the mobile robot in a given two-dimensional coordinate space, and $U_{att}(q)$ is the attractive potential field function generated by the target point, then

$$U_{att}(q) = \frac{1}{2}\eta\rho^k(q, q_g), \quad k = 1, 2$$

where η is a proportional gain coefficient, q_g is the goal's coordinate, and $\rho(q, q_g) = ||q - q_g||$ is the Euclidean distance between q and q_g. The value of k determines the shape of the function, i.e. when $k = 1$, the field presents a

fridaya@uef.fi

conic shape whereas when $k = 2$, the field presents a parabolic shape. In the latter case, the corresponding attractive force is determined by the negative gradient of the attractive potential, i.e.

$$F_{att}(q) = - \bigtriangledown U_{att}(q) = -\eta(q - q_g)$$

which decays linearly to zero as the robot approaches the goal. The repulsive potential field function generated by the obstacle as proposed in [16] takes the form

$$U_{rep}(q) = \begin{cases} \frac{1}{2}\eta \left[\frac{1}{\rho(q, q_{ob})} - \frac{1}{\rho_0}\right]^2 & \text{for} \quad \rho(q, q_{ob}) \leq \rho_0 \\ 0 & \text{for} \quad \rho(q, q_{ob}) > \rho_0 \end{cases}$$

where η is a proportional gain coefficient, q_{ob} is the obstacle reference point, $\rho(q, q_{ob}) = ||q - q_{ob}||$ is the shortest distance between q and q_{ob}, ρ_0 is a positive constant that represents the distance of the obstacle's influence. The corresponding repulsive force $F_{rep}(q)$ is then determined by

$$F_{rep}(q) = - \bigtriangledown U_{rep}(q) = \begin{cases} \frac{1}{2}\eta \left[\frac{1}{\rho(q, q_{ob})} - \frac{1}{\rho_0}\right] \frac{q - q_0}{\rho^3(q, q_{ob})} & \text{for} \quad \rho(q, q_{ob}) \leq \rho_0 \\ 0 & \text{for} \quad \rho(q, q_{ob}) > \rho_0 \end{cases}$$

This in turn enables the robot to move towards the goal point while avoiding any collision with enroute obstacles.

Although most of the aforesaid techniques are completely matured and can be well simulated using computers, the applications of the algorithms on real mobile robots may not neccessarily give the expected outcomes. This is largely due to limitations of the robots such as actuator constraints and external factors that include frictions and sensor noise. In this paper, the focus is on design and development of the APF-based obstacle avoidance strategy of a nonholonomic wheeled mobile robot with motion constraints. The method was validated via a series of experiments with different goal coordinates and number of obstacles. From the numerical and experimental results, we showed the effectiveness of the proposed strategy as it was able to drive the mobile robot to the goal coordinates with very small steady-state position errors without collisions with enroute obstacles.

2 Methodology

A nonholonomic wheeled mobile robot is considered in this work which is equipped with a forward-looking sonar sensor for obstacle detection. The movement of the robot is constrained in the sense that it is either moving with a translational velocity or a rotational velocity. Figure 2 illustrates the mobile robot in an xy plane where orange line arrow represents the translational velocity, and θ_1 is the heading angle of the robot which is also the angle from the x-axis.

fridaya@uef.fi

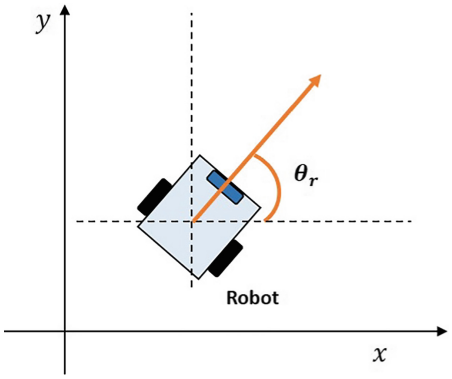


Fig. 2. Sketch of the nonholonomic wheeled mobile robot with its sensing area.

The rotational movement of the robot is constrained as it only moves to a certain set of θ_r, i.e. $\theta_r \in \Omega$ where

$$\Omega = \{\theta_r = \frac{k\pi}{8}, \quad k = 0, 1, \ldots, 8\}$$

To this end, we use a time-based method to control the orientation of the robot. Let $\bar{x} = (\omega_L, \omega_R, \delta t)$ be the state variables where ω_L, ω_R and δt are the left wheel speed, right wheel speed and the execution time respectively. The state $\bar{x}$ is controlled as follows:

$$\bar{x} = \begin{cases} (\omega_f, \omega_f, \Delta t_1) & \text{if} \quad \theta_r = 0 \\ (\omega_0, \omega_f, \Delta t_2) & \text{if} \quad \theta_r = \pi/4 \\ (\omega_0, \omega_f, \Delta t_3) & \text{if} \quad \theta_r = \pi/2 \\ (\omega_f, \omega_f, \Delta t_4) & \text{if} \quad \theta_r = 3\pi/4 \\ (\omega_b, \omega_b, \Delta t_1) & \text{if} \quad \theta_r = \pi \\ (\omega_f, \omega_0, \Delta t_4) & \text{if} \quad \theta_r = 5\pi/4 \\ (\omega_f, \omega_0, \Delta t_3) & \text{if} \quad \theta_r = 3\pi/2 \\ (\omega_f, \omega_0, \Delta t_2) & \text{if} \quad \theta_r = 7\pi/4 \end{cases} \tag{1}$$

where ω_f and ω_b are prespecified forward and backward angular velocities respectively, $\omega_0 = 0$, and Δt_i $(i = 1, 2, 3, 4)$ are prespecified execution times that match with the targeted heading angle θ_r.

The basic APF concept with a mobile robot, an obstacle and a goal point is illustrated in Fig. 3 where F_{att} is the attraction force of the target or goal point while the F_{rep} is the repulsive force from the obstacle. F_r is the resultant force between the attractive force and the repulsive force. In this work, all the forces are replaced with the robot's velocities as they are proportional to each other. The distance from the robot to the obstacle (as detected by the sensor) will be used to describe the strength of the repulsive velocity. In other words, the closer the robot to the obstacle, the bigger the repulsive velocity.

fridaya@uef.fi

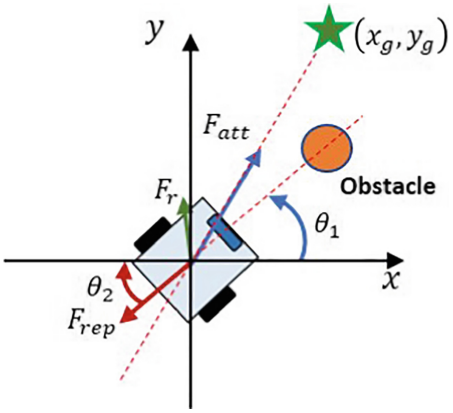


Fig. 3. Overview of the nonholonomic wheeled mobile robot with its sensing area.

Let $\bar{V}_1$ and $\bar{V}_2$ replace F_{att} and F_{rep} respectively. From Fig. 3, V_1 is the magnitude of the targeted translational velocity of the robot with angle θ_1 from the x axis ($\bar{V}_1 = V_1 \angle \theta_1$). When the sensor detects an obstacle on the left side of the robot, the repulsive speed with magnitude V_2 will be created with a direction θ_2. Let d be the distance from the robot to the obstacle, and ϵ_2 be the minimum distance allowed. We define V_2 as follows:

$$V_2 = \begin{cases} 0 & \text{if} \quad d > \epsilon_2 \\ -V_a & \text{if} \quad \epsilon_1 < d < \epsilon_2 \\ -V_b & \text{if} \quad \epsilon_0 < d < \epsilon_1 \end{cases} \tag{2}$$

where $\epsilon_0 = 0$, $\epsilon_1 < \epsilon_2$, and $V_b > V_a > 0$. The direction of the repulsive velocity is simply

$$\theta_2 = \theta_1 + \pi$$

Hence, the repulsive direction θ_2 is always π rad greater than θ_1. The resultant velocity, $\bar{V} = V \angle \theta$ can be calculated as follows:

$$\begin{aligned} V_{1x} &= V_1 \cos\theta_1 \\ V_{2x} &= V_2 \cos\theta_2 \\ V_{1y} &= V_1 \sin\theta_1 \\ V_{2y} &= V_2 \sin\theta_2 \end{aligned}$$

The velocities in the x and y axes directions will be

$$\begin{aligned} V_x &= V_{1x} + V_{2x} \\ V_y &= V_{1y} + V_{2y} \end{aligned}$$

fridaya@uef.fi

respectively. The magnitude and angle of the resultant velocity can then be calculated as

$$V = \sqrt{V_x^2 + V_y^2}$$
$$\theta = \arctan \frac{V_y}{V_x}$$

In this work, θ will be the command signals to the robot as described in (1). Hence, the robot will move with velocity described in the equation. In the case where the resultant velocity is equal to zero, the sensor will rotate to the nearest direction. Then, a new resultant velocity will be computed. If there is no obstacle, the robot will move towards the goal point. The overall obstacle avoidance algorithm can be illustrated as in Figs. 4 and 5 where (x_g, y_g) is the desired goal coordinate.

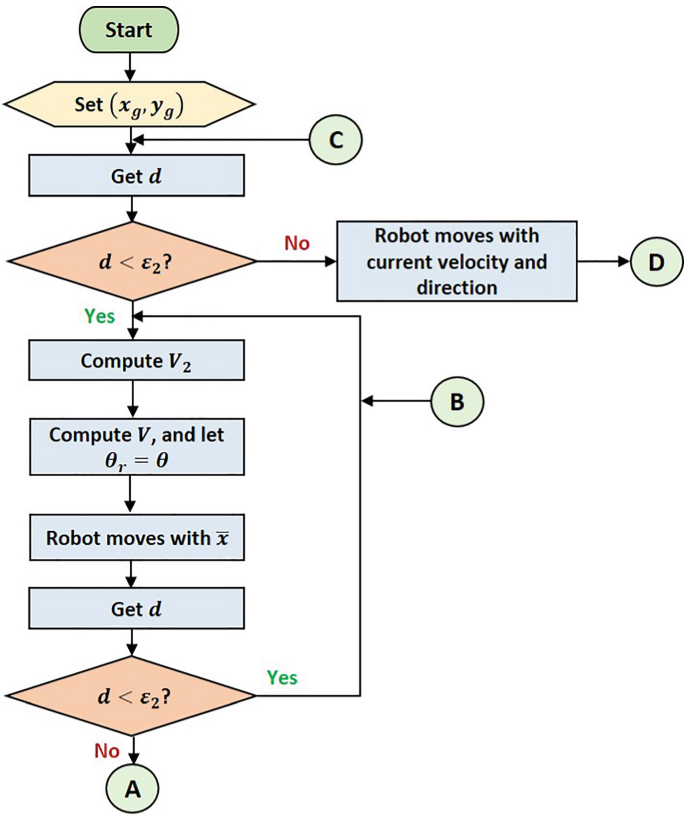


Fig. 4. Obstacle avoidance with a simplified APF algorithm.

fridaya@uef.fi

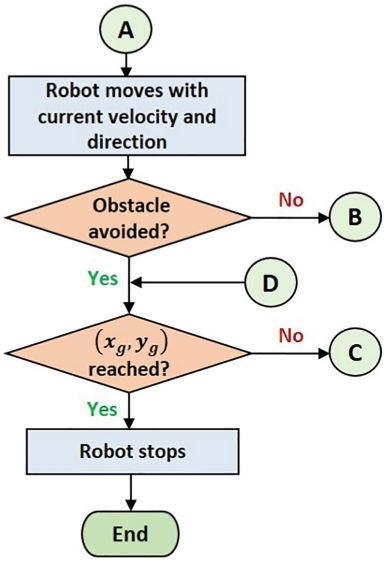


Fig. 5. Obstacle avoidance with a simplified APF algorithm *(continued)*.

3 Experimental Results

Figure 6 shows the nonholonomic wheeled mobile robot used for the experiments which is equipped with a forward-looking sonar sensor to compute the resultant command velocity. We used bluetooth for wireless data acquisitions and MATLAB to plot the robot's path. From (1), we define

$$\omega_f = 20\,\text{cm/s}, \quad \omega_b = 20\,\text{cm/s},$$
$$\Delta t_1 = 0\,\text{s}, \quad \Delta t_2 = 0.6\,\text{s}, \quad \Delta t_3 = 1\,\text{s}, \quad \Delta t_4 = 1.6\,\text{s},$$

and for (2), we set $V_a = 20$ cm/s, $V_b = 40$ cm/s, $\epsilon_0 = 0$ cm, $\epsilon_1 = 10$ cm, and $\epsilon_2 = 20$ cm.

A two-dimensional space of $150\,\text{cm} \times 150\,\text{cm}$ was used and three desired goal coordinates were considered for the experiments, i.e. $(x_g, y_g) = (100, 0)$, $(x_g, y_g) = (120, 50)$ and $(x_g, y_g) = (60, -60)$. For each goal coordinate, the robot started at the origin (i.e. $(0, 0)$), and two conditions were tested, i.e. when there was one obstacle and when there were two obstacles which were placed randomly between the starting and goal coordinates.

For Experiment 1 (i.e. $(x_g, y_g) = (100, 0)$), the path of the mobile robot for the two conditions are shown in Fig. 7, whereas for Experiment 2 ($(x_g, y_g) = (120, 50)$) and Experiment 3 ($(x_g, y_g) = (60, -60)$), the path of the mobile robot are shown in Figs. 8 and 9 respectively. The small red circle indicates the starting position while the small green circle denotes the goal point.

