

# **INTERNATIONAL CONFERENCE FOR INTERNET TECHNOLOGY AND SECURED TRANSACTIONS (ICITST-2017)**

December 11-14, 2017  
Cambridge, UK

## **ICITST-2017 PROCEEDINGS**

**The 12nd International Conference for Internet Technology and Secured Transactions  
(ICITST-2017)**

**University of Cambridge, UK**

**December 11-14, 2017**

## Message from the Steering Committee

Welcome to the International Conference for Internet Technology and Secured Transactions (ICITST-2017). The ICITST provide highly professional and comparative academic research forum that promotes collaborative excellence between academia and industry. The Tables 1, 2, and 3 show all the submissions received and accepted:

Table 1. Extended Abstracts

Conferences	Countries	Extended Abstracts Submissions	First Review	Peer Review	Accepted Extended Abstracts
ICITST	84	236	192	90	24

Table 2. Full Papers

Conferences	Countries	Paper Submissions	First Review	Peer Review	Accepted Papers
ICITST	91	656	232	90	73

Table 3. Workshops

Conferences	Countries	Workshops Submissions	First Review	Peer Review	Accepted Workshops
ICITST	29	19	5	5	2

The ICITST double blind paper evaluation method was adopted to evaluate each submission and selected papers will appear in high impact International Journals.

Many people have worked very hard to make this conference possible. We would like to thank all who have helped in making ICITST a success. The Programme Committee members each deserve credit for their excellent job. We would like to thank all our Keynote and Invited Speakers who have contributed to ICITST: Professor Charles Kim, Dr Evangelos Pournaras, Dr Gary Wills, Professor Anne James, Professor Steven Furnell and Dr Enrico Malfa, for agreeing to participate in this year conference. We would also like to acknowledge our appreciation to the following organisations for their support: Infonomics Society, IEEE UK and RI Computer Chapter, BCS, IET and IAP.

On behalf of the Programme Committees, we would like to encourage you to contribute to the future ICITST as authors, speakers, workshop organisers, panellists, poster presenters and volunteer conference organisers. We wish you a pleasant stay in Cambridge and please feel free to exchange ideas with others colleagues.

General Chair  
Professor Ella Pereira, Edge Hill University, UK

Steering Committee Chair  
Professor Charles A. Shoniregun, Infonomics Society, UK and ROI



## Contents Page

Welcome Message	3
Contents Page	4
Programme Committees	12
<b>Keynote Speakers</b>	14
Keynote Speaker 1: Professor Charles Kim	15
Keynote Speaker 2: Dr Evangelos Pournaras	16
Keynote Speaker 3: Dr Gary Wills	17
Keynote Speaker 4: Professor Anne James	18
Keynote Speaker 5: Professor Steven Furnell	19
Keynote Speaker 6: Enrico Malfa	20
<b>Workshops</b>	21
<b>Workshop 1: Chaos-based Data Protection, Data Security and Hiding in Multimedia Communications (CDP-DSHMC 2017)</b> (Organisers: Safwan El Assad and Charles Shoniregun)	22
Title: RARE: A Robust Algorithm for Rapid Encryption (Authors: Tasnime Omrani, Rabei Becheikh, Olfa Mannai, Rhouma Rhouma, Safya Belghith)	23
Title: LSB-Hamming based Chaotic Steganography (LH-Steg) (Authors: Marwa Saidi, Houcemeddine Hermassi, Rhouma Rhouma, Safya Belghith)	29
Title: New Keyed Chaotic Neural Network Hash Function Based on Sponge Construction (Authors: Nabil Abdoun, Safwan El Assad, Khodor Hammoud, Rima Assaf, Mohamad Khalil, Olivier Deforges)	35
Title: Reveal false names of accounts as a result of hackers' attacks: Security systems with heightened safety of information (Authors: Desislav Andreev, Simona Petrakieva, Ina Taralova)	39
Title: Lightweight Signcryption Scheme Based on Discrete Chebyshev Maps (Authors: Ta Thi Kim Hue, Thang Manh Hoang, An Braeken)	43
<b>Workshop 2: The 8th International Workshop on Cloud Applications and Security (CAS'17)</b> (Organisers: Ella Pereira and Michael Mackay)	48
Title: Forward Secure Searchable Symmetric Encryption (Authors: Muhammad Saqib Niaz, Gunter Saake)	49
Title: CPU Workload forecasting of Machines in Data Centers using LSTM Recurrent Neural Networks and ARIMA Models (Authors: Deepak Janardhanan, Enda Barrett)	55

Title: An Advanced Reinforcement Learning Approach for Energy-Aware Virtual Machine Consolidation in Cloud Data Centers (Authors: Rachael Shaw, Enda Howley, Enda Barrett)	61
Title: Predicting Host CPU Utilization in Cloud Computing using Recurrent Neural Networks (Authors: Martin Duggan, Karl Mason, Jim Duggan, Enda Howley, Enda Barrett)	67
<b>Sessions</b>	73
<b>Session 1: Information Security</b>	74
Title: Towards a Security and Privacy Co-Creation Method (Authors: Christophe Feltus, Erik HA Proper)	75
Title: Secured Transactions Technique Based on Smart Contracts for Situational Awareness Tools (Authors: Roman Graf, Ross King)	81
Title: Speech Encryption Based on Hybrid Chaotic Key Generator for AMR-WB G.722.2 Codec (Authors: Messaouda Boumaraf, Fatiha Merazka)	87
Title: A Blind Watermarking Technique based on DCT Psychovisual Threshold for A Robust Copyright Protection (Authors: Ferda Ernawan, Muhammad Nomani Kabir, Zuriani Mustaffa)	92
Title: Data Protection by Design in Systems Development From legal requirements to technical solutions (Authors: Fredrik Blix, Salah Addin Elshekeil, Saran Laoyookhong)	98
<b>Session 2: Cyber Security</b>	104
Title: A Comparison Between API Call Sequences and Opcode Sequences as Reflectors of Malware Behavior (Authors: Saja Alqurashi, Omar Batarfi)	105
Title: Android Botnet Detection: An Integrated Source Code Mining Approach (Authors: Basil Alothman, Prapa Rattadilok)	111
Title: An Analysis of Home User Security Awareness and Education (Authors: Fayez Alotaibi, Nathan Clarke, Steven Furnell)	116
Title: Monitoring Darknet Activities by Using Network Telescope (Authors: Shaikha AlShehyari, Chan Yeob Yeun, Ernesto Damiani)	123
Title: Enhancing Cyber Security Awareness with Mobile Games (Authors: Faisal Alotaibi, Steven Furnell, Ingo Stenge, Maria Papadaki)	129
<b>Session 3: Wireless Networking and Communication</b>	135
Title: On Blockchain-Based Authorization Architecture for Beyond-5G Mobile Services (Authors: Shinsaku Kiyomoto, Anirban Basu, Mohammad Shahriar Rahman, Sushmita Ruj)	136
Title: Hybrid Public Key Authentication for Wireless Sensor Networks (Authors: Daehee Kim, Jakeun Yun, Sungjun Kim)	142

Title: Intelligent Safety Management System for Crowds Using Sensors (Author: Norah Farooqi)	144
Title: A Hybrid Approach for Femtocell Co-tier Interference Mitigation (Authors: Abdullah Alhumaidi Alotaibi, Marios C. Angelides)	148
<b>Session 4: Internet Application and Technology</b>	154
Title: Entity Identity, Performance, and Storage (Author: Aspen Olmsted)	155
Title: Automated Course Management System (Authors: Ashik Mostafa Alvi, Md. Faqru Islam Shaon, Prithvi Ranjan Das, Manazir Mustafa, Mohammad Rezaul Bari)	161
Title: Change Management and the Integration of Information Technology: Research Notes from Selected African Universities (Authors: Omotayo Kayode Abatan, Manoj Maharaj)	167
Title: Web Service Injection Attack Detection (Authors: Victor Clincy, Hossain Shahriar)	173
Title: Mobile Business Performance Metrics: Framework and Case Study (Authors: Ahyoung Kim, Junwoo Lee)	179
<b>Session 5: Internet Applications and Technology</b>	185
Title: Evaluate action primitives for Human Activity Recognition using unsupervised learning approach (Authors: Luis F. Mejia-Ricart, Paul Helling, Aspen Olmsted)	186
Title: Classification of Music by Composer Using Fuzzy Min-Max Neural Networks (Authors: Pasha Sadeghian, Casey Wilson, Stephen Goeddel, Aspen Olmsted)	189
Title: Assessment of Fuzzy Min-Max Neural Networks for Classification Tasks (Authors: Pasha Sadeghian, Aspen Olmsted)	193
Title: Low-cost Detection of Backdoor Malware (Authors: Huicong Loi, Aspen Olmsted)	197
Title: Docker vs. KVM: Apache Spark application performance and ease of use (Authors: Walter Blair, Aspen Olmsted, Paul Anderson)	199
<b>Session 6: Internet Applications and Technology</b>	202
Title: Meaningful Sandbox Data (Authors: Ryan Lile, Aspen Olmsted)	203
Title: Classifying Influenza Outbreaks by Analyzing and Filtering Twitter Data (Authors: Elizabeth Healy, Husna Siddiqui, Aspen Olmsted)	205
Title: Handling an Organization's Communication Needs with a Single Web Service (Authors: Casey Wilson, Aspen Olmsted)	208

Title: Mobile Multi-Factor Authentication (Authors: Andrew Bissada, Aspen Olmsted)	210
Title: Three Factor Authentication (Authors: William Kennedy, Aspen Olmsted)	212
<b>Session 7: Infonomics and e-Technology</b>	214
Title: A Secure Enterprise Architecture Focused on Security and Technology-transformation (SEAST) (Authors: Md. Tomig Uddin Ahmed, Nazrul Islam Bhuiya, Md. Mahbubur Rahman)	215
Title: Distributed Query Processing and Data Sharing (Authors: Ahana Roy, Aspen Olmsted)	221
Title: Factors affecting the implementation of information assurance for eGovernment in Indonesia (Authors: Rio Guntur Utomo, Robert J. Walters, Gary B. Wills)	225
Title: The Study of Public Organization's Intention to Use an Open government Data Assessment Application: Testing with an applicable TAM (Authors: Chatipot Srimuang, Nagul Cooharajanone, Uthai Tanlamai, Achara Chandrachai, Kanokwan Atchariyachanvanich)	231
<b>Session 8: Wireless Networking and Communication</b>	237
Title: Propagation modelling and performance assessment of aerial platforms deployed during emergencies (Authors: Faris A. Almalki, Marios C. Angelides)	238
Title: A Comparative Analysis of MANET Routing Protocols through Simulation (Authors: Thomas Nash, Callum Brill, Aspen Olmsted)	244
Title: Multi-Channel Steganographic Protocol for Secure SMS Mobile Banking (Authors: Omega Obinna, Eckhard Pfluegel, Charles A. Clarke, Martin J. Tunncliffe)	248
<b>Session 9: Cyber Security</b>	254
Title: On the Cost of Cyber Security in Smart Business (Authors: Igor Ivkic, Stephan Wolfauer, Thomas Oberhofer, Markus G. Tauber)	255
Title: Evaluation of Ensemble Machine Learning Methods in Mobile Threat Detection (Authors: Sanjay Kumar, Ari Viinikainen, Timo Hamalainen)	261
Title: Evaluation of AV Systems Against Modern Malware (Authors: Abidullah Zarghoon, Irfan Awan, Jules Pagna Disso, Richard Dennis)	269
Title: Security Enhancements to TLS For Improved National Control (Authors: Lamy Alqaydi, Chan Yeob Yeun, Ernesto Damiani)	274
<b>Session 10: Information Security</b>	280
Title: Towards using Transfer Learning for Botnet Detection (Authors: Basil Alothman, Prapa Rattadilok)	281