Let (x_s, y_s) be the final position of the mobile robot from the sensor data, and (x_m, y_m) be the final position of the mobile robot from the real measurement.

fridaya@uef.fi

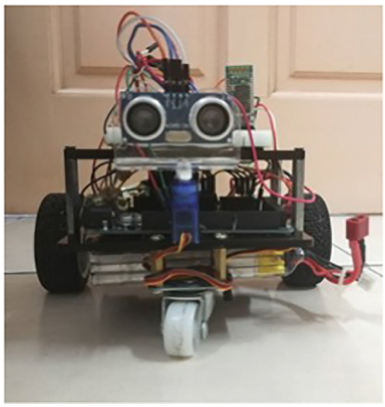
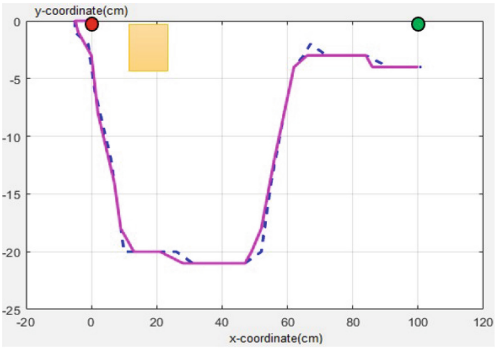
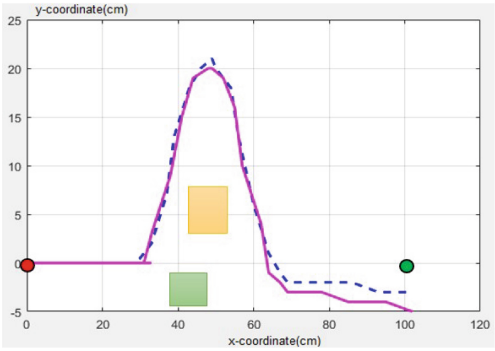


Fig. 6. The nonholonomic wheeled mobile robot used for the experiments.



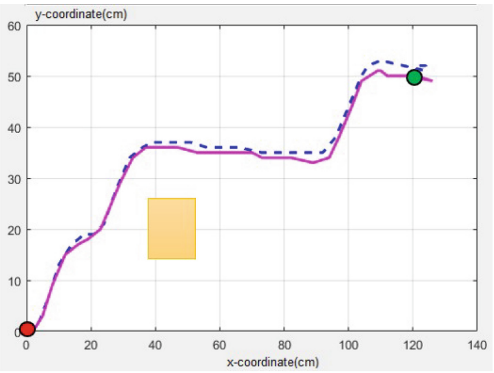
(a) One obstacle



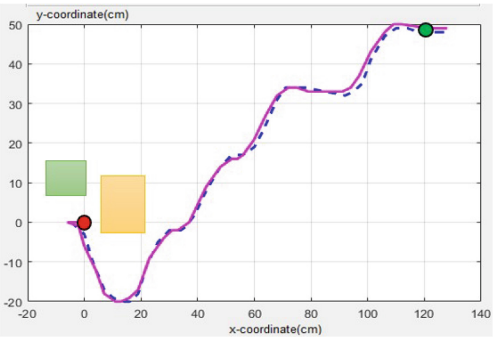
(b) Two obstacles

Fig. 7. Experiment 1: goal at $(100, 0)$.

fridaya@uef.fi



(a) One obstacle



(b) Two obstacles

Fig. 8. Experiment 2: goal at $(120, 50)$.

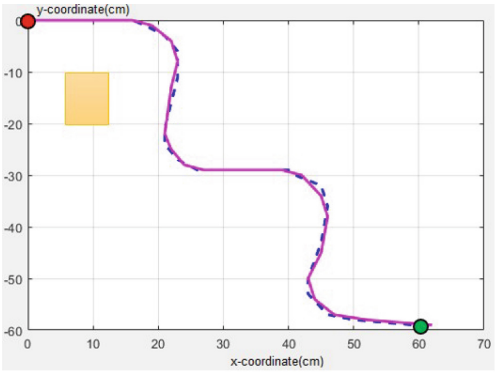
Table 1. Final positions of the mobile robot from Experiment 1 from the sensor data and real measurements.

Test	O1		O2	
	(x_s, y_s)	(x_m, y_m)	(x_s, y_s)	(x_m, y_m)
1	$(101, -4)$	$(100, -4)$	$(102, -3)$	$(102, -5)$
2	$(101, -3)$	$(101, -2)$	$(103, -6)$	$(102, -4)$
3	$(101, -4)$	$(101, -4)$	$(102, -3)$	$(103, -1)$
Average	$(101, -3.7)$	$(100.7, -3.3)$	$(102.3, -4)$	$(102.3, -3.3)$

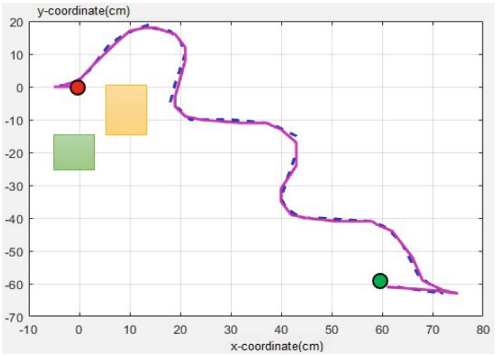
For each experiment, three tests were carried out and the results including the average final positions are tabulated in Tables 1, 2 and 3. We write "O1" to indicate one obstacle, and "O2" to indicate two obstacles.

Define e_s and e_m as steady-state position errors from the sensor data and real measurements respectively, the results for the three experiments are tabulated in Table 4. From the results of all experiments, it can be observed that the simplified

fridaya@uef.fi



(a) One obstacle



(b) Two obstacles

Fig. 9. Experiment 3: goal at $(60, -60)$.

Table 2. Final positions of the mobile robot from Experiment 2 from the sensor data and real measurements.

Test	O1		O2	
	(x_s, y_s)	(x_m, y_m)	(x_s, y_s)	(x_m, y_m)
1	$(122, 52)$	$(120, 52)$	$(121, 48)$	$(119, 49)$
2	$(122, 52)$	$(121, 50)$	$(121, 48)$	$(118, 52)$
3	$(122, 55)$	$(123, 55)$	$(121, 48)$	$(119, 50)$
Average	$(122, 53)$	$(121.3, 52.3)$	$(121, 48)$	$(118.7, 50.3)$

APF algorithm proposed is able to drive the robot to the goal points without colliding with the enroute obstacles. The resulting steady-state position errors via sensor data and real measurements are considered quite small as they are no greater than 5 cm.

fridaya@uef.fi

Table 3. Final positions of the mobile robot from Experiment 3 from the sensor data and real measurements.

Test	O1		O2	
	(x_s, y_s)	(x_m, y_m)	(x_s, y_s)	(x_m, y_m)
1	$(61, -59)$	$(61, -59)$	$(61, -61)$	$(60, -61)$
2	$(61, -59)$	$(62, -59)$	$(61, -61)$	$(62, -60)$
3	$(61, -59)$	$(64, -59)$	$(61, -61)$	$(61, -61)$
Average	$(61, -59)$	$(63.3, -58.3)$	$(61, -61)$	$(61, -60.7)$

Table 4. Steady-state position errors for Experiments 1,2 and 3 from the sensor data and real measurement. All values are in "cm".

Test	Exp. 1				Exp. 2				Exp. 3			
	O1		O2		O1		O2		O1		O2	
	e_s	e_m	e_s	e_m	e_s	e_m	e_s	e_m	e_s	e_m	e_s	e_m
1	4.12	4.00	3.61	5.39	2.83	2.00	2.24	1.41	1.41	1.41	1.41	1.00
2	3.16	2.24	6.71	4.47	2.83	1.00	2.24	1.00	1.41	2.24	1.41	2.00
3	4.12	4.12	3.61	3.16	5.39	5.83	2.24	1.00	1.41	5.00	1.41	1.41
Average	3.80	3.45	4.64	4.34	3.68	2.94	2.24	1.75	1.41	2.88	1.41	1.47

4 Conclusion and Future Works

In this work, a simplified APF-based algorithm has been proposed for the obstacle avoidance strategy of a nonholonomic wheeled mobile robot with motion constraints. Experiments with different goal points and number of obstacles were conducted to test the performance of the proposed algorithm, and it has been shown that the robot was able to avoid collisions with enroute obstacles and reach the destinations with very small steady-state position errors.

For future works, the smoothness of the trajectory may be further improved by modifying the constraints of the rotational and translational movements of the robot, and different type of sonar sensors may be employed to increase to accuracy of the measurement of the potential field or the resultant velocity.

Acknowledgments. The authors would like to thank Universiti Sains Malaysia for the financial support under the Research Universiti (RUI) Grant Scheme (1001/PELECT/8014055).

References

1. Simoens, P., Dragone, M., Saffiotti, A.: The internet of robotic things: a review of concept, added value and applications. Int. J. Adv. Rob. Syst. **15**(1), 1–11 (2018)

fridaya@uef.fi

2. Xidias, E.K., Aspragathos, N.A., Azariadis, P.N.: Robotics: Concepts, Methodologies, Tools, and Applications: Mission Planning of Mobile Robots and Manipulators for Service Applications. IGI Global, New York (2014)
3. Hoy, M., Matveev, A.S., Savkin, A.V.: Algorithms for collision-free navigation of mobile robots in complex cluttered environments: a survey. Robotica **33**(3), 463–497 (2015)
4. Siegwart, R., Nourbakhsh, I.R.: Introduction to Autonomous Mobile Robots. A Bradford Book. The MIT Press, Cambridge (2004)
5. Skogestad, S., Postlethwaite, I.: Multivariable Feedback Control. Wiley, Hoboken (2005)
6. Özdemir, A., Sezer, V.: A hybrid obstacle avoidance method: follow the gap with dynamic window approach. In: 1st IEEE International Conference on Robotic Computing (IRC), pp. 257–262, April 2017
7. Fox, D., Burgard, W., Thrun, S.: The dynamic window approach to collision avoidance. IEEE Robot. Autom. Mag. **4**(1), 23–33 (1997)
8. Hong, Z., Chun-Long, S., Zi-Jun, Z., Wei, A., De-Qiang, Z., Jing-Jing, W.: A modified dynamic window approach to obstacle avoidance combined with fuzzy logic. In: 14th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), pp. 523–526, August 2015
9. Borenstein, J., Koren, Y.: Real-time obstacle avoidance for manipulators and mobile robots. IEEE Trans. Robot. Autom. **7**(3), 278–288 (1991)
10. Lumelsky, V.J., Stepanov, A.A.: Path-planning strategies for a point mobile automation moving amidst unknown obstacles of arbitrary shape. Algorithmica **2**(1), 403–430 (1987)
11. Khatib, O.: Real-time obstacle avoidance for manipulators and mobile robots. Int. J. Robot. Res. **5**(1), 90–98 (1986)
12. Veelaert, P., Bogaerts, W.: Ultrasonic potential field sensor for obstacle avoidance. IEEE Trans. Robot. Autom. **15**(4), 774–779 (1999)
13. Xie, L., Chen, H., Xie, G.: Artificial Potential Field Based Path Planning for Mobile Robots Using Virtual Water-flow Method. Springer, Heidelberg (2007)
14. Harik, E.H.C., Korsaeth, A.: Combining hector slam and artificial potential field for autonomous navigation inside a greenhouse. Robotics **7**(2), 22 (2018)
15. Lozano-Pérez, T., Wesley, M.A.: An algorithm for planning collision-free paths among polyhedral obstacles. Commun. ACM **22**(10), 560–570 (1979)
16. Ge, S.S., Cui, Y.J.: New potential functions for mobile robot path planning. IEEE Trans. Robot. Autom. **16**(5), 615–620 (2000)

fridaya@uef.fi

Immersive Analytics Through HoloSENAI MOTOR Mixed Reality App

André L. M. Ramos[1], Thiago Korb[2(✉)], and Alexandra Okada[3]

[1] SENAI, Tubarão, Brazil
[2] SENAI, Florianópolis, Brazil
thiagokorb@sc.senai.br
[3] Open University, Milton Keynes, UK
ale.okada@open.ac.uk

Abstract. This study examines the use of HoloSENAI MOTOR and LiteMap annotation as novel approach for preparing learners and professionals for the industry 4.0. This new mixed reality technology was developed with UNIT3D and C# language for the Microsoft HoloLens®. This educational resource enables the projection of 3D scenes of a real electric motor into the natural world environment. It was used by students from an Engineering course in Brazil. Our aim was to identify the potential benefits and barriers to promote immersive analytics and authoring skills through HoloSENAI MOTOR for learning and teaching. Immersive Analytics is a novel framework that combines real-time interaction with visualization techniques for students to explore and analyze information about the motor in their physical environment. This study is based on Responsible Research and Innovation approach. It revealed that the key benefits for learners were to increase their motivation, curiosity and understanding of features, properties and functionalities of the motor, including better acquisition of information and data analysis skills fun learning. The key barriers were the expensive equipment, time and staff for developing other APPs and new pedagogical approaches for assessment.

Keywords: Immersive analytics · Microsoft HoloLens · C# · Fun learning · Responsible Research and Innovation · Annotation · Mapping · VISION tools

1 Introduction

This study presents the HoloSENAI MOTOR which is a Mixed Reality application developed with UNIT3D and C# language for the Microsoft HoloLens®. This educational resource was created by the technical team from SENAI to help teaching and learning in the area of engineering and computer science.

This study examines Mixed Reality and Visual Analytics which were connected together to promote immersive analytics (IA) through pedagogical activities. Immersive Analytics is a learner centered approach that combines real-time interaction with visualization techniques for students to explore and analyze information about complex objects in their physical environment. Immersive analytics aims to create meaningful opportunities for students to develop analytical reasoning and decision making by

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1259–1268, 2019.
https://doi.org/10.1007/978-3-030-22868-2_89

fridaya@uef.fi

immersing themselves on informational environment with real-world analytics tasks. Immersive Analytics builds on augmented reality environments, 3D large touch surfaces, 360 virtual environments with sensor devices and natural user interface devices.

2 Background

Immersive learning with mixed realities and holograms is a recent field in Education. The Microsoft HoloLens was released in March 2016, therefore literature is very limited.

Mixed reality is a novel approach for bringing students, places, and objects together from physical and digital worlds. It combines the immersive space of Virtual Reality (VR) with the real-world environment of Augmented Reality (AR) with authentic objects for responsive interaction.