Title: Transparent Authentication: Utilising Heart Rate for User Authentication (Authors: Timibloudi S Enamamu, Nathan Clarke, Paul Haskell-Dowland, Fudong Li)	283
Title: Performance vs. Security: Implementing an Immutable Database in MySQL (Authors: Thomas Nash, Aspen Olmsted)	290
<b>Session 11: Cloud Security</b>	292
An Architecture for Privacy-preserving Sharing of CTI with 3rd party Analysis Services (Authors: Fabio Giubilo, Ali Sajjad, Mark Shackleton, David W. Chadwick, Wenjun Fan, Rogério de Lemos)	293
Title: From E-government to Cloud-government: Challenges of Jordanian Citizens' Acceptance for Public Services (Authors: Abeer Alkhwaldi, Mumtaz Kamala, Rami Qahwaji)	298
Title: Data Subject Rights in the Cloud A grounded study on data protection assurance in the light of GDPR (Authors: Alaa Altorbaq, Fredrik Blix, Stina Sörman)	305
Title: Distributed Computing Framework in Security: Case Study of Encryption Method (Authors: Shuaiyi Bu, Shuxin Yang, Haoming Ji)	311
<b>Session 12: Infonomics and e-Technology</b>	319
Title: Seeking Academic Information on the Internet: Doctoral Students' Internet Self-efficacy and Emotions (Author: Lee Yen-Mei)	320
Title: Customer Churn Analysis: A Case Study on the Telecommunication Industry of Thailand (Author: Paweena Wanchai)	325
Title: The Influence of Advanced and Secure E-Commerce Environments on Customers Behaviour: The Case of Saudis in the UK (Authors: Haya Alshehri, Farid Meziane)	332
<b>Session 13: Infonomics and e-Technology</b>	338
Title: Effects of Lexicon Size on Solving Bananagrams (Authors: Paul Helling, Ahana Roy, Aspen Olmsted)	339
Title: Work in Progress: Nobody Knows Who You Really Are Online/On the World Wide Web – Amazon-Recognition-Service based 'Virtualnet' Supported Authentication: How certain are you it is indeed you in front of a screen? (Authors: Anderson Holguin Avila, Brigitte Rodríguez Mendoza, Maria Bohorquez Sotelo, Delgadillo Loaiza Juan Sebastián)	342
Title: Considerations for OSV over Linux-based Virtual Machines (Authors: Wayne Chen, Aspen Olmsted)	346
Title: Using Analysis of Temporal Variances within a Honeypot Dataset to better predict Attack Type Probability (Authors: Seamus Dowling, Michael Schukat, Hugh Melvin)	349

<b>Session 14: Internet Applications and Technology</b>	355
Title: Client Side Calculation of 'Bacon Number' (Authors: Thomas Briggs, Aspen Olmsted)	356
Title: intel-LEACH: An Optimal Framework for Node Selection using Dynamic Clustering for Wireless Sensor Networks (Authors: Prathap Siddavaatam, Reza Sedaghat, Aakriti Tarun Sharma)	359
Title: Enhancing Security in the Cloud: when Traceability meets Access Control (Authors: Clara Bertolissi, Omar Boucelma, Worachet Uttha)	365
Title: Database Multi-factor Authentication via Pluggable Authentication Modules (Authors: Cameron Hamilton, Aspen Olmstead)	367
<b>Session 15: Internet Applications and Technology</b>	369
Title: Improve CRUD performance on hierarchical data Nested Interval model vs. nested set model (Authors: Blake Badders, Aspen Olmsted)	370
Title: Efficient Hardware Implementation of Itubee For Lightweight Application (Authors: Juhua Liu, Wei Li, Guoqiang Bai)	372
Title: Weighing the shopping benefits of a smarter refrigerator (Authors: Stephen Goeddel, Pasha Sadeghian, Aspen Olmsted)	377
Title: Optimizing Synchronization of Cloud Storage Services: Combining Benchmark Monitoring and Learning-Based Framework (Authors: Preston T. Owens, Aspen Olmsted)	379
<b>Session 16: Cloud Security and Digital Forensics</b>	381
Title: Preventing vendor lock-ins via an interoperable multi-cloud deployment approach (Authors: Roland Pellegrini, Patrick Rottmann, Georg Strieder)	382
Title: A Novel Multimedia-Forensic Analysis Tool (M-FAT) (Authors: Shahlaa Mashhadani, Hiba Al-kawaz, Nathan Clarke, Steven Furnell, Fudong Li)	388
Title: Best of Two Worlds: Secure Cloud Federations meet eIDAS (Authors: Thomas Zefferer, Dominik Ziegler, Andreas Reiter)	396
<b>Session 17: Infonomics and e-Technology</b>	402
Title: Venmo: Exposing a User's Lifestyle (Authors: Husna Siddiqui, Aspen Olmsted, Brendan Keane)	403
Title: Challenge Based Visual Speech Recognition Using Deep Learning (Authors: Philip McShane, Darryl Stewart)	405
Title: The Study of the Local Community Products (OTOP) Website Characteristics toward Buyer Decision using Eye Tracking (Authors: Nagul Cooharajanone, Krittika Akasarakul, Thipsuda Wongkhamdi, Phonkornkrit Pruetthiwongwanich, Kanokwan Atchariyachanvanich)	411

Title: Data visibility and trust enhancement of enterprise customers in cloud computing services 417  
(Authors: W.P.Yuen, K.B.Chuah)

**Session 18: Information and Cyber Security** 423

Title: Automation of Cyber-Reconnaissance: A Java-based Open Source Tool for Information Gathering 424  
(Authors: Ahana Roy, Louis Mejia, Paul Helling, Aspen Olmsted)

Title: Towards a Security Baseline for IaaS-Cloud Back-Ends in Industry 4.0 427  
(Authors: Elisabeth Bauer, Oliver Schluga, Silia Maksuti, Ani Bicaku, David Hofbauer, Igor Ivkic, Markus G. Tauber, Alexander Wöhrer)

Title: Secure E-Mail Communication – Comparison and Selection of Encryption Solutions Using an Utility Value Analysis Approach 433  
(Authors: D. Fischer, B. Markscheffel, K. Scherr)

Title: Towards Comparing Programming Paradigms 436  
(Authors: Igor Ivkic, Markus G. Tauber, Alexander Wöhrer)

**Session 19: Intelligent Control** 438

Title: Applying Dynamic Verification Tagging to the k-Anonymity Model 439  
(Authors: Ahmad Bennakhi, Mohamed A. Jeragh)

Title: Study of Factors Effecting Thailand Talent Mobility Programme: Case of University and Food Technology Industry 444  
(Authors: Kritika Kongsoontornkijkul, Rath Pichyangkura, Pakpachong Vadhanasindhu, Kanlaya Vanichbuncha)

Title: shinySDM: Point and Click Species Distribution Modeling 450  
(Authors: Thomas Nash, Aspen Olmsted)

Title: Efficient Retrieval of Information from Hierarchical REST Requests 452  
(Authors: Seth Stoudenmier, Aspen Olmsted)

**Session 20: Information Security and Intelligent Control** 455

Title: Multiple Assignment Secret Sharing Scheme Using Hierarchical Threshold Scheme 456  
(Author: Kouya Tochikubo)

Title: Bot or Not 462  
(Authors: Husna Siddiqui, Elizabeth Healy, Aspen Olmsted)

Title: Intelligent Laboratory Management System Based on Internet of Things 464  
(Authors: Yichen Ma, Fuyao Wang, Zhuozheng Wang)

**Session 21: Infonomics and e-Technology** 468

Title: Spark framework for transcriptomic trimming algorithm reduces cost of reading multiple input files 469  
(Authors: Walter Blair, Aspen Olmsted, Paul Anderson)

Title: Avoiding unnecessary deaths Drag-Back, a deadlock avoidance model 472  
(Authors: Luis Mejía-Ricart, Aspen Olmsted)

Title: Role of Transformational Leadership on E-Government Switching: Multi-Channel and Digital Divide (Authors: Khurram Mahmood, Zainab Nayyar, Hafiz Mushtaq)	475
Title: Shared Situational Awareness in Information Security Incident Management (Authors: Keshnee Padayachee, Elias Worku)	479
Title: A Collaborative System for Corporate Performance Evaluation using Gamification and the Learning Vectors Model (Authors: Michelle G. Cacais, Gilvandenys L. Sales)	484
<b>Session 22: Infonomics and e-Technology</b>	490
Title: Designing an Assembly Language Using MIT App Inventor (Authors: Casey Wilson, Anna Sandifer, Aspen Olmsted)	491
Title: Undecidable Problems in Malware Analysis (Authors: Ali Aydin Selçuk, Fatih Orhan, Berker Batur)	494
Title: Ethereum Transaction Graph Analysis (Authors: Wren Chan, Aspen Olmsted)	498
Title: Detection of fake online hotel reviews (Authors: Anna V. Sandifer, Casey Wilson, Aspen Olmsted)	501

## ICITST-2017 Programme Committees

### Honorary Chair

Kevin Warwick, Coventry University, UK

### General Chair

Ella Pereira  
Edge Hill University, UK

### General Vice-Chair

Michael Mackey  
Liverpool John Moores University, UK

### Steering Committee Chairs

Charles A. Shoniregun, Infonomics Society, UK and ROI  
Frank Wang, University of Kent, United Kingdom  
Paul Hofmann, Space-Time Insight, United States

### Steering Committees

Chan Yeob Yeun, Khalifa University of Science, Technology and Research, UAE  
George Ghinea, Brunel University, United Kingdom  
Ion Tutanescu, University of Pitesti, Romania  
Liang-Jie (LJ) Zhang, Kingdee International Software Group, China  
René Lozi, University of Nice, Sophia-Antipolis, France  
Nicu Bizon, University of Pitesti, Romania  
Safwan El Assad, Ecole polytechnique de l'université de Nantes, France

### Publication and Publicity Chair

Galyna Akmayeva, Infonomics Society, UK and ROI

### Conference Coordinators

Holly Green, Infonomics Society, UK  
Audrey Wang, Infonomics Society, UK  
Christina Pawlowska, Infonomics Society, UK