The holographic mixed reality with the Microsoft HoloLens is a platform that provides a personalised immersion with a head-mounted display. Users can explore physical real-world and manipulate virtual components through "holograms". Holograms are three-dimensional images formed by the interference of light beams from a light source. Content and information can be linked to the holograms and become available only when users want to access it.

The interactive experiences with holograms occur through built-in sensors. Users can use their gaze to move the cursor and select holograms, information and data. Through their gestures, they can select and size items as well as drag and drop holograms in their world. Their voice can be used to navigate, select and command the app.

Through Augmented Reality physical objects are overlaid with real scenes to create the opportunity for students to explore and acquire information whenever they need. VR objects are merged with information and data to create an opportunity for interactive visualisation in context [1]. The key issue in AR and computer-generated information is the limited layout of objects which may cause reduce the visualization of informational zones.

A growing field related to AR is VA – visualization approaches, which offers interactive methods to solve the challenges of big data. Its aim is to reduce the complexity of data presentation in the limited mixed reality layout (see Fig. 1). The VA tools are designed for users to visualise the patterns and relationships within datasets. The key difference between VA and traditional interactive tools is focusing on open-ended interactive exploration of data, which is combined with sophisticated analytics [2, 3].

Mixed realities which combines AR, VR and VA offers to students a clear interface, whose layers of data are presented according to their needs, so they can explore based on their own perspective and having access to information and data 'context-driven'. Studies about the adoption of VA concepts for data exploration as a display space revealed a potential approach to support learning goals including situated analytical thinking skills [3, 4].

fridaya@uef.fi

Fig. 1. Punch Club Studio (2018) Making of: Dimension Data in Assoc. with Microsoft HoloLens AV – VFX (https://vimeo.com/221232251)

3 HoloSENAI MOTOR

The HoloSENAI MOTOR was designed by a multidisciplinary team of SENAI with UNIT3D, C# language and the Microsoft HoloLens system [5, 6]. SENAI refers to the National Industrial Learning Service, which is supported by the Brazil's national industry confederation (CNI). SENAI is one of the five largest professional education systems in the world and the largest in Latin America. One of its aim is to stimulate industrial innovation through education, consulting, applied research and technical and technological services.

The HoloSENAI MOTOR was the first APP with Microsoft HoloLens in Brazil. It was developed to enable students to visualise and explore the internal structure of an electrical motor equipment during the technical course about Automation and Mechatronics. This open educational APP provides an interactive and responsive environment for the future professional engineers. It enables students to visualize and work with digital content as part of a real professional world. Data and information were connected to holograms for each student to interact naturally within this hybrid environment based on their needs, interests and pace.

Students frequently consider equipment such as motors and valves as black boxes, so what happens inside is a mystery for them. Teaching complex topics with abstract concepts requires various lessons. The immersive interaction between students and AR might increase their understanding of motors' operation, features, functionalities, and properties. However, literature about immersive technologies show that students might feel dizzy or nauseated with AR or VR headset [7, 8].

The HoloSENAI MOTOR lesson was planned by the pedagogical team for each student to view and manipulate the equipment and analyse data inside and outside

fridaya@uef.fi

when equipment is in operation. Students should report their positive or negative experience.

This exploratory study examines three questions:

What are the students' views about the HoloSENAI MOTOR?

What are the teachers' opinions about learning gains and barriers?

What is the impact of HoloSENAI MOTOR for immersive analytics?

Fig. 2. HoloSENAI MOTOR –immersive analytics experience https://www.youtube.com/watch?v=nnebPmlqmw0

4 Methodology

This study follows the approach of Responsible Research and Innovation (RRI) by which researchers and society discuss how technological innovations can better address the societal needs.

4.1 Participants

HoloSENAI MOTOR was used by a group of students in the course of Mechanical Engineering. There were 20 males and 3 females. In terms of age, there were seventeen students from 16 to 19 years old and four from 20 to 29 years old, one student from 30 to 39 years old and one student from 40 to 49 years old.

4.2 Method and Procedures

This exploratory work focused on users' learning experience with the AR HoloLens headset and HoloSENAI MOTOR. Students replied the pre-and post- semi-structure questionnaires and completed their learning activities with the Microsoft HoloLens headset. All of them authorised data collection for research purposes. Participants of the videoclip also signed the consent form to make the data public for dissemination of the learning experience with immersive analytics.

The HoloSENAI application (Fig. 2) was related to Electrotechnics and Electrical Machines Fundamentals unit. The learning objectives included being able explore data, information and knowledge about magnetism and electromagnetism in electric

fridaya@uef.fi

machines and electric motors of alternating current. Students should explore the electric motor – its parts and functions, analyze the concepts of magnetism and electromagnetism and explain how these forces contribute to the operation of the motor.

The learning situation for immersive analytics was guided by the following questions: How does an electric motor work? What is the relationship between magnetism and the electric motor? Which parts make up an electric motor? What role do these parts play within the engine?

Students received the instructions and explanations to visualize and manipulate an electric motor through the AR glasses. The manipulation should be done by voice command and by gestures made with their hands. Before using HoloSENAI, students saw a similar activity was developed by Microsoft, using Microsoft HoloLens, where the user should manipulate an airplane turbine.

4.3 Data Analysis

Data analysis were based on three categories to reflect about mixed realities with holograms in Education [1, 5].

- **Data integration**: Enhances the blending of virtual information with AR object
- **Scene manipulation**: Offers options to augment data (colour, zoom, position)
- **Context-driven**: Highlights visual information based on the context.

Students' reflection and views about the visual information were analysed through a set of inquiry skills for RRI [4] (Table 1):

Table 1. Inquiry skills for RRI – ENGAGE project [4]

Inquiry skills for RRI	Attributes
Devise questions **(?)**	refers to a problem or an issue
Communicate Ideas (!)	presents (informed) ideas related to the issue
Critique claims **(−)**	highlights counter-argument that refutes an idea
Justify opinions **(+)**	explains opinions linked to knowledge, facts or data
Examine consequences **(+/−)**	shows benefits (support) or concerns (cons)
Interrogate sources (Q)	shows details about reliable evidence

The LiteMap tool (Fig. 3) for annotation and collaborative mapping was used to analyse the discussion among participants and their views. Data from staff and students including videoclip were annotated, tagged and visualised with the LiteMap Visual Analytics Dashboard.

fridaya@uef.fi

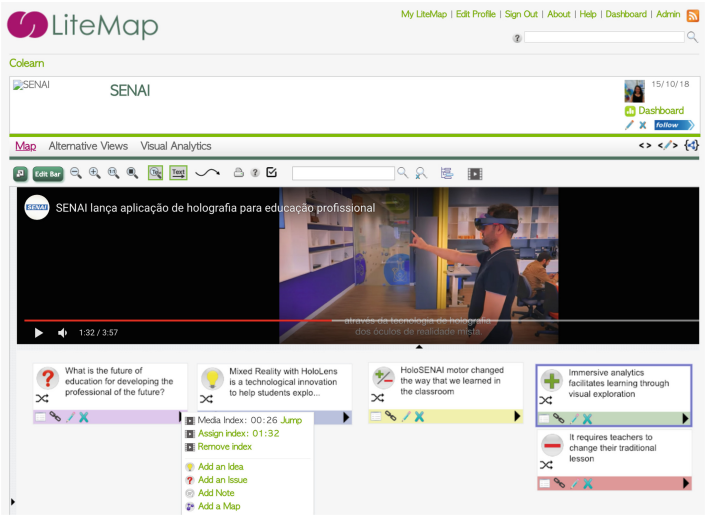


Fig. 3. LiteMap – collaborative annotation and mapping tool

5 Findings

Data from the pre-questionnaire revealed that 12 students were aware of different applications for immersive learning, 17 students knew about SENAI APPs developed by the Institution and 14 students had already used other APPs (see Fig. 4).

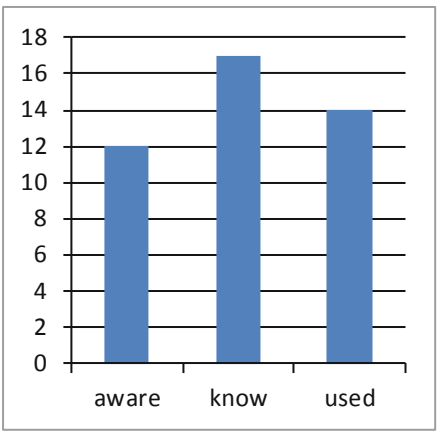


Fig. 4. Other mixed reality Apps

The majority of students agreed that immersive analytics with AR HoloLens headset increased their visual understanding of the “electric motor”.

fridaya@uef.fi

All students found it easy-to-use. They all highlighted it provided visible data, the content was significant and helpful. Most of them (23 out of 24) found the audio clear and 15 out of 24 found that this app cannot be replaced by other type of technology (see Fig. 5).

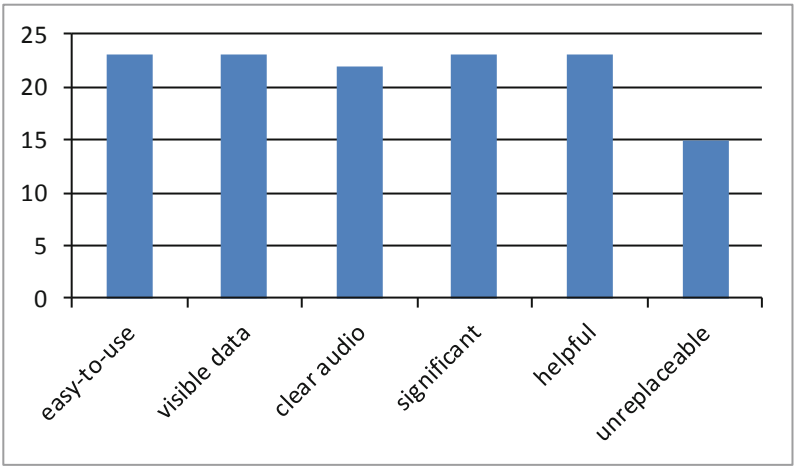


Fig. 5. The HoloMOTOR SENAI App

Data also revealed that the learning gains were positive in terms of acquisition of knowledge and skills. Various students were also able to provide examples. They found that HoloSENAI MOTOR was useful for them to develop data visualisation skills (5 out of 24). They also highlighted it facilitates data comprehension, supports critical thinking and meaningful learning.

In terms of potential barriers, both pedagogical and technological team listed some issues, such as the high cost equipment, the need of technical development of applications and the pedagogical approaches for assessment.

To identify the impact of AR HoloLens for immersive analytics, the Kalkofen et al. [1] categories were applied for thematic analysis.

In terms of **Data Integration**, students found the experience dynamic, innovative, immersive, and topical. The AR HoloLens "offered new ways of learning with excellent informational quality". It enabled them to explore "how the parts move and explore data when these are working without any danger". "It is an important learning tool for engineering learners and workers".

With reference to **Scene manipulation**, students found the interface easy-to-use and practical. They "had no effect of dizziness or anything similar" and "enjoyed the experience".

In respect of **Context-driven**, they found the immersive analytics "useful and fun for understanding abstract concepts". They found "easier to learn difficult concepts through visual information and "exploring the parts linked to data was cool, fun and useful way of learning".

fridaya@uef.fi

Data annotated and mapped in LiteMap from all participants (see Fig. 6) presented 9 questions, 14 ideas, 32 supporting arguments, 22 counter arguments, 6 arguments and 35 notes. Participants were very engaged with the APP, topic, methodology and map. LiteMap shows that 23 items of the dialogue map were viewed, 80 added, and 60 items edited.

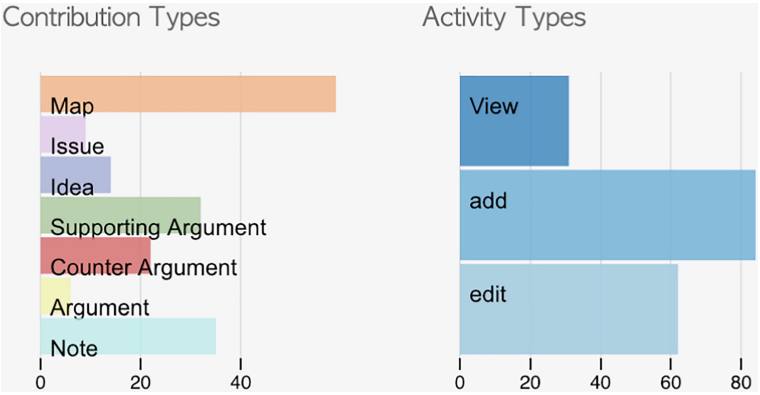


Fig. 6. LiteMap visual analytics dashboard.

6 Discussion

This study provided an overview about students and teachers' opinions of mixed reality with immersive analytics. The methodological approach adopted was exploratory research with the purpose of exploring the potential use of the HoloSENAI MOTOR to better equip students for the digital transformation for the industry 4.0.

One of the key strengths of mixed reality is to enable students to explore authentic scenarios augmented with interactive virtual components to enrich situated learning. The mixed reality with holograms enables them to learn with more engagement as it activates various senses through sensors: movement, voice, text, audio, holograms and visual information with data. This creates opportunities for learners to immerse in a hybrid environment to practice analytical thinking skills. This immersive learning environment, which are situated and in context, were never before available within the traditional lessons.

However, one of the key weaknesses about the HoloSENAI MOTOR highlighted by the institutional leaders was the cost of the Microsoft headset and the APP production. Although the APP can be produced under open license to open up more opportunities of immersive analytics for all, it takes time and requires trained team to produce a high-quality APP. In addition, the high cost of the Microsoft HoloLens headset makes very difficult to scale up this novel learning approach across the curriculum and courses.

To address these challenges, it is important to examine pedagogical strategies to make this experiment more open and inclusive; through collaboration within and

fridaya@uef.fi

outside the institution. Further studies will be necessary for reflecting on new ways that this research might be undertaken. More practical examples and experiments must be examined with other technologies, new pedagogical approaches for learning and embedded assessment.