### International Chair

Ayahiko Niimi, Future University-Hakodate, Japan

### PhD Student Forum Chair

Ina Taralova, IRCCyN, France

### Posters and Demos Chair

Roberto Pereira, University of Campinas (UNICAMP), Brazil

### Workshop and Tutorial Chairs

Bernd Markscheffel, Imenau University of Technology, Germany  
Fredrick Mtenzi, Dublin Institute of Technology, Ireland

**Technical Program Chair**

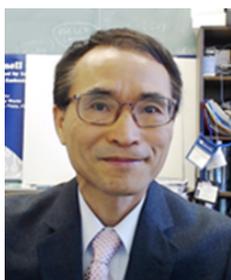
Daniel Fischer, Technische Universität Ilmenau, Germany

**Technical Program Committees**

Akhtar Hussain, KSU, Saudi Arabia  
Alex Akinbi, University of Salford, UK  
Ali Farooq, University of Turku, Finland  
Adnan Gutub, Umm Al-Qura University, Saudi Arabia  
Adrian Winckles, Anglia Ruskin University, UK  
Amr Youssef, Concordia University, Canada  
Atilla Elci, Suleyman Demirel University, Turkey  
Ayahiko Niimi, Future University-Hakodate, Japan  
Bernd Markscheffel, TU Ilmenau, Germany  
Besim Mustafa, Edge Hill University, UK  
Byoungcheon Lee, Joongbu university, Korea  
Daniel Fischer, Technische Universität Ilmenau, Germany  
Dimitrios Karras, Chalkis Institute of Technology, Athens, Greece  
Ella Pereira, Edge Hill University, United Kingdom  
Francesco Colace, University of Salerno, Italy  
Francisco Falcone, Universidad Pública de Navarra, Spain  
Hatem Halaoui, Haigazian University, Lebanon  
Héctor D. Puyosa P., Universidad Politecnica de Cartagena, Spain  
Ion Tutanescu, University of Pitesti, Romania  
Jean-Marc Robert, Ecole de Technologie Supérieure, Canada  
Luis Unzueta, Intelligent Transport Systems and Engineering Area of Vicomtech-IK4, Spain  
Mohammed Saeed, University of Chester, United Kingdom  
Mumtaz Kamala, University of Bradford, UK  
Nagwa Badr, Ain Shams University, Egypt  
Paolo Napoletano, University of Salerno, Italy  
Pavol Zavorsky, Concordia University of Edmonton, Canada  
Peter Matthew, Edge Hill University, UK  
Petre Anghelescu, University of Pitesti, Romania  
Qi Chai, University of Waterloo, Canada  
Rene' Lozi, University of Nice-Sophia Antipolis, France  
Ricardo J. Rodríguez, University of Zaragoza, Spain  
Ridha Hamila, Electrical Engineering, Qatar  
Safwan El Assad, Ecole polytechnique de l'université de Nantes, France  
Vic Grout, Glyndwr University, Wales, United Kingdom  
Zuzana Oplatkova, Tomas Bata University in Zlin, Czech Republic

## Keynote Speakers

## Keynote Speaker 1



Charles Kim is a Professor in Electrical Engineering and Computer Science at Howard University. Professor Kim's research includes fault and failure anticipation, detection, and location in aero-, naval-, and ground systems of electrical and electronic devices and networks. In line with his research on fault diagnostics, he has worked for safety and security for safety-critical systems in automotive, energy, aerospace, and nuclear industries. In control system's cybersecurity, he has extensively published on cyber-robust diversified architectures for fail-operate resilient control systems even under compromised situations. Professor Kim received a PhD in Electrical Engineering from Texas A&M University, and he is a Senior Member of IEEE).

**Title:** Cyber-Resilient ICS through Diversified Redundancy and Intrusion Detection

**Abstract:** Technological advancement of Industrial Control Systems (ICS) and control systems automation over the past decade has brought greater interconnections of the control devices. The control devices are interconnected via modern control communication bus such as ModbusTCP which now leverages Ethernet to allow interoperability between different solutions and vendors. The enhanced exchange of information has created cyber security vulnerabilities such as entry points for hackers. This presentation discusses diversified redundant architecture as a cyber-resilient system for networked control devices which maintains their normal operation even under compromised situations. This operation-based resilient architecture adds an isolated device, functionally equivalent to the networked primary device but with essential safe-mode operation features only, which is naturally immune to cyber incidents. In addition, a supervisor component monitors the operation of both networked and isolated devices, and transfers control to the safe-mode isolated device if the networked device operates abnormally, assuming that the abnormal operation is resulted from malicious hacking which has not been detected from the network cyber security defense. However, even with its resiliency, the diversified architecture has its own blind side – unawareness of presence of hacking traffic on the control bus when operational changes are not made by the hackers. Therefore, to solve this problem of unawareness of and undetected intrusion, we improve the architecture with an additional feature of intrusion detection utilizing open source software called Snort. The security features of Snort are incorporated as an intrusion detection and awareness solution, and are customized into the supervisor component of the diversified architecture for Modbus TCP/IP traffic monitoring. In laboratory tests, the added feature of intrusion detection has successfully initiated a safe-mode control transfer to the isolated device upon detection of unusual data traffic on the control bus. With this bus monitoring and intrusion detection, the diversified architecture would become a truly resilient and secure networked industrial control system.

## Keynote Speaker 2



Dr Evangelos Pournaras is a senior scientist in the Professorship of Computational Social Science, at ETH Zurich, Zurich, Switzerland. He was earlier at Delft University of Technology and VU University Amsterdam in the Netherlands, where he completed his PhD studies. He holds a MSc with distinction in Internet Computing from University of Surrey, UK and a BSc on Technology Education and Digital Systems from University of Piraeus, Greece. Evangelos has also been a visiting researcher at EPFL in Switzerland and has industry experience at IBM T.J. Watson Research Center in the USA, where he worked for the Pacific Northwest Smart Grid demonstration project. He serves the editorial board and the program committees of several international conferences and journals. He has several publications in high-impact journals and conferences in the area of self-managed distributed systems, including a best journal paper award. Smart Grids, Smart Cities and social sensing/mining are some application domains of his expertise.

**Title:** Participatory Self-management Systems for Resilient Smart Cities

**Abstract:** The pervasiveness of the Internet of Things in Smart Cities has brought paramount opportunities for improving the citizens' every life: energy and transport systems can be used more efficiently, waste collection and recycling actions can be optimized and the citizen is empowered to participate in online sharing economies for the peer-to-peer exchange of goods, e.g. energy, vehicles, tools, etc. Nevertheless, these opportunities come with several challenges that put the resilience of smart cities into risks. Several of these challenges stem from the centralized design of information and communication technologies: the concentration of data collection, storage and processing raises privacy concerns over citizens' sensitive data that allow discriminatory and profiling actions, which in turn question the autonomy and freedom of citizens, for instance, behavioral influence of online recommendations and personalized content creation. This talk raises the following question: Can participatory self-management systems orchestrated by the collective intelligence and wisdom of the crowd be used as a tactical utility to protect the citizen from these threats? What are the techno-socio-economic challenges of designing and applying decentralized systems in real-world? This talk sheds light on how open and transparent data sharing; data analytics, optimization and learning can be used by citizens, for citizens, to empower resilient smart cities. Some application scenarios discussed are the following: accurate, real-time and privacy-preserving measurements of urban qualities, load-balancing of bike-sharing stations, energy demand-response systems and self-repairable power grids using smart transformers.

### Keynote Speaker 3



Gary Wills is an Associate Professor at the University of Southampton. His research focuses on secure information systems, this includes information assurance, cloud computing, IoT and distributed ledgers. Gary is part of the Cyber Physical research group in the department of Electronics and Computer Science. Gary is also part of the Southampton's GCHQ Academic Centre of Excellence for Cyber Security Research. He is chartered engineer with the MIET.

**Title:** Looking after our IP in Supply Chain

**Abstract:** It is necessary to pass intellectual property (designs, specification, etc.) along a supply chain, but it is also known that these also 'leaks', often through poor security. This becomes more complex when the IP crosses international borders. We will examine three methods, each using security by design principles to protect our data and IP. Preserving information security policy during data integration, protecting documents when they leave an organisation, and when information is passed along the supply chain. I will also briefly look at how we move these ideas from academia to commercialisation.

## Keynote Speaker 4



Anne James is Professor and Head of Computing and Technology at Nottingham Trent University. Holding BSc and PhD degrees in computer science and data modelling, she has been engaged in higher education for many years and has focused her research in the area of data and distributed systems. She has co-authored many peer-reviewed publications and has edited books, journals and conference proceedings. She has successfully directed many PhD projects, whilst also being active in international research collaborations and serving on the committees of a number of international conferences. She is currently a member of the Network Infrastructures and Emerging Technologies (NIET) research group at Nottingham Trent University. Her current projects include Cloud Forensics and Biometric Public Key Infrastructure. Formerly Anne held the position of Professor of Data Systems Architecture at Coventry University where she oversaw research degrees across her faculty.

**Title:** Cloud Forensics

**Abstract:** As cloud technology continues to evolve with vast amount of data being transmit daily, it has added another form of complexity in forensic investigation. It is very difficult to analyse system logs during forensic investigation as cloud service providers may not be willing to share their customers' information with investigators. Furthermore a virtual machine set up in the cloud which hosts attacks might be shut down and thus logs associated with software and network access from the virtual machine would be lost. Additionally, cloud system logs transmitted over UDP or TCP packets without a robust encryption mechanism can be tampered. A further issue is that of dependency chains in the cloud where a user may use a service in one particular cloud, which in turn uses a service provided by another cloud and so on. Time and location disparities add further to the complexity. In a traditional physical network, users have significant control on their service providers (ISP) through contract agreements and policies. Cloud users lose control due to their dependency on the services which in turn depend on other services. The keynote presentation will outline the difficulties of Cloud forensics and offer some solutions.

## Keynote Speaker 5



Steven Furnell is a Professor of Information Security and leads the Centre for Security, Communications and Network Research at Plymouth University. He is also an Adjunct Professor with Edith Cowan University in Western Australia and an Honorary Professor with Nelson Mandela University in South Africa. His research interests include usability of security and privacy, security management and culture, and technologies for user authentication and intrusion detection. He has authored over 290 papers in refereed international journals and conference proceedings, as well as books including *Cybercrime: Vandalizing the Information Society* and *Computer Insecurity: Risking the System*. Professor Furnell is the current Chair of Technical Committee 11 (security and privacy) within the International Federation for Information Processing, and a member of related working groups on security management, security education, and human aspects of security. He is also a board member of the Institute of Information Security Professionals, and chairs the academic partnership committee and southwest branch.