One of the key achievements of this study was to examine that immersive analytics become unique and unreplaceable learner-centre approach. This occurs particularly in complex scenarios whose concepts are abstract, functionalities cannot be observed, and its components are intangible or inaccessible. However, a key issue is to define some criteria for selecting significant scenarios for immersive analytics as well as a framework to integrate assessment to ensure the quality of learning. Studies of HoloLens will be emerging in various contexts. Examples from practitioners [9] are growing particularly in the area of professional training, medicine, culture and arts. For instance, learners explore fine arts and objects from museums and galleries with holograms at home. They can also dance with holograms on a real stage.

7 Conclusions

In summary, this study presented an open educational resource developed with the Microsoft HoloLens in Brazil by SENAI, which is one of the biggest professional development centers for industry in the world.

AR applications with HoloLens for immersive analytics helped students develop knowledge and key skills such as visualizing information, manipulating objects and data and analyzing abstract concepts. However, the key challenge is to examine less expensive approaches for widening up professional development in Brazil for digital transformation. Our next study, which is part of the RRIdata.com VISION project focuses on immersive analytics for open and inclusive STEM education [4].

Acknowledgments. This study is funded by eSTEeM – The OU centre for STEM pedagogy. Project Reference 17E-AOADLAB-LTIKMIEI-01.

References

1. Kalkofen, D., Sandor, C., White, S., Schmalstieg, D.: Visualization Techniques for Augmented Reality. Springer-Verlag, New York (2011)
2. Chandler, T., Cordeil, M., Czauderna, T., Dwyer, T., Wilson, E.: Immersive analytics. In: 2015 Big Data Visual Analytics (BDVA), pp. 1–8. IEEE (2015)
3. ElSayed, N., Thomas, B., Marriott, K., Piantadosi, J., Smith, R.: Situated analytics. In: 2015 Big Data Visual Analytics (BDVA), pp. 1–8. IEEE, Hobart, Australia (2015)
4. Okada, A., Kowalski, R.P.G., Kirner, C., Torres, P.L.: Factors influencing teachers' adoption of AR inquiry games to foster skills for Responsible Research and Innovation. Interact. Learn. Environ. **27**(3), 324–335 (2019)
5. Chen, H., Lee, A.S., Swift, M., Tang, J.C.: 3D collaboration method over HoloLens™ and Skype™ end points. In: Proceedings of the 3rd International Workshop on Immersive Media Experiences, pp. 27–30. ACM (2015)

fridaya@uef.fi

6. Evans, G., Miller, J., Pena, M.I., MacAllister, A., Winer, E.: Evaluating the Microsoft HoloLens through an augmented reality assembly application. In: Degraded Environments: Sensing, Processing, and Display 2017, vol. 10197, p. 101970V). International Society for Optics and Photonics (2017)
7. Cui, N., Kharel, P., Gruev, V.: Augmented reality with Microsoft HoloLens holograms for near infrared fluorescence based image guided surgery. In: Molecular-Guided Surgery: Molecules, Devices, and Applications III, vol. 10049, p. 100490I. International Society for Optics and Photonics (2017)
8. Kalantari, M., Rauschnabel, P.: Exploring the early adopters of augmented reality smart glasses: The case of Microsoft HoloLens. In: Augmented Reality and Virtual Reality, pp. 229–245. Springer, Cham (2018)
9. Workman, S.: Mixed reality: a revolutionary breakthrough in teaching and learning. Educase Review (2018). https://er.educause.edu/articles/2018/7/mixed-reality-a-revolutionary-breakthrough-in-teaching-and-learning

fridaya@uef.fi

Explainable Artificial Intelligence Applications in NLP, Biomedical, and Malware Classification: A Literature Review

Sherin Mary Mathews(✉)

McAfee LLC, Santa Clara, CA 95054, USA
sherin_mathews@mcafee.com, sherinm@udel.edu

Abstract. Deep learning algorithms have achieved high performance accuracy in complex domains such as image classification, face recognition sentiment analysis, text classification, and speech understanding. Due to the nested non-linear structure of deep learning algorithms, these highly successful models are usually applied in a black-box manner, i.e., no information is provided about what exactly causes them to arrive at their predictions. The effectiveness of these systems is thus limited by the machine's current inability to explain its decisions and actions to human users. Such a lack of transparency can be a major drawback. For instance, in medical applications the development of methods for visualizing, explaining, and interpreting deep learning models has recently attracted increasing attention. Explainable Machine Learning (XAI) or Interpretable Machine Learning (IML) programs aim to create a suite of machine learning techniques that produce more explainable models while maintaining the high level of accuracy. It also enables human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners. Our work summarizes recent developments in this field, and we present three classification tasks wherein we make use of LIME (Local Interpretable Model-Agnostic Explanations) to explain predictions of deep learning models. LIME attempts to make these complex models at least partly understandable by evaluating using three classification tasks. Our first evaluation is performed on natural language processing application for tweet data classification and our second evaluation is on biomedical signal classification cancer detection. Our third evaluation is on Windows PC malware classification in a cybersecurity domain. In all three case studies, the interpreted results are presented in terms of visualization.

Keywords: Explainable machine learning · Interpretable machine learning · Explainable artificial intelligence · Explainability · Interpretability · Deep neural networks · Black-box models · Sensitivity analysis · Tumor classification · Tweet classification · Natural language processing · Biomedical signal classification · Windows PC malware classification

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1269–1292, 2019.
https://doi.org/10.1007/978-3-030-22868-2_90

fridaya@uef.fi

1 Introduction

1.1 Background

The artificial intelligence (AI) field is booming, and according to one survey of experts at a recent conference [1], it is predicted that by 2050, AI through intelligence machines will be able to perform any intellectual task a human can perform. AI and data-focused machine learning (ML) have been able to achieve exceptional performance on important and difficult learning problems. Progress in automation, self-management systems, voice recognition, and deep learning (DL) has enabled computers to make a series of major breakthroughs that were previously impossible. For example, ML algorithms have recently shown remarkable success at learning to play ATARI games [2], categorizing images [3], writing poetry automatically [4], beating professional poker players, developing scientific theory, defeating champions at the Go Game [5] and even excelling humans on IQ tests [4]. ML technologies have been able to achieve impressive accuracies in computer vision for applications such as recognizing objects in images, recognizing emotions in images of faces, recognizing emotions in speech, aiding translation between languages, assessing autonomous driving, and flying a drone in a field. Deep learning (DL) algorithms have achieved near-human accuracy levels in various types of prediction tasks and classifications including those of images of text, speech, and video data. Advances in ML and DL have had a profound impact on many "low-level" tasks such as object recognition and behavior monitoring. Researchers have also begun to explore how these approaches can be used in "high-level" domains such as healthcare, the criminal justice system, finance, security, and military decision making [6].

Healthcare is one of the high-profile areas in which the successes of AI have demonstrated the greatest practical value. AI has specifically revolutionized the medical field of cancer diagnosis. Also, it has been shown to be efficient in discovering new uses for existing drugs, revealing cancer in tissues, detecting cardiac arrhythmia, and predicting hypoglycemic events in diabetics three hours in advance of the medical industry average. DL is a powerful and widely-used tool for making potentially important decisions in a wide range of medical diagnosis issues. An example of the impact of machine learning is as follows: The IBM supercomputer Watson is already able to see deviation in the health of an individual which has eluded the attention of experienced diagnosticians. Statistically, in about 30% of its cases, Watson provides additional diagnoses in patients which have been missed by medical people [7].

Ml and DL algorithms also offer unparalleled flexibility and high accuracy in a variety of security domain applications such as intrusion detection systems, spam e-mail filtering, malware detection, malicious application detection, and identity theft prevention. Such algorithms have been reporting impressive performances on benchmark datasets, even when trained only on statistical features extracted from the application, such as system calls and permissions (Table 1).

fridaya@uef.fi

Table 1. Summary of recent works carried out using deep learning algorithms in different domains

Author	Application	DL Algorithm	Accuracy
Sun et al. [8]	Action recognition	Spatio Temporal Conv. Nets	88.1%
Ng et al. [9]	Action recognition	Conv. Pooling +LSTM	88.6%
Geng et al. [10]	Gesture recognition	CNN	99.5%
Fraiwan et al. [11]	Sleep state identification	Autoencoders	80.4%
Schirrmeister et al. [12]	EEG decoding	CNN	89.8%
Spampinato et al. [13]	Discriminate brain activity	CNN	86.9%
Acharya et al. [14]	Seizure detection	CNN	88.67%
Zheng et al. [15]	Emotion classification	DBN	87.62%
Lou et al. [16]	Heartbeat classification	DNN	97.5%
Acharya et al. [17]	Heart beat classification	CNN	94.03%
Muduli et al. [18]	Fetal-ECG signal reconstruction	Stacked Autoencoder	99.5%
Acharya et al. [19]	Arrhythmia detection	CNN	92.5%
Zheng et al. [20]	ECG Identification	DNN	94.39%
Ye et al. [21]	Malware Detection	DeepAM	98.2%
Yuan et al. [22]	Android malware detection	Deep Belief Networks	96.76%
Yin et al. [23]	Intrusion detection	RNN	83.49%
Shone et al. [24]	Intrusion detection	Deep Autoencoder	97.90%
Muna et al. [25]	Anomaly detection system	Deep Auto-Encoder	99%

Despite the impressive performance of ML algorithms, it can make mistakes, some more costly than others. As such, issues of trust and safety of ML are important themes [26–28]. While improving overall accuracy has been the main goal of the ML community's effort, it may not be enough; we need also to better understand the strengths and limitations of ML techniques.

1.2 Significance: The Need for Explainable Artificial Intelligence (XAI)

Impressive accomplishments in machine learning have led to an outburst of computer vision, natural language processing, and AI applications. These advancements have resulted in autonomous systems that perceive, learn, decide, and act on their own. AI, however, makes use of deep look-ahead search and stochastic optimization of huge neural network techniques to fit equally huge datasets, often resulting in complex behavior difficult for people to understand. A key impediment to the use of AI-based systems is that they often lack transparency. Indeed, while the black-box nature of these systems allows powerful predictions, they cannot be directly explained. The efficacy of these systems is thus limited by the machine's current inability to explain its own decisions and actions to human users.

In order to be able to trust system behavior, we must make machine learning explainable, either by using inherently interpretable models or by developing new methods for explaining and controlling otherwise overwhelmingly complex

fridaya@uef.fi

decisions by means of local approximation and interactive explanation. Our work aims to approach the problem of "explainability in artificial intelligence" and to illustrate the reason black box models are not acceptable for certain applications, e.g., in the medical domain where wrong decisions by the system can be harmful. The strategy for achieving this goal is to develop new or modified machine-learning techniques that will produce more explainable models. These models will be combined with state-of-the-art human-computer interface techniques capable of translating these models into understandable and useful explanation dialogues for an end user.

XAI aims to create a suite of machine learning techniques that produce more explainable models, while maintaining a high level of learning performance and thereby enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners. With the power of XAI or IML, new machine-learning systems will have the ability to explain their rationale, characterize their strengths and weaknesses, and thus provide an understanding of their future behaviour. We pursue a variety of techniques in order to generate a portfolio of methods that will provide future developers with a range of design options covering the performance-versus-interpretability trade space. Our work demonstrates that explainability of machine learning models is vital and highlights fundamental directions for research through three case studies of the use of explainability on ML classification tasks.

1.3 Contribution of Our Work

This paper consists of two main parts. Section 1 (above) is an introduction to the topic and demonstrates the significance of XAI. Section 2 is a critical review of an earlier articles on XAI that defined key concepts, specified some significant stages in a data-science life cycle where explainabillity is important, and identified algorithms that can be used in each of those phases. This survey provides an entry point for interested researchers and practitioners to learn key aspects of the young and rapidly growing body of research related to XAI. Through the lens of the literature, we review existing approaches to the topic, discuss trends surrounding its sphere, and present major research trajectories.

The second part is an experimental study demonstrating the implementation of an explainability framework on three domains: natural language processing, biomedical signal processing, and cybersecurity. Section 3 also covers detailed investigation of feature importance plots generated by explainer functions as well as providing analysis of misclassified cases, a step which would help detect and combat any adversarial attack attempting to disrupt the classifier learning.

The main goal of our work is to foster awareness of the necessity to understand the capabilities of machine learning and artificial intelligence. Our work also showcases the fact that interpretability is not only a powerful tool for detecting flaws in the model and biases in data, for verifying predictions, for improving models, but also in helping gain insights into the classification problem at hand.

fridaya@uef.fi

2 Literature Review of Explainable Artificial Intelligence

This section is divided into the following subsections: We first examine a sample of relevant and concepts related XAI that would help contextually to define the field (Sect. 2.1). Explainability is closely related to the concept of interpretability. Interpretable systems are explainable if their operations can be understood by humans. We note that even though "Explainable" is a keyword in the XAI appellation, in the ML community the term "interpretable" is preferred. In Sect. 2.2, we cover the conference/journal publications on XAI, and Sect. 2.3 covers the workshops and tutorials on the same area.