**Title:** Taming Security Technology

**Abstract:** Security threats now confront us everywhere, and there are numerous technologies and controls that we should use in response. However, the manner in which these are presented and operated can mean that they do not deliver a favourable user experience. In fact, security is often delivered in ways that actively make it harder for users to understand and apply. This talk examines the problem, with examples from typical user-facing technologies, and presents experimental findings to illustrate the benefits that can result from supporting the user more effectively.

## Keynote Speaker 6



Enrico Malfa is Research and Development Director of the Metal Division of Tenova, a Techint Group company, worldwide partner for innovative, reliable and sustainable solution in the Metal and Mining industries. Before taking the current position in Tenova, he has responsibilities at Centro Sviluppo Materiali (CSM), now Rina Consulting, as Business Line Manager, Department Manager and Senior Scientist for environment solutions. In this position he worked developing solution for energy intensive industrial processes, including combustion systems with low environmental impact (flameless technology and regenerative system), solid materials gasification process, waste valorization, optimization of chemical energy utilization in Electrical Arc Furnaces (oxy fuel burners/injectors and control system based on off gas measurement, energy recovery). He is inventor or co-inventor in 17 patents and author/co-author of more than 60 publications. From 1996 to 2002, Enrico has held the position of Senior Scientist within the Energy System Technology area in ABB R&D Centre in Italy. Before he worked in the Cement Industry as process engineer within Combustion Department of Technology Division of Italcementi Group and in Aerospace Industry (Aermacchi) as CFD engineer. Enrico received a Master's Degree in Aerospace Engineering from Politecnico of Milan, and currently he is the Chairman of European Steel Technology Platform Working Group Planet, Member of the Technical Group Steel – Steelmaking (TGS2) of the Research Found for Steel and Coal (European Commission Directorate G – Industrial Technologies Research Fund for Coal and Steel), Member of the board of Associazione Lombarda Fabbrica Intelligente (AFIL), the cluster of intelligent manufacturing in Lombardy and Secretary of “Environmental and health committee” of AIM, the Italian Association of Metallurgy.

**Title:** Steel made in Europe: the backbone of sustainability

**Abstract:** Steel has historically been central to modern economies, synonymous with growth and progress. Modern society would be impossible without steel: Europe's reconciliation after World War II was built on unified coal and steel industries. Today the steel sector in Europe has an annual turnover of EUR 166 billion and it is responsible for 1.3% of EU GDP. Moreover steel is the essential material for a circular economy, not only for its recyclability, but because it is a material that remains available to be reintroduced into a production process in order to give birth to products or materials (permanent material). Therefore the sector has been recognized as one of three areas, along with space and defense, where the European Commission proposes specific policy measures.

At the same time Europe's steel industry has been under severe pressure, squeezed between brutal market conditions and the resolve to mitigate climate change with the associated shift to a carbon-limited world. To conquer these challenges, apart from creating and maintaining a level playing field, the European steel industry has to rely on its highly skilled workforce and on its ability to deliver technological breakthroughs.

Starting from the consideration that the sector finds itself very close to the physical limits of CO<sub>2</sub> emissions reduction from conventional steelmaking technologies, the European steel industry is fully committed to the mitigation of greenhouse gas emissions and to helping meet the objectives of the Paris Agreement working on the following main pathways towards the smart, low carbon industry of the future:

- Carbon Direct Avoidance (CDA), which substitutes carbon for hydrogen and/or via the use of electricity.
- Low Carbon Without CO<sub>2</sub> Emissions (LCWCE), which further optimises carbon-based Metallurgy and applies the circular use of waste carbon in synergy with other industrial sectors and the use of carbon storage methods to mitigate greenhouse gas emissions.
- Enhancing the recycling of steel and its by-products, helping to improve resource efficiency and reinforcing the circular economy.

The European steel industry's development plans are ambitious, and this comes at a cost, potentially of several billion euros. Thus, it is important to note that large-scale projects must have the option to apply for additional EU and national funds on top of the funding by a single joint technology initiative. Only joint initiatives with other industrial sectors, the EU institutions and the member states to support the necessarily time-consuming and expensive R&D, will foster the emergence of such breakthrough solutions. The 'Big Scale' initiative – i.e. the work on a joint initiative on low carbon steel – is a key component which will be needed to accelerate carbon reduction over the entire steel value chain.

This should also contribute to the creation of the coveted circular economy in Europe, given the huge potential of steel.

## Workshops

## **Workshop 1: Chaos-based Data Protection, Data Security and Hiding in Multimedia Communications (CDP-DSHMC 2017)**

Title: RARE: A Robust Algorithm for Rapid Encryption  
(Authors: Tasnime Omrani, Rabei Becheikh, Olfa Mannai, Rhouma Rhouma, Safya Belghith)

Title: LSB-Hamming based Chaotic Steganography (LH-Steg)  
(Authors: Marwa Saidi, Houcemeddine Hermassi, Rhouma Rhouma, Safya Belghith)

Title: New Keyed Chaotic Neural Network Hash Function Based on Sponge Construction  
(Authors: Nabil Abdoun, Safwan El Assad, Khodor Hammoud, Rima Assaf, Mohamad Khalil, Olivier Deforges)

Title: Reveal false names of accounts as a result of hackers attacks: Security systems with heightened safety of information  
(Authors: Desislav Andreev, Simona Petrakieva, Ina Taralova)

Title: Lightweight Signcryption Scheme Based on Discrete Chebyshev Maps  
(Authors: Ta Thi Kim Hue, Thang Manh Hoang, An Braeken)

# RARE: a Robust Algorithm for Rapid Encryption

Tasnime Omrani, Rabei Becheikh, Olfa Mannai, Rhouma Rhouma, Safya Belghith  
*RISC Laboratory, ENIT, University of Tunis El-Manar*  
 Tunis Tunisia

**Abstract**—Regarding the intensive use of images in the context of IoT, their security becomes crucial. Despite that several lightweight ciphers have been proposed, they are not suitable for images data due to their features such as correlation, redundancy and voluminosity. In this context, we treat in this paper the weaknesses of the existent lightweight ciphers concerning the constraints of images. Additionally, we propose an appropriate lightweight image cryptosystem that takes all the features of this kind of data. The experimental results show the effectiveness of the new scheme compared to existent ones.

**Index Terms**—Lightweight cipher, chaos, correlation, redundancy, diffusion, confusion, outer, inner.

## I. INTRODUCTION

The use of the IoT has seen an enormous development during the last decade. The IoT is integrated in several domains in our daily life such as smart home and it is used in sensitive domains like medical area, military surveillance. These kind of objects characterize by their limited resources in term of memory, CPU capability and energy supply. Several recherches have been done in this area in order to ensure the security of the exchanged data. Although, many encryption algorithms exist in the literatures, they are unsuitable for IoT even the standards ciphers such as AES [1]. For this reason, there is a birth of new type of encryption algorithms that targets the constrained devices and named lightweight encryption algorithms. Among these cryptosystems, we can cite Present [2], Simon [3], [4], Speck [3], [4], Twine [5], RoadRunner [6], Rectangle [7], Prince [8], Pride [9] and LBlock [10]. Although, the security and the high performance provided by these cryptosystems, the explosion of using the multimedia data in the applications of our daily life (like videoconferencing, surveillance sensor in the environmental or military field, and medical sensors and applications), show the weaknesses of this kind of algorithms. Indeed, a multimedia data differs from other types by its high level of correlation and redundancy. An image with high redundancy and correlation allows to reduce the search keyspace related to the statistical attacks.

In order to prove the limitations of those lightweight cryptosystems in the context of encryption of image data types, we have applied each one of the aforementioned cryptosystems in a ECB mode to encrypt a medical image depicted in Fig. 1.(a) characterized by its high correlation and redundancy. The results of this experiment is exhibited in Fig. 1 showing the insufficiency of Lightweight cryptosystems when they are

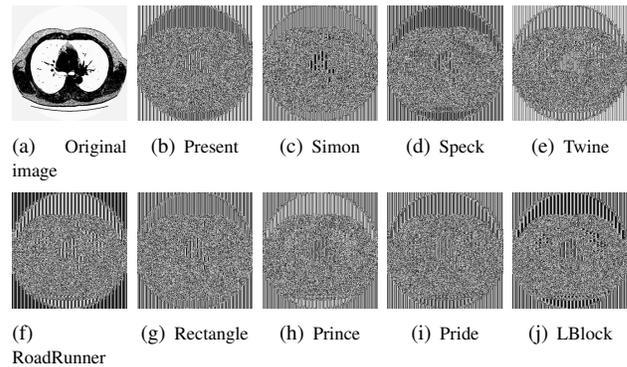


Fig. 1. Results of the encryption of a highly correlated and redundant medical image by Lightweight cryptosystems in a ECB mode.

used to encrypt images. Afterward, We encrypted the same image by the diverse algorithms by adopting CBC mode. The objective of using the last mode is to reduce the correlation and the redundancy level. Although by following the CBC mode the redundancy and the correlation issues are addressed, the diffusion criteria regarding a change in the plain-image remind not satisfied. To prove this fact, we have taken 100 images and applied Present in a CBC mode to those images. Actually, the evaluation of the the diffusion criteria is as follows: a bit is altered in an image and comparison between the ciphered original image and the ciphered modified is carried out by counting the number of differ bits. The last test allows to examine the influence of altering one bit on the ciphered image. Therefore, to investigate the effect of the other bits on the image, one should change other bits one by one and repeat the same test. To be more accurate the test must be applied on the remaining chosen images. Results given by Figure 2 show that diffusion level is always under the optimal value which is 50% although the CBC mode is used in the encryption of the plain images.

The found results showing the insufficiency of lightweight cryptosystems applied for images to elude correlation, redundancy and to enhance diffusion level constitute our motivation to propose a lightweight block encryption method specially designed to treat images. We describe in section 2 the specification of our proposed. In section 3, an experimental results regarding correlation, redundancy, voluminosity and, confusion and diffusion criteria are presented. Furthermore, a cryptanalysis comparison between the proposed cipher and existents lightweight cryptosystems are given. In section 4, the

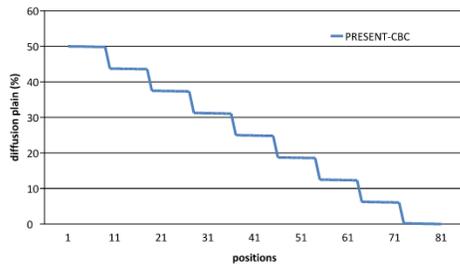


Fig. 2. Diffusion level regarding change in the plain image using CBC mode with PRESENT

performance of our proposal is evaluated and compared with the existing cryptosystems. Finally the conclusion is given in section 5.