2.1 Terms and Definitions Used in Explainable Artificial Intelligence

- **Artificial intelligence (AI)**: AI is branch of computer science dealing with the simulation of intelligent behavior in computers and the capability of a machine to imitate intelligent human behavior [29].
- **Machine learning (ML)**: ML is an application of AI that provides systems with the ability to automatically learn and improve from experience without being explicitly programmed. The process of learning begins with observations of data (such as examples, direct experience, or instruction) to find patterns that allow for better future decisions based on the examples that we provide. The primary aim is to allow the computers to learn automatically without human intervention or assistance and adjust actions accordingly [30].
- **Supervised Machine Learning algorithms**: In Supervised algorithms, the learning algorithm produces an inferred function to make predictions about output values starting from the analysis of a known training dataset. The learning algorithm can also compare its output with the correct, intended output to find errors in order to modify the model accordingly, and can predict labels for any new input after sufficient training [30].
- **Unsupervised Machine Learning algorithms**: These algorithms analyze the way systems can infer a function to describe a hidden structure from unlabeled data. The system does not determine the right output but explores the data to allow inferences from datasets in order to describe hidden structures from unlabeled data [30].
- **Semi-supervised machine learning**: These algorithms fall somewhere in between designations of supervised and unsupervised learning, since they use both labeled and unlabeled data for training typically - a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy [30].
- **Reinforcement machine learning algorithms (RL)**: These learning methodologies interact with their environment by producing actions that discover errors or rewards. Their use is an approach to machine learning inspired by behaviorist psychology which contrasts with other ML approaches in that the algorithm is not explicitly told how to perform a task but works through the problem in its own. The RL method allows machines and software agents

fridaya@uef.fi

to automatically determine the ideal behavior within a specific context in order to maximize performance [30].
- **Deep Learning**: This method is an AI function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of machine learning that has networks capable of unsupervised learning from unstructured or unlabeled data [31].
- **Black Box**: A device, process, system, or program module that is designed to perform a particular task, whose inputs and outputs are known but in applying it a user has no knowledge of its internal structure and can only use it by passing a clearly defined set of input parameters [32].
- **Explainable artificial intelligence (XAI)**: This term, central in AI, refers to efforts to make sure that artificial intelligence programs are transparent in their purpose. It refers to the capability of understanding the work logic in ML algorithms. The idea behind explainable AI is that AI programs and technologies should not be strictly "black-box models" that people cannot understand. Thus, XAI is a key part of applying ethics to AI because it shows how machines are performing computational work and reduces imprint of unconscious biases and increase trust in model decisions [33].

2.2 Review of Prior Art (Conference/Journal Publications) on Explainable Artificial Intelligence

Model-based (Intrinsic) explainability and post hoc explainability are the main stages where explainability comes into play in machine learning model development. Explainability considerations often come into play in the modelling step related to the choice between simpler, easier-to-interpret models and more complex, black-box models which may fit the data better. Model-based (Intrinsic) explainability involves using a simpler model to fit data which can negatively affect predictive accuracy but yield higher descriptive accuracy. Post hoc explainability involves using methods to extract information from a trained model (with no effect on predictive accuracy) [34]. The model's ability to fit the data is measured through predictive accuracy. When selecting an interpretation method, it is not enough for the method to have high accuracy - the extracted information must also be relevant. The ability of the interpretations to properly describe what the model has learned is denoted by descriptive accuracy. The context provided by the problem and data stages guides what kinds of relationships a practitioner is interested in learning about and, by extension, the methods that should be used (Fig. 1).

The authors in [34] introduce the Predictive, Descriptive, Relevant (PDR) framework for evaluating and constructing interpretations: predictive accuracy, descriptive accuracy, and relevancy (Fig. 2). Model-based (Intrinsic) explainability can also be implemented in any of the following stages: "data before modeling", "building an explainable model", or "explainability analysis of established model". Because models are trained using data, efforts to explain the behavior of a model can start by analyzing the data used to train the model. In the "data before modeling" stage, a few representative characteristics can be selected from

fridaya@uef.fi

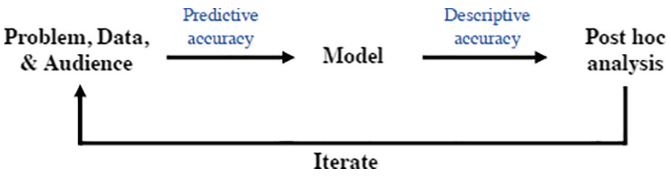


Fig. 1. Overview of different stages (black text) in a data-science life cycle where interpretability is important [34].

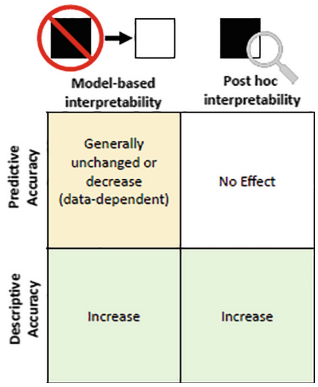


Fig. 2. Impact of interpretability methods on descriptive and predictive accuracies. Model-based interpretability involves using a simpler model to fit the data which can negatively affect predictive accuracy, but yields higher descriptive accuracy. Post hoc interpretability involves using methods to extract information from a trained model (with no effect on predictive accuracy). These correspond to the model and post hoc stages in Fig. 1 [34].

the data to build the model during training in order to explain the input and output of the model.

In the "Building an explainable model" stage, an explainability method is used to supplement the AI structure by combining it with traditional ML. This combination can balance the effectiveness of the learning result with the explainability of the learning model and provide a framework for explainable learning. A common important theoretical foundation of traditional ML methods is that of statistics, an approach widely used to provide sound explainability in such computer fields as natural language processing, voice recognition, image recognition, information retrieval, and biological information recognition.

In the "Explainability analysis of established models" stage, an approach is developed that seeks to analyze the dependencies between the input, output, and intermediate information of AI models to analyze and verify the models' logic. In the literature, there are both general model analysis methods applicable to multiple models, such as Local Interpretable Model-Agnostic Explanations (LIME) [18], and specific model analysis methods that can analyze the construction of a

fridaya@uef.fi

specific model in depth. LIME is a robust framework as it is not model specific and can be applied to a wide range of ML and DL algorithms.

We describe recent work on improving the interpretability/explainability of machine learning and deep learning models in Table 2. We classify each work according to the dimensions of interpretability outlined as intrinsic or post hoc explainability approaches. Note that this review is not fully comprehensive but represents a survey of methods and results that we consider particularly pertinent to our future research goals in machine learning and deep learning interpretability.

2.3 Related Workshops/Tutorials on Explainable Artificial Intelligence

The current generation of AI systems offer substantial benefits, but their effectiveness will be limited by the machine's inability to explain its decisions and actions to users. DL is one of the most successful approaches in machine learning but has been disapproved for not being able to provide answers that cannot be completely trusted and pointing to model vulnerabilities in various applications such as language, natural language process, and vision models. Spoofability and biasedness have been demonstrated for face recognition, visual recognition, and natural language processing but no robust solution has so far been found. The European Union introduced a right-to-explanation in GDPR as an attempt to remedy the potential problems given the rising importance of ML algorithms.

Explainability is the first step in achieving a transparency goal which can be achieved by checking systems explanations to determine whether they satisfy desirable trust criteria. Because machine learning models are opaque, non-intuitive, and thus difficult for people to understand, an explainable AI will be essential if users are to understand, appropriately trust, and effectively manage this incoming generation of artificially intelligent partners. DARPA and LIME were the first two most notable efforts to create explainable AI. The US Department of Defense's Defense Advanced Research Projects Agency (DARPA) launched the Explainable Artificial Intelligence (XAI) project to develop a software library toolkit for explainable AI. Researchers demonstrated initial implementation of their explainable AI systems. A complete evaluation of those systems is presented in [33].

There is a rapidly growing literature on techniques that can be used to interpret more complex machine learning models such as neural networks. A tool, **LIME** (Local Interpretable Model-Agnostic Explanations), created by Marco Ribeiro and his colleagues, can highlight the features in an image or a text responsible for an ML model's predictions for either that image or the text. It attempts to determine the most salient features of the model or the key factors driving any one decision by feeding through the model inputs similar to the original ones and observing how the predictions change. This method has the benefit of yielding simple explanations, such as whether a particular word in a document or a shape in a photo is driving the model's predictions. Other methods of explaining complex models, such as **DeepLIFT** (Deep Learning

fridaya@uef.fi

Table 2. Summary of explainability techniques

Parent technique	Sub category details of technique	Specific to/Advantages	Examples/References
Integrated interpretability methods (Intrinsic)	User based techniques	Cater to models with lower complexity inexperienced user	-Linear model [35] -decision tree [35] -decision table [35]
	Transparent approaches	Specific to transparent model familes with low complexity	-evolutionary programming [36] -interpretable decision sets [37] -tree sparse additive models [38] -prototype selection [39]
	Hybrid approaches	Combine transparent with black box model	-Logistic Regression+ SVM [40] -model trees + black box [40]
Post-hoc Methods	Transparent proxy techniques	Use of rules/advanced training schemes	-model specific DT used to learn single DT from ensemble DT [41] -makes use of rules extracted from SVM [42] -rules extracted from ensemble NN [43] -Bayesian regression mixture on DNN, SVM, RF [44] -Adversarial Training scheme on DNN [45]
	Indictive approaches	Use of visualization, conceptual representation	-visualization techniques on high dimensional data [46], RNN [47], NN via unit salience [48], deconvolutional network for CNN [49]
	Sensitivity analysis	Use of barplots, variable effect charts, feature importance, partial dependency plots, individual conditional expectation, local approximation methods	-Pedagogical visualization using instance level explanation [50] -layerwise relevance propagation [51] -taylor decomposition on DNN [52] -Reinforcement learning based discriminative loss function for DNN explanation [53] -LIME [54], LOCO [55] -Shaply additive explanations(SHAP) [56]
	Counterfactual explanations	Use of unconditional counterfactual explanations	-minimum conditions that led to alternative decision for NN [57]
	Prototypes and criticism	Use of prototypes and critics to characterize dataset	-unsupervised algorithm to find prototypes (MMD-critic) [58]
	Model distillation approach	Transfer knowledge from deep networks to shallow networks	-Interpretable Mimic Learning [59] -knowledge distillation method (DARK SIGHT) [60]

Important FeaTures) [61], peer into the inner workings of a neural network to extract those parameters driving the model's output. **SHAP** (SHapley Additive exPlanations) [56], a recent addition to this set of methods, attempts to unify these prior attempts at interpreting model output.

fridaya@uef.fi

XAI is being used now in certain limited circumstances, and prototypes are emerging from applied research which can report what is in a photo or describe why an autonomous system did one thing rather than something else. Annual international conference workshops and tutorial series have been dedicated exclusively to the topic of explainable artificial intelligence such as Fairness, Accountability, and Transparency (FAT-ML) workshop at KDD 2014–2018 [62]; ICML Workshop on Human Interpretability in Machine Learning (WHI) 2016–2018 [63]; XCI 2017 on Explainable Computational Intelligence [64]; CD-MAKE 2018 Workshop on Explainable Artificial Intelligence [65]; ICAPS 2018 Workshop on Explainable AI Planning [66]; HRI 2018 Workshop on Explainable Robotic Systems [67]; ACM Intelligent User Interfaces (IUI) 2018 workshop on Explainable Smart Systems (EXSS 2018) [68]; IPMU 2018 on Advances on Explainable Artificial Intelligence [69]; and ICCBR 2018 XCBR: First Workshop On Case-Based Reasoning For The Explanation Of Intelligent Systems [70].

The subject has also become a key panel discussion topic in specific sessions at major conferences such as at the Deep Learning and Security workshop at 40th IEEE Symposium on Security and Privacy 2019 [71] and the Workshop on Explainable Smart Systems (EXSS) [68]. The year 2019 will also witness a wide range of dedicated workshops on explainablity such as a special session on Advances in explainable Artificial Intelligence at FUZZ-IEEE 2019 [72], CD-MAKE 19 [65], ML Prague 2019 [73], a workshop on XAI at AppliedML-Days [74], and ACM IUI 2019 [75] where the Human-Computer Interaction (HCI) community meets the AI community.

Although several workshops have previously addressed the problem of explainable AI, none has had a focus on semantic technologies such as ontologies and reasoning. The SEMEX workshop [76] aims to bring together international experts interested in the application of semantic technologies for explainability of artificial intelligence to stimulate research, engineering, and evaluation, all steps towards making machine decisions transparent, re-traceable, comprehensible, interpretable, explainable, and reproducible. Semantic technologies have the potential to play an important role in the field of explainability since they lend themselves quite well to the task by enabling model users conceptualizations of a problem domain.

In view of recent advancements, these workshops will not only bring together international cross-domain experts interested in AI but will also stimulate research, engineering and evaluation in and for explainable AI -towards making machine decisions transparent, comprehensible, interpretable, and thus explainable.

3 Explainable Artificial Intelligence (XAI) Applications

XAI can bring a significant benefit to a large range of Domains such as Biomedical, HealthCare, Legal, Finance, Transportation, Malware, Robots, Cities, and NLP. In our analysis, we make use of a model agnostic LIME interpretability framework on three classification tasks. We introduce the LIME Framework in Sect. 3.1 followed by case studies in Sect. 3.2.

fridaya@uef.fi

3.1 Details of LIME Framework

LIME Framework, short for Local Interpretable Model-Agnostic Explanations, refers to local fidelity, i.e., the explanation must reflect the behavior of the classifier "around" the instance being predicted [54]. LIME is able to explain any model without needing to "peek" into it, making it model-agnostic. Many classifiers use representations that are not intuitive to users at all (e.g., word embedding for text classification). LIME explains such classifiers in terms of interpretable representations (words), even if that is not the representation used by the classifier.

The steps in LIME Framework are as follows:

- Permute data
- Calculate distance/similarity-score metric between original and permutated observations
- Make predictions on new data using complex model
- Pick minimum number of features/predictors that give the maximum likelihood of class as predicted by the black box
- Fit a simple model to the permuted data with m features and similarity scores as weights

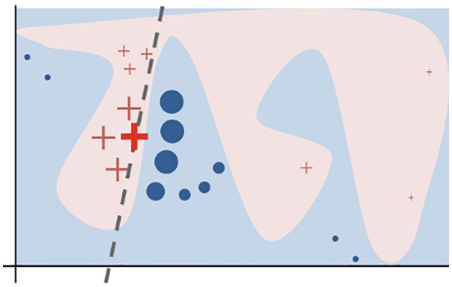


Fig. 3. Illustration of explanation process for LIME Framework. Original non-linear model's decision function represented by blue/pink background. Bright red cross is the instance being explained (X). Dashed line is linear model [54]

Weights from the model make explanations for the complex model's local behavior (Fig. 3). The original model's decision function is nonlinear. We sample perturbed instances around X, and weight them according to their proximity to X (weight is represented by size). We get original model's prediction on these perturbed instances, and then learn a linear model that approximates the model well in the vicinity of X. Note that although the explanation in this case is not faithful globally, it is faithful locally around X.

fridaya@uef.fi

For our experimental analysis, we implemented LIME interpretability framework on the following case studies:

- First: natural language processing application for tweet data classification
- Second: biomedical tumor data classification
- Third: Windows PC malware classification

3.2 Implementation Details and Results of Case Studies

(A) Explainability in Natural Language Processing Domain
Previous work on interpretability for NLP was primarily focused on text classification problems. These include studies by Ribeiro et al. [77] in which they try to identify an interpretable model over an interpretable representation locally faithful to the classifier. Samek et al. [51] directly build connections between the inputs and the outputs, using heatmap analysis to visualize how much each hidden element contributes to predicted results. Liu [78] builds a generative explanation framework for text classification, proposing an explanation factor and introducing minimum risk training for this hybrid generative-discriminative framework.

Although these systems are promising in some ways, they typically do not consider the fine-grained information (e.g., textual explanations for the labels) that may contain sufficient hints for interpreting behavior of models. Unfortunately, existing text classification models typically utilize the original review texts to predict overall results directly [79–81], but none provides fine-grained explanations as to why the decisions are made. From human perspective, we could hardly understand or trust such models that provide almost scant reasons for decision making. Therefore, it is crucial to build an explainable text-classification model capable of explicitly generating fine-grained information for predictions.

To achieve those goals, we implement an example of generative explanation framework for a tweets classification problem. Our implemented model, based on [82], is capable not only of providing classification predictions but also of generating fine-grained information as explanations for decisions. For this tweet classification task, we used the data from Kaggle which had close to 3,000 of the most recent tweets and associated metadata from Hillary Clinton and Donald Trump [83].

For data preprocessing and feature modeling, we tokenized the data and transformed it to Document-Term Matrices (DTM) for word-level tokens. The Extreme Gradient Boosting tree classifier performed the best in terms of classification accuracy (84%) using standard parameters. To understand the clues that the classifier picked, we further implemented LIME Framework to observe the explainer function on the correctly predicted instances while ignoring misclassified observations.

The result of the first first observations for interpretation of tweet classification is illustrated in Fig. 4. The features plot developed using LIME framework provides us with a better understanding of explainability of the classifier model

fridaya@uef.fi

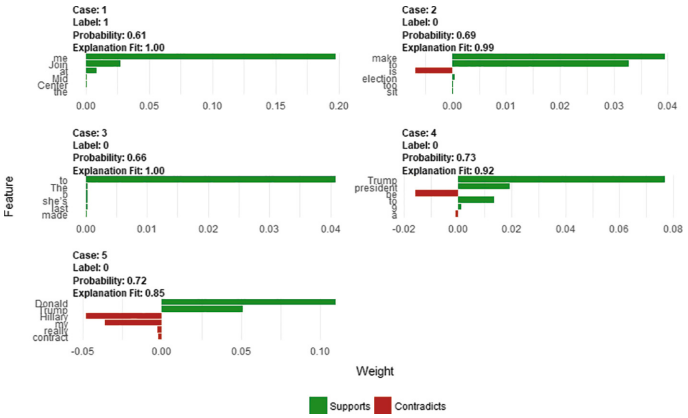


Fig. 4. Result of the first five observations for interpretation of tweet classification

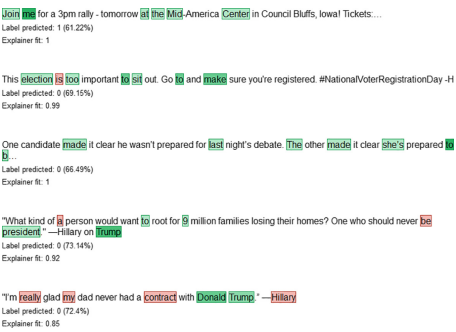


Fig. 5. Results of the text explanations which indicate the direction but not strength of the relationship of support or contradiction of certain words for a given label

by indicating which features matter. LIME also provides a clearer view of support and a contradiction of certain words for a given class label using LIME explainer function as demonstrated in Fig. 5. Overall, LIME interpretability allows us substantial improvement in understanding how the "black box" model works and why it allows us to make an educated guess as to which features matter.

(B) Explainability in the Health Care Domain

As ML and DL has become an active research area in health care, the increasingly widespread applicability of these models necessitates the need for explanations to hold such models accountable. Specifically, in health care a critical application such as the prediction of a patient's end of life may have much more stringent conditions for explanation fidelity as compared to merely predicting costs of a procedure. Health care offers unique challenges in which the demands for explainability, model fidelity, and performance are generally much higher as compared to most other domains.

fridaya@uef.fi

In the medical domain there have only been a few workshops specifically focused on interpretability/explainability, such as [84]. Previous work in interpretability includes the work of Lee in [85], in which the authors present an understandable deep-learning system that detects acute intracranial haemorrhage (ICH) and can classify five ICH subtypes from unenhanced head computed-tomography scans. The author in [85] reviews the notion of interpretability within the context of health care, the various nuances associated with it, the challenges related to interpretability unique to health care, and the future of interpretability in this domain. These useful findings, however, miss generating fine-grained information for predictions that makes prediction results more understandable and the classifier more explainable.

In order to remedy this deficiency, we implement the LIME Framework for the task of Automated Classification between Benign and Malignant Tumor Class [82]. For this task, we implemented random forest (RF) method, gradient boosting machines (GBM), generalized linear models (GLM), and deep learning (DL) models on cancer datasets from R library to produce a leaderboard based on the best stopping metric. Deep Neural network achieved the best classification results with accuracy of 94%, recall of 86.84%, and precision of 97.05%. We made use of LIME to explain predictions of deep learning models. For evaluation, we split the data into correct and wrong predictions to better comprehend what may have confused the model regarding misclassified observations, and also plotted feature importance plots (Fig. 6).

From the results of LIME Framework, we can observe that the smaller and more regular cells with low values of bare nuclei (`bare_nuclei`) correctly indicate benign cells, whereas big, irregular cells with higher values of clump

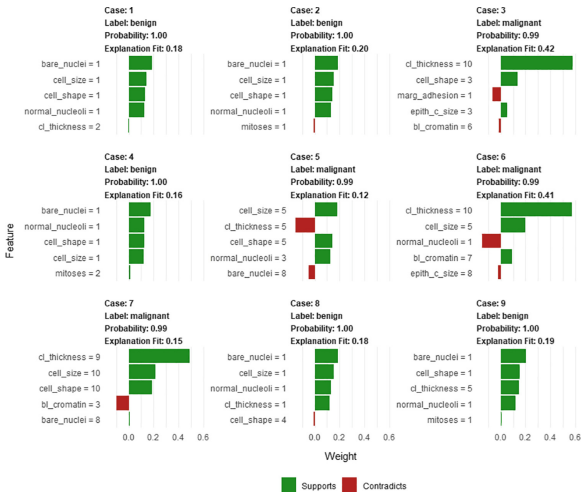


Fig. 6. Illustration of results from LIME explainer function for the correctly predicted labels

fridaya@uef.fi

thickness (`cl_thickness`) support malignant label, an observation that makes prediction results understandable.

Analysis of misclassifications for automated tumor classification is presented in Fig. 7. Looking retrospectively at the misclassification (i.e., cases predicted to be benign while actually malignant), LIME Framework indicates that they consisted of mainly small and quite regular cells, although some malignant characteristics were still present (as observed by higher values of the "bare nuclei" and "clump thickness" feature). The true power of the lime package is established when we look at misclassified labels and try to understand what made the model misclassify them.

(C) Explainability in a Malware Detection/Security Domain
Machine-learning models have been reporting to demonstrate impressive performance for detecting malicious Android applications, intrusions, and malware in security domains. Recent findings on adversarial machine learning have showcased the fact that very few changes to the content of Android malware may suffice to evade detection. This lack of explainability leaves openings that can be exploited by adversarial machine learning methods such as evasion, poisoning, and backdoor attacks. Thus, in order to ensure that a malware detector that is performing well on benchmark data will continue to do so when deployed in an operating environment, researchers recently have made use of linear, explainable machine-learning models to easily identify the most influential features contributing to each decision.

Inspired by DARPA's XAI program, Luca Vigano and his group proposed a new paradigm in security research called Explainable Security (XSec) in [86]. They discuss the "Six Ws" of XSec (Who? What? Where? When? Why? and How?) arguing that XSec has unique and complex characteristics in as much as it involves several different stakeholders (i.e., the system's developers, analysts,

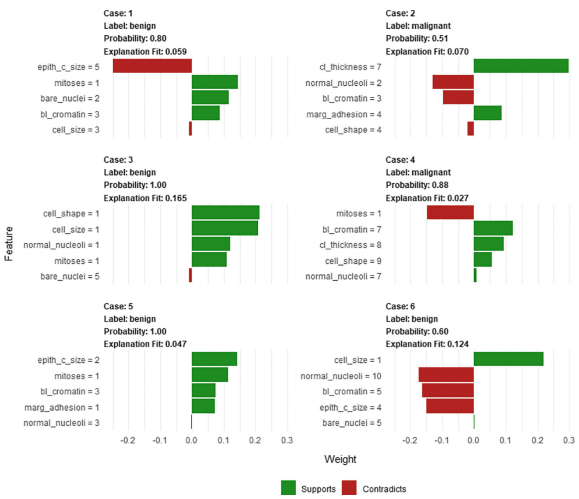


Fig. 7. Illustration of results from LIME explainer function for the misclassified labels

fridaya@uef.fi

users, and attackers) and is multi-faceted by nature (requiring reasoning about system model, threat model, properties of security, privacy, and trust as well as about con. The paper proposes that using the growing amount of results on XAI together with the pioneering research on explanations in security and trust as stepping stones, will help to "bridge the gap" between various research questions and communities, as may be seen in the work of Hoffman on planning for penetration testing [86]. This year will witness a session on Explainable Deep Learning at the workshop "Deep Learning and Security" organized by IEEE Symposium of Security and Privacy [71].

Other previous work on interpretability in the security domain includes the work by Marco Melis on explaining malicious black-box android applications [87]. This work leverages a gradient-based approach to identify the most influential local features. It also enables use of nonlinear models to potentially increase accuracy without sacrificing interpretability of decisions on the DREBIN Dataset. Drebin as such explains its decisions by reporting, for any given application, the most influential features, i.e., those features present in a given application and assigned the highest absolute weights by the classifier. It is easy to see, from its most influential features, that a malware sample is correctly identified as it connects to a suspicious URL and uses SMS as a side channel for communication. This work by Marco on interpretability extends this approach to nonlinear models. For each detected application the respective patterns can be extracted, mapped to meaningful descriptions, and then provided to the user as explanation for the detection.

In the domain of intrusion detection, an adversarial learning-based approach was proposed for generating explanations for misclassified samples in [88]. The authors here presented an adversarial approach to find the minimum modifications needed in order to correctly classify the misclassified samples. These modifications are then used to find and visualize the relevant features responsible for the misclassification on an Intrusion Detection benchmark dataset. The explanations were presented using intuitive plots that can be easily interpreted by a user, thus providing insightful and satisfactory explanations for the misclassification of samples with results that match expert knowledge. Other works on explainable android malware detection do not specifically use explainable machine learning models but do make use of other feature analyses to reduce uncertainty of information. For example, paper [89] uses static analysis and probability statistics-based feature extraction analysis to detect and analyze malicious Android apps.

To the best of our knowledge, explainability has not been introduced for a Windows PC malware classification. We would be the first to present the explainability framework on the Windows PC malware classification. Modern malware is designed with mutation characteristics (namely polymorphism and metamorphism) which cause an enormous growth in the number of variants of malware samples. For this analysis, we used the data released by Microsoft (as part of Microsoft Malware Classification Challenge 2015) which has a huge dataset of nearly 0.5 terabytes of data, containing more than 20 K malware samples. We mainly focused our classification task on distinguishing between the Ramnit

fridaya@uef.fi

Malware Family and the Lollipop Malware Family. For each malware sample, we compute not only a set of content-based features by relying on state-of-the-art mechanisms, but also propose the extraction of powerful complementary statistical features that reflect the structure of portable executable (PE) files inspired by [90]. The features extracted include both those from the hex view and the assembly view in order to exploit complementary information brought in by these two representations.

Hex dump-based features include N-gram, size, address of first-byte sequence as metadata features, entropy, ASCII strings, local binary patterns, and haralick features from the visualization of the byte code. Disassembled file features include operation codes that are based either on their commonness, the frequency of use of the registers (REG), frequency of use of Windows Application Programming Interfaces (API), section-based features, and some discriminative sets of features specific for certain malware samples not containing any API call.

To better exploit both the richness of the available information in the number of the malware samples for training the classifier and the number of features used to represent the samples, we resorted to the stacked ensemble classification technique. We evaluated our system on two malware families from data provided by Microsoft for their malware challenge [91] hosted at Kaggle. The algorithm achieved an impressive accuracy of 99.83%, recall of 99.57%, and misclassification rate of 0.165%. The explainer function in the LIME Framework provided the feature importance plots, different for each malware sample (Fig. 8).

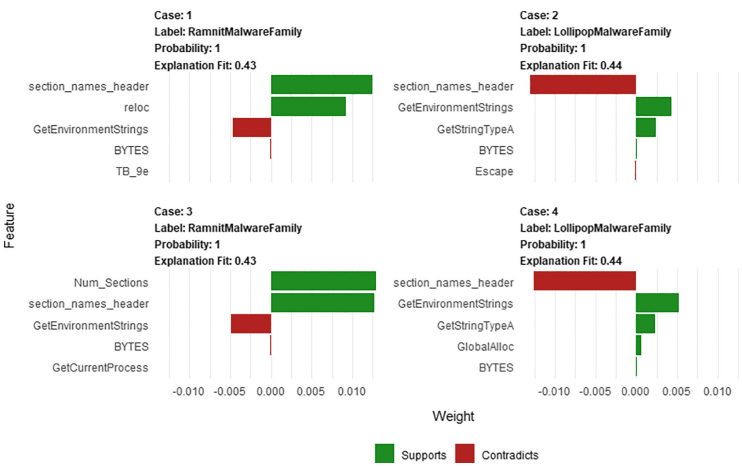


Fig. 8. Illustration of results from LIME explainer function for the correctly predicted labels

This result of explainer function indicate that high value of `section_name_header, num_sections` correctly indicate Class "Ramnit Malware", whereas "environment string", "string type features" `(getstringtypeA)` and `global`

fridaya@uef.fi

`alloc` support "Lollipop Malware" Class. A PE consists of some predefined sections like.text, .data, bss, idata, .rsrc, tls, and .reloc. The features `section_names_header`, `:num_sections` are for different characteristics from the sections (SEC) feature. The true power of LIME package is reiterated by understanding what made the model missclassify labels (Fig. 9).