## II. SPECIFICATION OF THE PROPOSED CRYPTOSYSTEM

### A. Structure

Processing the entire image as a single block is essential to ensure the diffusion property of Shannon on the whole image while reducing the correlation and the redundancy. However, cryptosystems that follow this structure use a high number of rounds or hybride chaotic function in order to ensure the security of their cryptosystems. This affects negatively the real-time concept. In order to take advantage of this method while assuring real-time scenario, we propose a fast pre-processing phase at the beginning of the encryption, which consists of ensuring diffusion and reducing the correlation and redundancy using a single round. We call this phase the "Outer phase". It will be followed by an "Inner-Phase" which converts the image into sub-blocks and treats them using a ECB mode of block encryption in order to lighten the processing of a voluminous data as the image, while increasing the security level by treating each block in enough number of rounds. We have called this structure "Inner-outer struct" and it is represented in Figure 3. In the following, we describe step by step this newly proposed structure through the presentation of our cryptosystem.

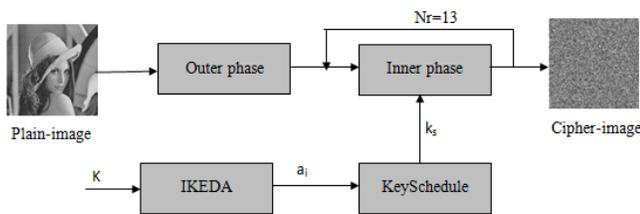


Fig. 3. Outer-Inner Structure.

### B. Outer phase

This external phase consists of two diffusion operations. An operation to scan the image horizontally by applying a modular addition as shown in Fig. 4.a. And an operation to scan the image vertically by applying slightly similar

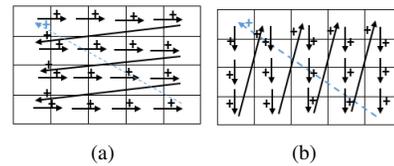


Fig. 4. Operations of outer phase: (a) Horizontal ADD-Diffusion(HAD); (b) Vertical ADD-Diffusion(VAD)

operation but in vertical orientation as shown in Fig. 4.b. The application of these two operations successively contributes on the one hand, on reducing the correlation and the redundancy constraints of the image and on the other hand, on ensuring the diffusion property of Claude Shannon. In order to validate these hypotheses, we take 100 medical images and we use the correlation coefficient as a metric to calculate the correlation level. And we used the Shannon entropy to evaluate the level of redundancy. To evaluate the diffusion level, for each image  $P_i$ , we generate an image  $P'_i$  which is equal to  $P_i$  but with a slight difference in a single bit. We then encrypt the set of images  $P_i$  and  $P'_i$  in order to obtain a set of images  $C_i$  and  $C'_i$ . Finally, the mean weight  $W_i$  of each  $R_i = C_i \oplus C'_i$  is evaluated for each image. The result of these tests are illustrated in Table I. We can conclude from the results that the use of the two operations HAD and VAD provides an optimal level of diffusion regarding a modification in the plaintext and the best reduction of correlation and redundancy.

TABLE I  
MEASURE OF CORRELATION COEFFICIENT, ENTROPY LEVEL AND DIFFUSION LEVEL AFTER APPLICATION OF HAD AND VAD OPERATIONS

	Correlation coefficient	Entropy	Diffusion level % Min/Max
HAD	0.1649	7.9927	0.008/24.91
VAD	0.1647	7.9935	0.008/24.91
(HAD) then (VAD)	0.0029	7.9971	24.9/50

### C. Inner phase

In this subsection, we will describe the building blocks involved in the inner phase

1) *S-box*: This transformation follows the mixing operation and it uses an  $4 \times 4$  S-box which is shown in Table II. This choice is due to the fact that this operation ensures Claude Shannon's confusion property.

2) *P-box*: In order to ensure the diffusion property of Claude Shannon, a permutation operation  $P(\cdot)$  is necessary. Indeed, our aim is to spread the result of each S-box on a maximum number of sub-blocks. We have for this purpose used a bit-by-bit permutation to spread each 4-bit coming out of an S-box on four adjacent sub-blocks. As shown in Fig. 5. The permutation function is designed in a way to enlarge the number of active S-boxes through the rounds of the inner-phase. This is to survive linear and differential cryptanalysis that we will evaluate them later.

TABLE II  
 THE S-BOX USED IN OUR PROPOSITION

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
S-box(i)	13	14	0	8	11	1	9	15	6	3	7	10	2	5	12	4

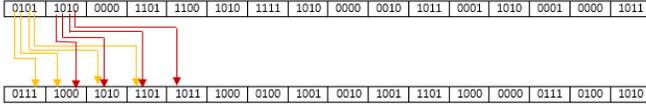


Fig. 5. Working process of the bit-Permutation

#### D. Key schedule

Our cryptosystem uses a principal key of 80-bits. This key will be used as an input to the chaotic function with  $N$  dimensions named IKEDA system [13], [14], [15], [16] which represents the core of the key schedule function (see Fig. ??). This function has the following form:

$$\begin{aligned} a_i &= a_{i+1} & i &= 1, \dots, N-1 \\ a_{N+1} &= a_N + \alpha(-\beta a_N + m \sin(a_1)) \end{aligned} \quad (1)$$

where  $N$  represents the dimension of the system,  $\alpha$ ,  $\beta$  and  $m$  denote the parameters of the system. Actually, a main key  $K$  will be used to initialize the system which in turn will be used to generate the subkey of each rounds. Indeed, the process of the initialization is designed in way that the diffusion key is ensured.

It was proven in a previous work in [17] that with  $N = 7$  and  $\alpha = 0.5$  and  $\beta = 1$ , the system shows chaotic behavior. To initialize the system we do the following:

- 1) The main key  $K$  is divided into 8 segments of length 10-bits labeled  $k_1 = K^{\{1, \dots, 10\}}, \dots, k_8 = K^{\{70, \dots, 80\}}$
- 2) The generated segments  $k_1, \dots, k_8$  will be used to yield a new variables labeled  $w_1, \dots, w_8$  as below:

$$w_i^0 = k_i \oplus (k_{i+1} \ggg 4) \oplus (k_{i+2} \ggg 2),$$

where  $\ggg$  represents right rotation operation.

- 3) To ensure that each variable  $w_i$  depends on the whole bits of the main key  $K$  the following equation is iterated 3 times

$$w_i^{t+1} = w_i^t \oplus (w_{i+1}^t \ggg 4) \oplus (w_{i+2}^t \ggg 2)$$

- 4) Once the last step is performed, the items of the IKEDA system is initialized as follows:

$$\begin{cases} a_i = 0.1 + w_i^3 / 2^{10} \text{ for } i = 1, \dots, 7 \\ m = 10 + w_8^3 / 2^5 \end{cases}$$

Once the system is initialized, the system will be iterated to generate the round subkeys. To do so, the system will be iterated 8 times for each round. Actually, in each iteration the system will output  $a_1$  while the states of the system are updated by shifting each state to the right except the leftmost

state  $a_7$  is update by  $a_7 + \alpha(-\beta a_N + m \sin(a_1))$ . The output  $a_1^i$  in round  $j$  will be quantified as below:

$$k_j^i = (a_1^i \times 2^{32}) \bmod 256.$$

Therefore, the subkey of round  $j$  is  $k_j^1 \parallel \dots \parallel k_j^8$ , where  $\parallel$  denotes the concatenation operation

Fig. 6 depicted the general structure of key schedule function. Note that the initial states of the system for each round except the first one represent the contain of system states in the last round.

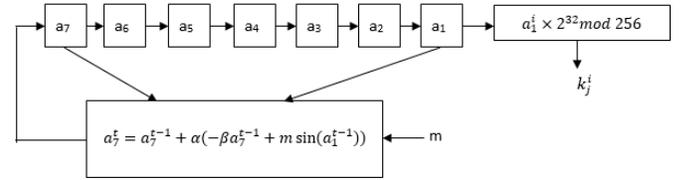


Fig. 6. KeySchedule function: iterate this function 8 times to obtained the subkey.

### III. CRYPTANALYSIS OF THE PROPOSED CRYPTOSYSTEM

In this section, we will study the response of the proposed cryptosystem to the correlation and redundancy constraints of the image. Then we will test confusion and diffusion levels. And finally, the robustness of our proposed against the linear and differential attacks will be described.

All these evaluations will be compared with the existing lightweight cryptosystems used in IoT context given their limited use of resources.

#### A. Statistical tests

Correlation, redundancy, confusion and diffusion can be used in statistical attacks, we have to be sure that the proposed cryptosystem will exhibit always cipher image with very low correlation and redundancy level and ensure the confusion and diffusion properties. To measure these properties, we have used 100 test images to be encrypted with the proposed cryptosystem and other existing lightweight cryptosystems. In table III, we present the mean value of the correlation coefficient and the entropy calculated over respective measures of those 100 test images. We report that the proposed cryptosystem shows the lowest correlation coefficient and the highest entropy level compared to other lightweight cryptosystems. The results of the confusion and diffusion measures are presented in the same table III and shows an optimal level of confusion and diffusion of the proposed scheme.

TABLE III  
CORRELATION, REDUNDANCY, CONFUSION AND DIFFUSION LEVEL OF THE PROPOSED CRYPTOSYSTEM COMPARED TO EXISTENT LIGHTWEIGHT CRYPTOSYSTEMS OVER 100 MEDICAL IMAGES

Cryptosystems	Correlation	Entropy	Confusion	Diffusion plain AVG/Min/Max	Diffusion key AVG/Min/Max
Present	0.502	6.981	52.84	0.01/0.01/0.01	49.88/44.51/55.02
Rectangle	0.489	6.977	52.54	0.01/0.01/0.01	49.66/45.55/55.14
Simon	0.441	6.927	49.65	0.01/0.01/0.01	49.67/45.31/54.09
Prince	0.449	6.931	45.89	0.01/0.01/0.01	50.18/45.29/56.42
Speck	0.435	6.987	44.58	0.01/0.01/0.01	50.02/45.67/55.48
Pride	0.435	6.978	49.43	0.01/0.01/0.01	49.96/44.37/55.23
Twine	0.442	6.974	49.43	0.01/0.01/0.01	50.19/45.86/55.13
RoadRunner	0.501	6.981	50.30	0.01/0.01/0.01	49.99/46.01/53.78
LBlock	0.491	6.977	50.05	0.01/0.01/0.01	49.89/45.77/54.70
RARE	0.003	7.997	50.003	49.989/49.81/50.01	50/49.99/50.02

B. Differential and linear cryptanalysis

Linear and differential cryptanalysis were used to almost all the existent block ciphers to measure their security [12]. Even encryption standard DES and AES were tested by those generic cryptanalysis attacks. They serve to estimate the effort needed to break the secret key or part of it. To say that a cryptosystem is resistant to these attacks, the complexity of these two attacks should be greater than the naive brute force attack.

1) *Linear cryptanalysis*: Linear cryptanalysis is a known plaintext attack. It tries to find linear approximation of the cipher between plaintext, ciphertext and the secret keys as described by the following equation:

$$\sum P_u \oplus \sum C_v = \sum K_w \tag{2}$$

with  $P_u$  is the  $u^{th}$  bit of the input  $P = [P_1, P_2, \dots, P_n]$  and  $C_v$  is the  $v^{th}$  bit of the output. With  $K_w$  is the  $k^{th}$  bit of the key.