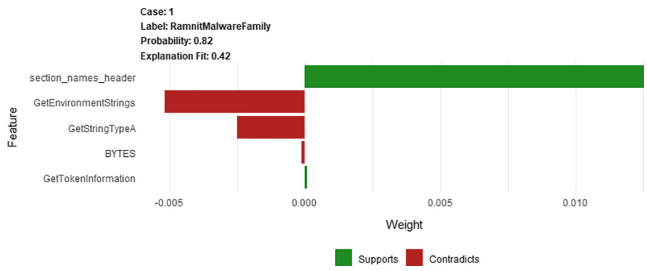


Fig. 9. Illustration of results from LIME explainer function for the misclassified labels

All misclassified cases were predicted to belong to "Ramnit Malware", while they were actually of "Lollipop Malware" Class. The explainer function reveals that these malware samples indicated particularly high value for "section-based features", a representative characteristic of "Ramnit Malaware". It shows quite low values for the features (i.e., environment string, string typefeatures (getstringtypeA)) which is contradictory to the characteristics of the "Lollipop Malware" Class.

These findings enhance our understanding of how the "black box" model works and why it makes mistakes. Even though they do not produce "fixed" feature importance plots (i.e., a general rather than a case-to-case view of which variables are most informative when making a prediction), they allow one an educated guess as to which features matter. The analysis of this dataset inspired the development of a novel paradigm that is not only effective in categorizing malware variants into their actual family groups but that is also able to explain why they were classified that way.

Our approach not only unpacks coverage for predictive features while identifying flaws in feature engineering, but highlights those global characteristics learned by the model in order to discriminate between the two classes. Finally, as shown by our empirical analysis on a popular Windows PC malware detection task, our method also helps identify potential vulnerabilities of both linear and nonlinear models from adversarial manipulation. The benefits of our proposed explainability framework are three-fold. We can identify illegitimate intrusion, malware, or other attack data, and even fix or delete adversarial examples based on the self-consistency of data in order to improve model robustness. Explainability in general also helps to identify bias in raw data and strategize the model optimization (Fig. 10). Additionally, the proposed explainability framework opens a new window for future research, showcasing specific explainability methods that provide an intriguing potential for accurate malware classification.

fridaya@uef.fi

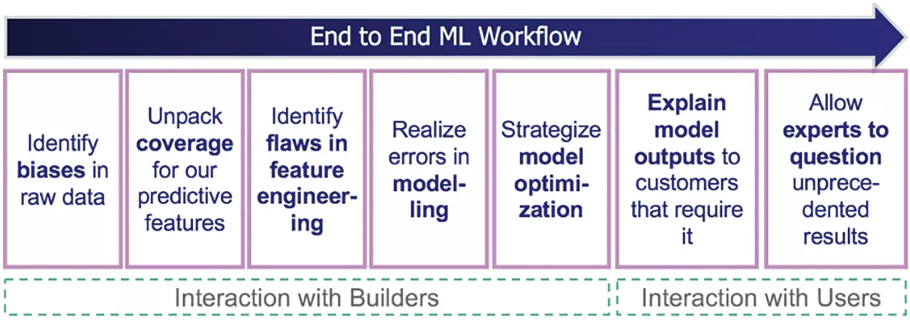


Fig. 10. Benefits provided by Interpretability throughout a data science workflow [92]

4 Conclusion

Black-box models should not be directly used in critical applications to domains such as medical or malware detection as wrong decisions of the system can result in harmful consequences. An explainability approach addresses this problem by developing a suite of deep learning and advanced machine learning techniques that produce more explainable models, while maintaining a high level of prediction performance. The explainability paradigm is presented and discussed in the present paper, where emphasis has been given to the phases related to explainability in the machine learning modelling process and in the selection of a set of novel algorithms that can be used for the effective representation and interpretability of misclassified samples.

We applied explainability framework to three domains: natural language processing domain for classifying tweets, signal processing domain for tumor classification, and cybersecurity domain to specifically detect and classify Windows PC malware. We find that an explainability framework as presented here furnishes a useful tool for identifying both fine-grained analysis at the feature level for malware and tumor data classification and as a text dataset introspection across documents and words that are important to a classifier's decision. This knowledge enables us to broaden the scope of applications of standard machine learning classifiers such as support vector machines or deep neural networks, by extending the primary classification result with additional information linking the classifier's decision back to input components.

Furthermore, based on LIME relevance, we expect this technique to also be suitable for types of applications based on other neural network architectures such as character-based or recurrent network classifiers, or on other types of classification problems (e.g., sentiment analysis). More generally, interpretability could contribute to the design of more accurate and efficient classifiers, not only by inspecting and leveraging the input space's relevance, but also through the analysis of intermediate relevance values "at classifier hidden layers". The present results have yielded encouraging results from our efforts in explanation of deep network processing, deep network representation, and system-level explanation production.

fridaya@uef.fi

References

1. 59 impressive things Artificial Intelligence can do today. https://www.businessinsider.com/artificial-intelligence-ai-most-impressive-achievements-2017-3. Accessed 7 Mar 2017
2. Mnih, V., et al.: Human-level control through deep reinforcement learning. Nature **518**(7540), 529 (2015)
3. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
4. 6-Amazing-AI-Achievements. https://themerkle.com/6-amazing-ai-achievements-that-will-blow-your-mind//. Accessed 24 Feb 2017
5. Silver, D., et al.: Mastering the game of Go with deep neural networks and tree search. Nature **529**(7587), 484 (2016)
6. AI Transforming the world. https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world//. Accessed 24 Apr 2018
7. IBM Watson Report. https://www.wsj.com/articles/ibm-bet-billions-that-watson-could-improve-cancer-treatment-it-hasnt-worked-1533961147. Accessed 24 Apr 2018
8. Sun, L., Jia, K., Yeung, D.-Y., Shi, B.E.: Human action recognition using factorized spatio-temporal convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4597–4605 (2015)
9. Yue-Hei Ng, J., Hausknecht, M., Vijayanarasimhan, S., Vinyals, O., Monga, R., Toderici, G.: Beyond short snippets: deep networks for video classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4694–4702 (2015)
10. Geng, W., Du, Y., Jin, W., Wei, W., Hu, Y., Li, J.: Gesture recognition by instantaneous surface EMG images. Sci. Rep. **6**, 36571 (2016)
11. Fraiwan, L., Lweesy, K.: Neonatal sleep state identification using deep learning autoencoders. In: 2017 IEEE 13th International Colloquium on Signal Processing & its Applications (CSPA). IEEE, pp. 228–231 (2017)
12. Schirrmeister, R.T., et al.: Deep learning with convolutional neural networks for EEG decoding and visualization. Hum. Brain Mapp. **38**(11), 5391–5420 (2017)
13. Spampinato, C., Palazzo, S., Kavasidis, I., Giordano, D., Souly, N., Shah, M.: Deep learning human mind for automated visual classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6809–6817 (2017)
14. Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adeli, H.: Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput. Biol. Med. **100**, 270–278 (2018)
15. Zheng, W.-L., Zhu, J.-Y., Peng, Y., Lu, B.-L.: EEG-based emotion classification using deep belief networks. In: 2014 IEEE International Conference on Multimedia and Expo (ICME). IEEE, pp. 1–6 (2014)
16. Luo, K., Li, J., Wang, Z., Cuschieri, A.: Patient-specific deep architectural model for ECG classification. J. Healthc. Eng., **2017** (2017)
17. Acharya, U.R., et al.: A deep convolutional neural network model to classify heartbeats. Comput. Biol. Med. **89**, pp. 389–396 (2017)
18. Muduli, P.R., Gunukula, R.R., Mukherjee, A.: A deep learning approach to fetal-ECG signal reconstruction. In: 2016 Twenty Second National Conference on Communication (NCC). IEEE, pp. 1–6 (2016)

fridaya@uef.fi

19. Acharya, U.R., Fujita, H., Lih, O.S., Hagiwara, Y., Tan, J.H., Adam, M.: Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Inf. Sci. **405**, 81–90 (2017)
20. Zheng, G., Ji, S., Dai, M., Sun, Y.: ECG based identification by deep learning. In: Chinese Conference on Biometric Recognition. Springer, pp. 503–510 (2017)
21. Ye, Y., Chen, L., Hou, S., Hardy, W., Li, X.: DeepAM: a heterogeneous deep learning framework for intelligent malware detection. Knowl. Inf. Syst. **54**(2), 265–285 (2018)
22. Yuan, Z., Lu, Y., Xue, Y.: Droiddetector: android malware characterization and detection using deep learning. Tsinghua Sci. Technol. **21**(1), 114–123 (2016)
23. Yin, C., Zhu, Y., Fei, J., He, X.: A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access **5**, pp. 954–961 (2017)
24. Shone, N., Ngoc, T.N., Phai, V.D., Shi, Q.: A deep learning approach to network intrusion detection. IEEE Trans. Emerg. Comput. Intell. **2**(1), 41–50 (2018)
25. Muna, A.-H., Moustafa, N., Sitnikova, E.: Identification of malicious activities in industrial internet of things based on deep learning models. Inf. Secur. Appl **41**, 1–11 (2018)
26. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**(1), 50–80 (2004)
27. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., Mané, D.: Concrete problems in AI safety. arXiv preprint arXiv:1606.06565 (2016)
28. Varshney, K.R., Alemzadeh, H.: On the safety of machine learning: cyber-physical systems, decision sciences, and data products. Big data **5**(3), 246–255 (2017)
29. Merriam-webster. https://www.merriam-webster.com/dictionary/artificial%20intelligence//. Accessed 27 Feb 2019
30. Expertsystems. https://www.expertsystem.com/machine-learning-definition/
31. Investopedia. https://www.investopedia.com/terms/d/deep-learning.asp. Accessed 9 Mar 2018
32. Software testing. https://www.softwaretestinghelp.com/black-box-testing//. https://www.softwaretestinghelp.com/black-box-testing//. Accessed 31 Dec 2018
33. DARPA. https://www.darpa.mil/program/explainable-artificial-intelligence
34. Murdoch, W.J., Singh, C., Kumbier, K., Abbasi-Asl, R., Yu, B.: Interpretable machine learning: definitions, methods, and applications. arXiv preprint arXiv:1901.04592 (2019)
35. Huysmans, J., Dejaeger, K., Mues, C., Vanthienen, J., Baesens, B.: An empirical evaluation of the comprehensibility of decision table, tree and rule based predictive models. Decis. Support Syst. **51**(1), 141–154 (2011)
36. Cano, A., Zafra, A., Ventura, S.: An interpretable classification rule mining algorithm. Inf. Sci. **240**, 1–20 (2013)
37. Lakkaraju, H., Bach, S.H., Leskovec, J.: Interpretable decision sets: a joint framework for description and prediction. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 1675–1684 (2016)
38. Wang, J., Fujimaki, R., Motohashi, Y.: Trading interpretability for accuracy: oblique treed sparse additive models. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 1245–1254
39. Bien, J., Tibshirani, R., et al.: Prototype selection for interpretable classification. Ann. Appl. Stat. **5**(4), 2403–2424 (2011)

fridaya@uef.fi

40. Van Gestel, T., Baesens, B., Van Dijcke, P., Suykens, J., Garcia, J., Alderweireld, T.: Linear and nonlinear credit scoring by combining logistic regression and support vector machines. J. Credit Risk **1**(4) (2005)
41. Van Assche, A., Blockeel, H.: Seeing the forest through the trees: learning a comprehensible model from an ensemble. In: European Conference on Machine Learning. Springer, pp. 418–429 (2007)
42. Martens, D., Baesens, B., Van Gestel, T., Vanthienen, J.: Comprehensible credit scoring models using rule extraction from support vector machines. Eur. J. Oper. Res. **183**(3), 1466–1476 (2007)
43. Zhou, Z.-H., Jiang, Y., Chen, S.-F.: Extracting symbolic rules from trained neural network ensembles. AI Commun. **16**(1), 3–15 (2003)
44. Guo, W., Zhang, K., Lin, L., Huang, S., Xing, X.: Towards interrogating discriminative machine learning models. arXiv preprint arXiv:1705.08564 (2017)
45. Dong, Y., Su, H., Zhu, J., Bao, F.: Towards interpretable deep neural networks by leveraging adversarial examples. arXiv preprint arXiv:1708.05493 (2017)
46. Maaten, L.v.d., Hinton, G.: Visualizing data using t-SNE. J. Mach. Learn. Res., pp. 2579–2605 (2008)
47. Karpathy, A., Johnson, J., Fei-Fei, L.: Visualizing and understanding recurrent networks. arXiv preprint arXiv:1506.02078 (2015)
48. Li, J., Chen, X., Hovy, E., Jurafsky, D.: Visualizing and understanding neural models in NLP. arXiv preprint arXiv:1506.01066 (2015)
49. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: European conference on Computer Vision. Springer, pp. 818–833 (2014)
50. Tamagnini, P., Krause, J., Dasgupta, A., Bertini, E.: Interpreting black-box classifiers using instance-level visual explanations. In: Proceedings of the 2nd Workshop on Human-In-the-Loop Data Analytics, p. 6. ACM (2017)
51. Samek, W., Wiegand, T., Müller, K.-R.: Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models. arXiv preprint arXiv:1708.08296 (2017)
52. Montavon, G., Lapuschkin, S., Binder, A., Samek, W., Müller, K.-R.: Explaining nonlinear classification decisions with deep taylor decomposition. Pattern Recognit. **65**, 211–222 (2017)
53. Hendricks, L.A., Akata, Z., Rohrbach, M., Donahue, J., Schiele, B., Darrell, T.: Generating visual explanations. In: European Conference on Computer Vision. Springer, pp. 3–19 (2016)
54. Ribeiro, M.T., Singh, S., Guestrin, C.: Why should i trust you?: explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Uiscovery and Data Mining, pp. 1135–1144. ACM (2016)
55. Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R.J., Wasserman, L.: Distribution-free predictive inference for regression. J. Am. Stat. Assoc. **113**(523), 1094–1111 (2018)
56. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, pp. 4765–4774 (2017)
57. Wachter, S., Mittelstadt, B., Russell, C.: Counterfactual explanations without opening the black box: automated decisions and the GDPR. Harv. J. Law Technol. **31**(2), 2018 (2017)
58. Kim, B., Khanna, R., Koyejo, O.O.: Examples are not enough, learn to criticize! criticism for interpretability. In: Advances in Neural Information Processing Systems, pp. 2280–2288 (2016)

fridaya@uef.fi

59. Che, Z, Purushotham, S., Khemani, R., Liu, Y.: Distilling knowledge from deep networks with applications to healthcare domain. arXiv preprint arXiv:1512.03542 (2015)
60. Xu, K., Park, D.H., Yi, C., Sutton, C.: Interpreting deep classifier by visual distillation of dark knowledge. arXiv preprint arXiv:1803.04042 (2018)
61. Shrikumar, A., Greenside, P., Kundaje, A.: Learning important features through propagating activation differences. In: Proceedings of the 34th International Conference on Machine Learning. JMLR. org, vol. 70, pp. 3145–3153 (2017)
62. Fairness, accountability, and transparency in machine learning. http://www.fatml.org/
63. ICML - WHI. https://sites.google.com/view/whi2018/home
64. XCI -2017. http://xci2017.arg.tech/
65. CD MAKE. https://cd-make.net/
66. ICAPS 2018 International Conference on Automated Planning and Scheduling. http://icaps18.icaps-conference.org/xaip/
67. Human Robot Interaction- Workshop on Explainable Robotic Systems. http://humanrobotinteraction.org/2018/
68. EXSS workshop on explainable smart systems. http://explainablesystems.comp.nus.edu.sg/
69. IPMU advances on explainable artificial intelligence. https://sites.google.com/view/xai-ipmu2018
70. ICCBR 2018 XCBR: First Workshop on Case-Based Reasoning for the Explanation of Intelligent Systems. http://gaia.fdi.ucm.es/events/xcbr/
71. DEEP LEARNING ANDSECURITY WORKSHOP. https://www.ieee-security.org/tc/spw2019/dls/
72. FUZZ IEEE International conference on fuzzy systems. http://sites.ieee.org/fuzzieee-2019/
73. ML - Prague. https://www.mlprague.com/
74. Applied ML Days. https://www.appliedmldays.org/
75. ACM IUI. https://iui.acm.org/2019/
76. Workshop on semantic explainability. http://www.semantic-explainability.com/
77. Ribeiro, M.T., Singh, S., Guestrin, C.: Model-agnostic interpretability of machine learning. arXiv preprint arXiv:1606.05386 (2016)
78. Liu, H., Yin, Q., Wang, W.Y.: Towards explainable NLP: a generative explanation framework for text classification. arXiv preprint arXiv:1811.00196 (2018)
79. Xu, W., Sun, H., Deng, C., Tan, Y.: Variational autoencoder for semi-supervised text classification. In: Thirty-First AAAI Conference on Artificial Intelligence (2017)
80. Tang, D., Qin, B., Liu, T., Yang, Y.: User modeling with neural network for review rating prediction. In: Twenty-Fourth International Joint Conference on Artificial Intelligence (2015)
81. Asghar, N.: Yelp dataset challenge: review rating prediction. arXiv preprint arXiv:1605.05362 (2016)
82. Kasia kulma blog. https://kkulma.github.io//
83. Kaggle Tweet Dataset. https://www.kaggle.com/benhamner/clinton-trump-tweets//
84. NIPS 2018. https://nips.cc/
85. Lee, H., et al.: An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. Nature Biomedical Engineering, p. 1 (2018)

fridaya@uef.fi

86. Viganò, L., Magazzeni, D.: Explainable security. arXiv preprint arXiv:1807.04178 (2018)
87. Melis, M., Maiorca, D., Biggio, B., Giacinto, G., Roli, F.: Explaining black-box android malware detection. In: 2018 26th European Signal Processing Conference (EUSIPCO), pp. 524–528. IEEE (2018)
88. Marino, D.L., Wickramasinghe, C.S., Manic, M.: An adversarial approach for explainable AI in intrusion detection systems. In: IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society, pp. 3237–3243. IEEE (2018)
89. Grosse, K., Manoharan, P., Papernot, N., Backes, M., McDaniel, P.: On the (statistical) detection of adversarial examples. arXiv preprint arXiv:1702.06280 (2017)
90. Ahmadi, M., Ulyanov, D., Semenov, S., Trofimov, M., Giacinto, G.: Novel feature extraction, selection and fusion for effective malware family classification. In: Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy. ACM, pp. 183–194 (2016)
91. Ronen, R., Radu, M., Feuerstein, C., Yom-Tov, E., Ahmadi, M.: Microsoft malware classification challenge. arXiv preprint arXiv:1802.10135 (2018)
92. CircleUpArticle. https://circleup.com/blog/2018/11/14/circleup-at-grace-hopper-the-degree-of-interpretability-in-ml-depending-on-your-problem-space

fridaya@uef.fi

Author Index

A
Abdelmaboud, Abdelzahir, 1030
Abuzitar, Raed, 771
Adama, Victor Ndako, 368
Adepojur, Solomon Adelowo, 368
Agbo, Friday Joseph, 1061
Agoulmine, Nazim, 474
Ahmad, Aleena, 190
Ahmad, Fatimah, 873
Ahmad, Muhammad, 885
Ahmad, Nur Syazreen, 1217, 1235, 1247
Akram, Muhammad Usman, 358
Al Fanah, Muna, 93
Alba-Flores, Rocio, 1181
Alja'am, Jihad M., 1117
Almahdawi, Amer J., 200
Al-Muhammed, Muhammed Jassem, 771
Al-Rwashidi, Hamed, 336, 345
Al-Shoshan, Abdullah I., 294
Al-Wahaibi, Fawziya, 336, 345
Antonino, Pablo Oliveira, 902
Arai, Kohei, 1193
Ashraf, Awais, 885
Aslam, Andleeb, 190
Avila-Pesantez, Diego, 1097
Azman, Azreen, 873

B
Babiker, Ammar E., 1030
Bakaoukas, Anastasios G., 253
Banerjee, Amit, 737
Bansod, Gaurav, 566, 834
Barneaud, Camille, 727
Bashir, Muhammad Bilal, 847
Bergues, Guillermo J., 1129
Borisenko, Pavel, 612
Boulmakoul, Azedine, 99
Brambilla, Nancy, 1129
Brayshaw, Mike, 1047
Butler, Thomas, 948

C
Cadzow, Alex, 817
Çalışkan, Emin, 622
Calvo, Luz, 43
Cambou, Bertrand, 494
Carrasco-Jiménez, José C., 43
Casas, Antonio, 1181
Chakraborty, Sourav, 62
Channabasava, H., 597
Chen, Peikun, 1202
Chen, Qinghua, 178
Chen, Ying-Jen, 1170
Chowdhury, Stiphen, 29
Chu, Chao-Ting, 1138
Cosma, Georgina, 166
Crockett, Robin, 242, 281
Cucchietti, Fernando M., 43

D
Dahiphale, Vijay, 834
Dang, Chuong, 639
Darwish, Roba, 797
de Amorim, Renato Cordeiro, 29
de Valroger, Thibault, 582

E
Egbokhare, F. A., 936
El Moniem, Dina Abbas Abd, 1

© Springer Nature Switzerland AG 2019
K. Arai et al. (Eds.): CompCom 2019, AISC 998, pp. 1293–1295, 2019.
https://doi.org/10.1007/978-3-030-22868-2

fridaya@uef.fi

F
Farzand, Habiba, 885
Feltus, Christophe, 474

G
Gambäck, Björn, 76
Garcia-Saez, Artur, 43
Ghafarian, Ahmad, 514
Ghafoor, Salman, 230
Ghazinour, Kambiz, 797
Ghose, Mrinal Kanti, 435
Ghosh, Shamik, 62
Goh, Patrick, 1217
Gordon, Neil A., 1047
Grey, Simon, 1047
Gu, Weiwei, 178
Gudla, Charan, 446
Gunnam, Leenendra Chowdary, 396

H
Haboobi, Heba, 217
Han, Yong, 419
Hasan, Mahamudul, 737
Hernandez, Alejandro S., 1078
Hongmei, Ge, 118
Hossain, Syeda Sumbul, 920
Huang, Dong, 1202
Hung, Chung-Wen, 1138

I
Ibrahim, Hellema, 1155
Idri, Abdelfettah, 99
Imeri, Adnan, 474
Imran, Ahmed, 1175
Imran-Daud, Malik, 13
Ismail, Ahmed O. A., 1030

J
Jiang, Jun, 386
Jibing, Hu, 118
Jonas, Karl, 306

K
Kadhum, Mohammad R., 217, 242, 281
Kadir, Rabiah Abdul, 873
Kafle, Him, 737
Kalimullina, Olga V., 157
Kanakis, Triantafyllos, 242, 281
Kanthimathi, S., 597
Khadraoui, Djamel, 474
Khan, Shoab Ahmad, 358
Khatoon, Noorjahan, 435
Klisowski, M., 654
Komatsu, Naoka, 1150
König, Sandra, 550
Kor, Ah-Lian, 145
Korableva, Olga N., 157
Korb, Thiago, 1259
Koutsomichalis, Marinos, 76
Kuhn, Thomas, 902
Kwong, Kai Man, 462
Kyriazis, Dimosthenis, 406

L
Laeeq, Kashif, 994
Lau, Chi Ho, 462
Lazarevski, Sanela, 145
Lee, Dong-Myung, 1230
Levina, Alla, 612
Li, Benji, 178
Liu, Chao, 419
Liu, Chun-Yu, 1170
Liu, Gia-Rong, 1138
Lopez-Rojas, Edgar A., 727
Lorenz, Birgy, 622
Lu, Yahan, 178

M
Mahmood, Ahmad K., 1030
Malik, Zeeshan Haider, 885
Mao, Wei-Lung, 1138
Marin, Guillermo, 43
Mashat, Arwa, 1187
Mathews, Sherin Mary, 1269
Memon, Zulfiqar Ali, 994
Miriam Avila P., L., 1097
Mitch, Reavis, 145
Mityakova, Viktoriya N., 157
Mokhtar, Hoda M. O., 1
Morgenstern, Andreas, 902
Mostovoy, Roman, 612

N
Nadeem, Aamer, 847
Naidu, Praveen, 566
Ndié, Thomas Djotio, 306, 318
Ng, Sing Yee, 1235, 1247
Nguyen, Hieu, 639
Nigro, Christian, 975
Nigro, Libero, 975
Nilavalan, Rajagopal, 345
Niu, Yuzhen, 1202
Noor, Fozia, 358

O
Okada, Alexandra, 1259
Okamoto, Manabu, 1150
Oloruntoba, Oluwabunmi, 166

fridaya@uef.fi

Oveh, R. O., 936
Oyelere, Solomon Sunday, 1061

P
Pan, Yong, 386
Pan, Zhenxi, 386
Paul, Sanchita, 62
Pearson, Esther, 866
Pengbo, Yang, 118

Q
Qamar, Usman, 190
Qasim, Anisa, 230
Qi, Wenfeng, 529
Qin, Ling, 529

R
Rado, Omesaad, 93
Rafat, Khan Farhan, 675
Rainer, Benjamin, 550
Ramos, André L. M., 1259
Rana, Md. Shohel, 446
Rass, Stefan, 550
Rios, Fernando, 1181

S
Saleh, Moutaz, 1117
Salunke, Rutuja, 566
Saqib, Pakizah, 190
Sawanobori, Kana, 1150
Sayama, Natsuki, 1150
Schauer, Stefan, 550
Schmeelk, Suzanna, 751
Schürrer, Clemar, 1129
Sciammarella, Paolo F., 975
Sembok, Tengku Mohd Tengku, 873
Sen, Malay, 62
Shah, Asadullah, 358
Shehu, Ibrahim Shehi, 368
Shi, Xiaoli, 419
Sleptsova, Daria, 612
Spence, Lisa R., 1078
Sulayman, A. Fatima, 368
Sung, Andrew H., 446
Sung, Guo-Ming, 396
Sung, Shan-Hao, 396
Symvoulidis, Chrysostomos, 406
Szabo, Robert, 708

T
Taktek, Ebtesam, 93
Teahan, William J., 200
Telcs, Andras, 708
Ting, Mary, 873
Tiong, Pei Kee, 1217
Tran, Hau, 639
Triplett, Stephanie, 1181
Tsoumas, Ilias, 406

U
Ustimenko, V., 654

V
Vaarandi, Risto, 622
Vaca-Cárdenas, Leticia, 1097
Vaca-Cárdenas, Mónica, 1097
Vo, Phuong, 639
Vu, Tu, 639

W
Wang, Xiaoyun, 529
Wu, Chensheng, 178
Wu, Jianhua, 419
Wu, Yan, 696

X
Xiao, Perry, 1155

Y
Yang, Le, 696
Yao, Shu-Nung, 1170
Yeung, Kai-Hau, 462

Z
Zambare, Ankur, 834
Zhang, Lei, 132
Zhang, Tianyi, 696
Zhang, Xue, 529
Zhao, Liutao, 386
Zheng, Zhongxiang, 529
Ziegler, Gabriela, 1009
Zikratiov, Igor, 612

fridaya@uef.fi