This linear approximation holds with a probability  $P$ . By knowing this probability, we can determine the needed pairs number of plaintext/ciphertext to break the key. This number can be estimated by the following equation:

$$LC = \frac{1}{|P - \frac{1}{2}|^2} \tag{3}$$

In order to find a linear approximation of the proposed cryptosystem, we have to construct an approximation table of each nonlinear component of the cryptosystem [12] which is in this case the used S-box in the inner-phase. To create the linear approximation table of the S-box, we pick up all the possible pairs (input,output) and we compute the number of times where a linear equation relevant to each (input output) pair holds. A linear equation is true when it satisfy Eq (2). Take the example of (input, output)=(2,7). The equation relevant of this example has the following form:  $E = P_3 \oplus C_2 \oplus C_3 \oplus C_4 = 0$ , whereby  $P_i$  is the bit number  $i$  of the input 2 and  $C_i$  is the bit number  $i$  of the output 7. Therefore, to calculate the number of times in which the equation hold, we take all the bit possible  $P_i$  as depicted in the four first lines of Table IV and we generate the  $C_i$  according to the S-box illustrated in Table II. Thereafter, we calculate

the linear equation 2 (see Table IV). Then, we compute the number of time where the equation E equal to = 0 (12 in our cas). The later will be assigned to Table V after applying a soustraction operation by 8.

TABLE IV  
SAMPLE LINEAR APPROXIMATIONS OF S-BOX

$P_1$	$P_2$	$P_3$	$P_4$	$C_1$	$C_2$	$C_3$	$C_4$	$E$
0	0	0	0	1	1	0	1	0
0	0	0	1	1	1	1	0	0
0	0	1	0	0	0	0	0	1
0	0	1	1	1	0	0	0	1
0	1	0	0	1	0	1	1	0
0	1	0	1	0	0	0	1	1
0	1	1	0	1	0	0	1	0
0	1	1	1	1	1	1	1	0
1	0	0	0	0	1	1	0	0
1	0	0	1	0	0	1	1	0
1	0	1	0	0	1	1	1	0
1	0	1	1	1	0	1	0	0
1	1	0	0	0	0	1	0	1
1	1	0	1	0	1	0	1	0
1	1	1	0	1	1	0	0	0
1	1	1	1	0	1	0	0	0

Once the table is created, we have to select the most weak combination of input that activate the minimal number of S-boxes through the multiple.

In our case, the input that gives the minimal number of activated S-boxes is 0000000C00000000 in its hexadecimal form.

We can now search of the linear equation combining the inputs, the output and the round key for each encryption round. Each linear approximation for each encryption round will holds with a probability  $P_i$  and biais  $\epsilon_i = P_i - \frac{1}{2}$ . Then to find the final linear approximation, we have to combine all the linear approximation of each encryption round and compined probability  $P$  using the following Pilling-up Lemma:

$$P = \frac{1}{2} + 2^{n-1} \prod_{i=1}^n (\epsilon_i) \tag{4}$$

We illustrate in Table VI, of the linear equation of the first 4 rounds as well as the probability, the bias and the complexity. The result of our experiment showed that by reaching the

TABLE V  
LINEAR APPROXIMATION TABLE OF OUR S-BOX

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	4	-4	0	0	4	4	0
2	0	-2	-2	-4	0	-2	-2	4	2	0	0	-2	2	0	0	-2
3	0	2	-2	0	0	2	-2	0	-2	4	0	-2	-2	-4	0	-2
4	0	2	-2	0	0	2	2	4	0	-2	-2	4	0	-2	2	0
5	0	-2	2	0	-4	2	2	0	4	2	2	0	0	-2	2	0
6	0	0	0	0	-4	-4	0	0	-2	2	-2	2	2	-2	-2	2
7	0	0	4	-4	0	0	0	0	-2	-2	-2	-2	-2	-2	2	2
8	0	-2	2	0	2	0	4	2	-4	2	2	0	2	0	0	-2
9	0	-2	2	0	2	0	-4	2	0	2	2	4	-2	0	0	2
10	0	0	0	0	-2	-2	-2	-2	-2	-2	2	2	0	0	4	-4
11	0	-4	0	4	-2	2	-2	2	-2	-2	-2	-2	0	0	0	0
12	0	4	4	0	-2	2	-2	2	0	0	0	0	2	2	-2	-2
13	0	0	0	0	2	2	-2	-2	0	0	0	0	6	-2	2	2
14	0	-2	-2	-4	-2	4	0	-2	-2	0	0	2	0	2	-2	0
15	0	-2	2	0	2	0	0	-2	2	0	-4	2	0	-2	-2	-4

round number 9, the found complexity of the attack already surpass the exhaustive search complexity which mean that the proposed cryptosystem should have at least 9 encryption rounds to withstand linear cryptanalysis.

TABLE VI

LINEAR CRYPTANALYSIS OF OUR CRYPTOSYSTEM : LINEAR EQUATION, IT PROBABILITY AND ITS CORRESPONDING BIAS IN THE 4 FIRST ROUND

N	Equation	P	$\epsilon$	Complexity
1	$X_{1,28} \oplus X_{1,29} \oplus Y_{1,40} \oplus \sum K = 0$	$\frac{3}{4}$	$\frac{1}{4}$	$2^4 < 2^{80}$
2	$X_{2,40} \oplus Y_{2,46} \oplus Y_{2,49} \oplus \sum K = 0$	$\frac{5}{8}$	$\frac{1}{8}$	$2^6 < 2^{80}$
3	$X_{3,46} \oplus X_{3,49} \oplus Y_{3,51} \oplus Y_{3,53} \oplus Y_{3,54} \oplus Y_{3,56} \oplus Y_{3,57} \oplus Y_{3,60} \oplus \sum K = 0$	$\frac{17}{32}$	$\frac{1}{32}$	$2^{10} < 2^{80}$
4	$X_{4,51} \oplus X_{4,53} \oplus X_{4,54} \oplus X_{4,56} \oplus X_{4,57} \oplus X_{4,6} \oplus Y_{4,2} \oplus Y_{4,5} \oplus Y_{4,6} \oplus Y_{4,54} \oplus Y_{4,57} \oplus Y_{4,58} \oplus Y_{4,60} \oplus \sum K = 0$	$\frac{255}{512}$	$\frac{1}{512}$	$2^{18} < 2^{80}$

We give in table VII a summary about linear cryptanalysis complexity of the proposed cryptosystem compared to other lightweight ciphers. Table VII confirm that our cryptosystem with a less number of rounds compared to other lightweight ciphers, can resist linear cryptanalysis.

2) *Differential cryptanalysis*: Differential cryptanalysis is a chosen plaintext attack. It studies the propagation of two inputs difference along the cipher encryption rounds and how it affects their corresponding outputs difference. To cryptanalyse the proposed cryptosystem by the differential attack, we need to construct a difference distribution table for each nonlinear component [12]. In this case, the only one nonlinear component is the S-box used in the inner-phase. We have to choose the appropriate input difference  $\Delta X$  which activate the minimal number of activated S-boxes through all the cipher to minimise the attack complexity. We have chosen the

TABLE VII

LINEAR CRYPTANALYSIS COMPLEXITY OF THE PROPOSED CRYPTOSYSTEM AND SOME LIGHTWEIGHT CRYPTOSYSTEMS.

Cryptosystem (block size/key size)	Number of rounds	Complexity
Present 64/80	28	$2^{84}$
Simon 32/64	21	$2^{31.69}$
Simon 64/96	27	$2^{62.53}$
Speck 32/64	11	$2^{30.1}$
Speck 64/96	16	$2^{63}$
Rectangle 64/80	10	$2^{90}$
Lblock 64/80	15	$2^{66}$
RARE 64/80	9	$2^{88.83}$

difference input  $\Delta X = 000000000000B000$  which satisfies the former condition of S-box activation. We can then study the propagation of this input difference in every round output until the penultimate output of the cipher. For each round, we calculate the probability holding each output difference, then we can combine the probability of all the rounds by calculating the product of each their relative holding probability (Eq. 5). The complexity of the attack is estimated to be  $\frac{1}{P}$ .

$$P = \prod_{k=1}^{NR} P_i \tag{5}$$

The complexity of the differential cryptanalysis of the proposed cryptosystem compared to its complexity for other lightweight cryptosystems is given in table VIII. These analyses show that for a lower round number in the proposed cryptosystem, it outperforms the security of Present, Simon, Rectangle et Lblock except for Speck. This also proves that is the proposed cryptosystem is immune against differential cryptanalysis.

TABLE VIII  
COMPLEXITIES OF DIFFERENTIAL CRYPTANALYSIS ATTACK OF THE  
PROPOSED CRYPTOSYSTEM AND EXISTENT LIGHTWEIGHT  
CRYPTOSYSTEMS.

Cryptosystem (block size/key size)	Round number	Complexity
Present 64/80	25	$2^{100}$
Simon 32/64	18	$2^{31.2}$
Simon 64/96	26	$2^{63}$
Speck 32/64	10	$2^{29}$
Speck 64/96	16	$2^{63}$
Rectangle 64/80	15	$2^{66}$
Lblock 64/80	15	$2^{64}$
RARE 64/80	13	$2^{103.53}$

#### IV. PERFORMANCE

The limited power of equipment used in the context of IoT imposes the importance of designing a cryptosystem that considers this constraint. Indeed, the existing lightweight cryptosystems responds to this constraint. However, they are intended for text or binary data and are not intended for multimedia data (see section 1). Hence the importance of evaluating this criteria in our cryptosystem by comparing it with the existing cryptosystem to determine if the cryptosystem corresponds to the real-time property or not. We calculated the number of clock cycles by byte for our proposed and for existing cryptosystems using C language for  $265 \times 256$  images sizes. The result is shown in the table IX and show that the proposed cryptosystem is faster then Present cipher. This implies that our cryptosystem responds to the constraint of limited power of devise used in IoT.

TABLE IX  
NUMBER OF CLOCK CYCLES BY BYTE OF THE PROPOSED CRYPTOSYSTEM  
AND EXISTENT LIGHTWEIGHT CRYPTOSYSTEMS.

Cryptosystem	Round number	Encryption	Decryption
Present	31	6708	6842
Rectangle	25	829	651
Simon	42	412	425
Prince	12	1042	1119
Speck	26	287	277
Pride	20	1279	1266
RoundRunner	10	1410	1457
Lblock	32	694	679
RARE	13	3280	3412

#### V. CONCLUSION

We have proposed a new lightweight cryptosystem for image encryption in the IoT environment. Its structure is specially designed to elude the correlation, redundancy and voluminosity constraints of the image. It follows a structure named outer-inner structure, the outer-phase pre-treats the

image rapidly to enhance confusion and diffusion criteria. The inner phase is composed of 13 rounds to resist linear and differential cryptanalysis. Comparison with existent up to date lightweight ciphers has been done. The analyses show that the proposed cryptosystem outperforms those ciphers in term of security with a lower number of rounds.

#### REFERENCES

- [1] Simon Heron, Advanced Encryption Standard (AES), In Network Security, Volume 2009, Issue 12, 2009, Pages 8-12, ISSN 1353-4858, [https://doi.org/10.1016/S1353-4858\(10\)70006-4](https://doi.org/10.1016/S1353-4858(10)70006-4).
- [2] Bogdanov, A., Knudsen, L. R., Leander, G., Paar, C., Poschmann, A., Robshaw, M. J. B., Seurin, Y. and Vikkelsoe, C., "PRESENT: An Ultra-Lightweight Block Cipher". Cryptographic Hardware and Embedded Systems - CHES 2007: 9th International Workshop, Vienna, Austria, September 10-13, 2007. Proceedings, 2007, 450-466.
- [3] Ray Beaulieu, Douglas Shors, Jason Smith, Stefan Treatman-Clark, Bryan Weeks and Louis Wingers, "SIMON and SPECK: Block Ciphers for the Internet of Things". Cryptology ePrint Archive, Report 2015/585, 2015, <http://eprint.iacr.org/2015/585>.
- [4] Ray Beaulieu, Douglas Shors, Jason Smith, Stefan Treatman-Clark, Bryan Weeks and Louis Wingers, "The SIMON and SPECK Families of Lightweight Block Ciphers". Cryptology ePrint Archive, Report 2013/404, 2013, <http://eprint.iacr.org/2013/404>.
- [5] Suzuki, Tomoyasu and Minematsu, Kazuhiko and Morioka, Sumio and Kobayashi, Eita, "TWINE: A Lightweight Block Cipher for Multiple Platforms". Selected Areas in Cryptography, 7707, 339-354, 2012, Springer. Adnan BaysalEmail authorShap ahin
- [6] Baysal, A. and Sahin, S., "Roadrunner: A small and fast bitslice block cipher for low cost 8-bit processors". International Workshop on Lightweight Cryptography for Security and Privacy, 58-76, 2015, Springer.
- [7] Zhang, W., Bao, Z., Lin, D., Rijmen, V., Yang, B. and Verbauwhede, I., "RECTANGLE: a bit-slice lightweight block cipher suitable for multiple platforms". Science China Information Sciences, 58, 12, 1-15, 2015, Springer.
- [8] Borghoff, J., Canteaut, A., and Güneysu, T., Kavun, E. B., Knezevic, M., Knudsen, L. R., Leander, G., Nikov, V., Paar, C., Rechberger, C. and others, "PRINCE—a low-latency block cipher for pervasive computing applications". International Conference on the Theory and Application of Cryptology and Information Security, 208-225, 2012, Springer.
- [9] Albrecht, M. R., Driessen, B., Kavun, E. B., Leander, G., Paar, C. and Yalçın, Tolga, "Block ciphers—focus on the linear layer (feat. PRIDE)". International Cryptology Conference, 57-76, 2014, Springer.
- [10] Wu, Wenling and Zhang, Lei, "LBlock: a lightweight block cipher". Applied Cryptography and Network Security, 327-344, 2011, Springer.
- [11] Des, Data Encryption Standard. In FIPS PUB 46, Federal Information Processing Standards Publication, 1977, 46-2.
- [12] Heys, Howard M, "A tutorial on linear and differential cryptanalysis". Cryptologia, 26, 3, 189-221, 2002, Taylor & Francis.
- [13] K. Ikeda, "Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system". Optics Communications, 3, 1979, 257-261.
- [14] Lakshmanan, Muthusamy, Senthilkumar, Vijayan, D. "Dynamics of Nonlinear Time Delay Systems", Springer, 2011.
- [15] J.D. Farmer, "Chaotic attractors of an Infinite Dimensional Dynamical system". Physica D, 1982, 366-393.
- [16] K. T. Alligood and T. D. Sauer and J. A. Yorke, "An introduction to dynamical systems". Springer Verlag, New York, USA, 1996.
- [17] O. Mannai and R. Bechikh and H. Hermassi and R. Rhouma and S. Belghith, "A new image encryption scheme based on a simple first-order time-delay system with appropriate nonlinearity". Nonlinear Dynamics, 82, 2015.

# LSB-Hamming based Chaotic Steganography (LH-Steg)

Marwa Saidi, Houcemeddine Hermassi, Rhouma Rhouma, Safya Belghith  
*RISC Laboratory, ENIT, University of Tunis El-Manar*  
 Tunis Tunisia

**Abstract**—This work introduces a novel secret-key spatial steganographic approach based on Hamming codes and LSB embedding. The embedding artifacts are restricted to rich textured zones of the cover image. We exploited the standard deviation measurement to select potential candidate block of pixels to hold the secret message. In the aim of ensuring the security aspect of our proposed method, we use the Skew Tent Map as a random generator to select exploited bits in the embedding procedure. Such a chaotic function shows an efficient performance in terms of randomness and perfect sensibility to slight alteration of the initial condition or the control parameter (sensibility:  $10^{-14}$ ). The proposed approach showed an effective resistance to the state-of-art steganalysis attacks.

**Index Terms**—Spatial Steganography, LSB embedding, Hamming codes, Steganalysis, Chaos, Feature extraction, Machine Learning.

## I. INTRODUCTION

The term Steganography [3], [4] means literally "Covered Writing". Its concept based on hiding secret messages within innocents multimedia support. Generally, two types of steganography can be presented: Pure steganography where an attacker has no information about the used method or Secret-key based steganography where the method applies to Kerckhoffs principle and the security of such approach depends only on the shared secret key.

Recently, a variety of steganographic approaches in different domains whether frequency or spatial domain have been proposed.

Moving Steganography from Lab to Real life unfortunately confronts with so many critical unsolved problems [5]. However, designers nowadays still working on inventing potential new approaches limited to lab's conditions, where the embedding distortion is minimized in the aim of achieving a more effective resistance towards steganalysis attacks.

In steganalysis, a warden (eavesdropper) can play two roles: whether an active role where he intend to detect and break the integrity of secret message, or he can remain as a passive warden and his power is limited to the detection process only. In the state-of-the-art literature, two types of attacks are generally introduced, targeted steganalysis and Blind steganalysis. The first type is based on designing accurate detectors (could be based on a learning strategy as well) with consideration to alteration and artifacts caused by a specific embedding approach. In terms of classification, they are super accurate however their limitations lies on their non-adaptivity to different image formats and non-effectiveness when it comes to

newly invented embedding schemes causing totally different patterns in the stego image which have never been treated before. Whereas blind steganalysis approaches [6] are mainly established through a learning technique with wider set of features describing alterations and artifacts caused by various data hiding methods. Designing universal/blind classifiers presents a challenging task since it is crucial to consider possibly all distortions caused the embedding method regardless the nature of the domain (spatial, frequency) where the data is proceeded. The block diagram of a blind steganalysis is illustrated in Figure I.

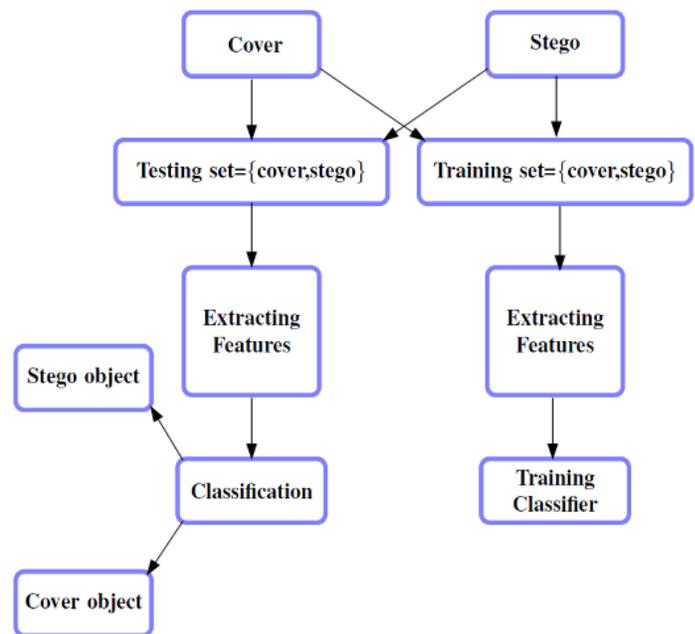


Fig. 1. Block diagram of Blind/Universal Steganalysis.

The feature extraction step can simply be described as extracting the most important properties of the image illustrating any embedding artifact. So instead of treating the data in its original form, it is more flexible and less time-consuming process to deal with a set of features. Once the feature set is extracted (stego and cover samples), the classification can be applied.

Generally, the feature designs started with basic efficient-informative alterations which are typically captivated through

dependencies measured over block of pixels in spatial domain such as analysis of variance (co-occurrence matrix, standard deviation, etc...) or dependencies between coefficients in transform domain. Typical examples of features are Image Quality Features [7], Calibration Based Feature and Calibration based Markov process Features [10], [11], Moment based Features [12], [13] and Correlation based Features [14], [15], [16]. In this paper, we exploit a pre-treatment concept where we spot the blocks of pixels verifying a certain high level of standard deviation so that the embedding will be established only within rich zones of the cover image. Stego objects created using our LH-Steg showed an interesting resistance towards steganalysis tool which implies the undetectability of the proposed approach across various payloads. Adapting the payload to each image aside and training the classifier on different payloads consists indeed of real life scenarios, in contradiction to Lab's condition which consists of training the classifier on fixed payloads.

Section II makes a brief description of three of the state of the art methods to which we compared our scheme. Section III, contains an illustration of the exploited chaotic system and its random behaviour analysis. In section IV, we introduce the embedding operation, the pre-processing phase of cover objects. The proposed LH-Steg approach is tested in Section V, where we report the classification decisions using the Ensemble Classifiers [19] and the 34671-dimensional Spatial domain Rich Model (SRM) features [17] applied over BOSS-base (images in pgm format) [18] with comparison to three of the state-of-the-art methods: WOW [20], S-UNIWARD [21] and MVGG [2]. Finally, the paper is concluded in Section VI.

## II. RELATED WORKS

Recently, various spatial adaptive steganographic approaches [8], [20], [9] have been proposed. These schemes are based on minimizing a modeled distortion function which quantifies in terms of costs the effect of changing a cover image into a stego one. Their effectiveness is shown through the correlation of designed function with statistical undetectability. In this work, we will briefly describe three different distortion-functions based approaches to which we compared our proposed scheme. Starting with the proposed WOW (wavelet obtained weights) steganography [20] which allocates high costs to more predictable pixels using directional filters residuals and low costs to less predictable pixels. S-UNIWARD [21] (Spatial Universal Wavelet Relative Distortion) has a slightly altered cost function comparing to WOW. In fact, both of the distortion functions are similar in the embedding cost computation phase, however, computing costs using WOW method makes it rather difficult to use for embedding in other domains such as the Transform domain (DCT/JPEG) whereas S-UNIWARD is more adaptive to any arbitrary domain. Another approach serves to model the cover medium through the employment of a Multivariate Generalized Gaussian (MVGG) [2] to model pixels. The parameters of the modeled cover are estimated then the probability of modifying each pixel is

analytically computed with consideration to minimizing the KullbackLeibler(KL)-divergence between the cover and the stego images. We compared our proposed LH-Steg to all of the aforementioned schemes.

## III. MATHEMATICAL BACKGROUND

### A. The Skew Tent Map

The Skew Tent Map, given by Equation (1), is a 1-D non linear function which reach a chaotic behavior for a value of control parameter  $p \in [0, 1]$  and an initial condition  $x_0 \in [0, 1]$  as given in Figure 2 which shows the chaotic aspect of this function for the indicated intervals. The Skew Tent map represents a high sensitivity to its control parameter and initial condition which make it useful in the data hiding context. In fact, the statistical properties of such function in a steganographic scenario allows the designer to protect the integrity of the secret message in case of an active warden who intend to spot potential hiding positions and alter the exchanged secret data. Thus using a chaotic function to select embedding spots will make it harder for a warden to get same spots chosen by the sender. To this end, a secret key containing the chaotic function parameters has to be exchanged before the communication would start.

$$f(x_n) = x_{n+1} = \begin{cases} x_n/p & \text{if } 0 \leq x_n \leq p \\ (1-x_n)/(1-p) & \text{if } p \leq x_n \leq 1 \end{cases} \quad (1)$$

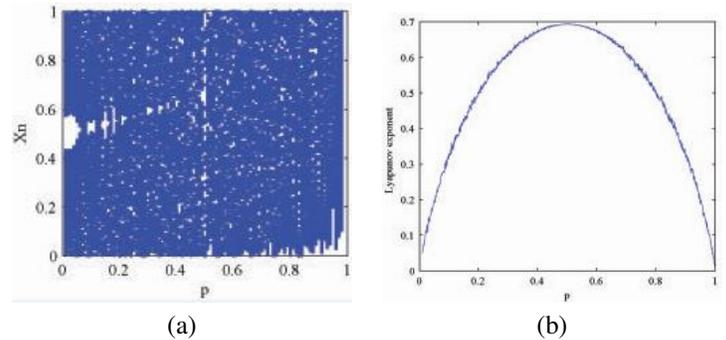


Fig. 2. The Skew Tent: (a): Bifurcation diagram (b): Lyapunov exponent

### B. Standard deviation of a discrete source

We used the standard deviation as a richness measurement (in terms of pixels values) to spot the blocks where the probability of detecting embedded data is lower comparing to other zones of the image. The standard deviation (2) describes how much a given set of data points is spread. So, an optimal threshold can be determined (optimal  $\sigma$  value) to choose only rich zones where altering pixels is an untraceable process.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (I(i, j) - mean)^2}{n \times m}} \quad (2)$$

Where  $I$  is a block of  $(n \times m)$  coefficients and  $mean$  is the mean is the average value given by equation(3).

$$mean = \frac{\sum_{i=1}^n \sum_{j=1}^m I(i, j)}{n \times m} \quad (3)$$

## IV. PROPOSED TECHNIQUE

To hide a secret image  $S$  into a cover image  $C$  of size  $(M \times N)$  we propose to apply a pretreatment process on the cover support before the embedding.

## A. Pre-treatment process

The recursive pretreatment phase, shown through Figure 3, aims to select the candidate blocks  $B_{i,j}$  which have a higher value of standard deviation (rich zones) comparing to poor regions (lower values). We describe in the next paragraph the detailed steps that an embedder has to follow to hide his secret data.

- 1) Divide the cover image  $C$  into  $(n \times n)$  blocks denoted by  $B_{i,j}$ . In our case,  $n=3$ .
- 2) For each block of pixels  $B_{i,j}$ , we determine the standard deviation  $Std(i, j)$  by applying Equation (2).
- 3) Generate a witness image related to the given cover object. We denote it by  $\omega_C$ . White blocks within  $\omega_C$  stands for candidate blocks to which a higher level of richness (amount of variation in pixel values) is assigned and black blocks stands for blocks to be eliminated from embedding. Thus,  $\omega_{C_{i,j}} = 255$  where  $Std(i, j) \geq Th$  and  $\omega_{C_{i,j}} = 0$  where  $Std(i, j) < Th$ .  $Th$  is a given Threshold initialized to  $Th_0$  and then incremented until the selection criteria of the blocks is verified.
- 4) Embed the secret message  $S$ , Generate the witness image of the stego object  $\omega_S$  and then select the candidate blocks through which the recipient will apply the extraction operation.
- 5) Extract the secret message  $\hat{S}$  by following the steps described in the subsection (IV-C).
- 6) Compare the secret message  $S$  to the estimated secret message  $\hat{S}$ . If  $\hat{S} = S$ , then the threshold  $Th$  is the optimal value which guarantees that the embedding didn't affect the standard deviation within candidate blocks. Thus same blocks are detected by the sender and the recipient. Else, the threshold is incremented by a step denoted by  $\alpha$  which has to be submitted to the receiver to perform the extraction process, so the new value of the threshold will be denoted by  $Th = Th + \alpha$ . Once the incrementing value of  $Th$  is saved, the process is performed again starting from step number 3. The convergence criteria of this recursive algorithm is  $\hat{S} = S$ .

We illustrate in 4, the case where the absolute difference between both of the witness images (stego and cover) does exist. In such scenario, the receiver is not capable to extract perfectly the secret message since the used blocks for embedding couldn't be identically detected by the receiver which is caused by not finding yet the optimal threshold up on which the classification of candidate blocks is achieved perfectly. However, Figure 5 shows the optimal output of the recursive process where same blocks are detected in the embedding and the extraction phase regarding to an optimal value of  $Th$ .

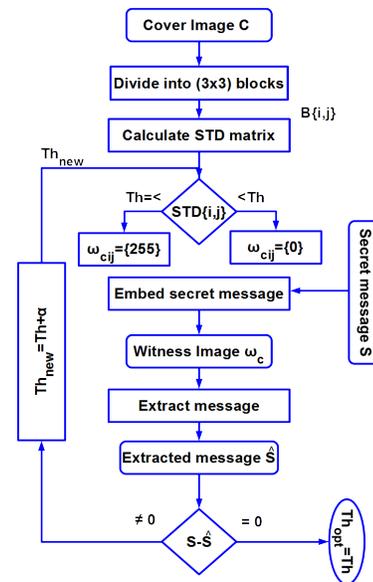


Fig. 3. The pre-treatment process

## B. The embedding process

The Embedding phase, shown by Figure 6, aims to insert a secret message  $S = \{0, 1\}$  into a cover image  $C$  of size  $(M \times N)$  by involving the hamming encoding and a chaotic generator  $G$ . In fact, while designing a steganographic system, we have to maintain and consider the protection of secret data by ensuring its integrity. So in our scheme, we used Hamming codes applied over 4 randomly chosen most significant bits (MSB) to generate 3 correcting bits, 2 of the previously generated bits will be later xored with 2 LSBs of the treated pixel and embedded in the LSBs of neighboring pixel. Such a process of correction will allow the receiver to ensure whether the secret message has been modified by an active warden or not. We denote the candidate  $(n \times n)$  blocks of the cover image by  $\tilde{B}_{ij}$ . The embedding process is performed through the following steps:

- 1) Convert diagonal pixels of each block  $\tilde{B}_{ij}(1, 1), \tilde{B}_{ij}(2, 2), \tilde{B}_{ij}(3, 3)$  from decimal encoding to 8-bits representation.
- 2) Choose 4 MSBs from the treated pixel by following a chaotic order generated through the Skew Tent map given in Equation (1). These bits are then used as an input to the Hamming encoder  $H(7,4)$ .
- 3) Calculate  $H(7,4)$ . Then choose  $2_{LSB}$  From the three generated Hamming bits.
- 4) Determine the secret bits to be hidden through an XOR operation applied over the chosen two bits of the Hamming code and two bits of the original secret message (Equation (4)). The obtained bits are embedded in the LSBs of neighboring pixels which don't belong



Fig. 4. Encountered problem in the Pre-Treatment phase: Measured absolute difference between witness cover  $\omega_C$  and witness stego  $\omega_S$ .

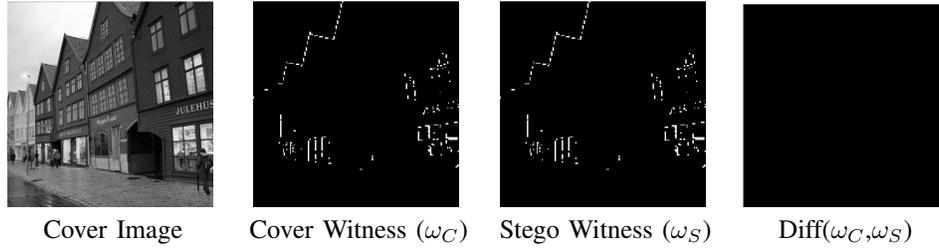


Fig. 5. Pre-Treatment phase: Optimal Threshold selected thus Identical witness images are obtained.

to the diagonal in the treated block  $B_{i,j}$ .

$$2_{LSB_{stego}} = 2_{Secret\_bits} \text{ XOR } 2_{LSB(H(7,4))} \quad (4)$$

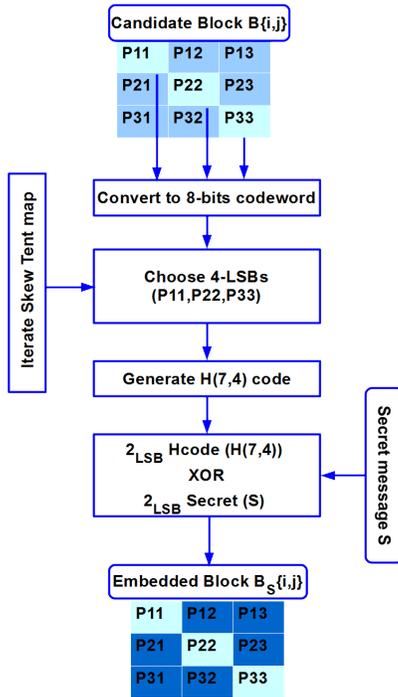


Fig. 6. The embedding process

### C. The extraction process

To extract the secret binary message  $S$  from the stego image  $SI$  we adopt the same pre-treatment steps to select upon the

Threshold  $Th$  candidate blocks used in the embedding process. In our experiments over the used image database, we choose to initiate  $Th$  value to a constant=10 and in the aim of ensuring a perfect extraction, optimal  $\alpha$  value has to be transmitted to the recipient. Following steps are applied to extract the message:

- 1) Divide the stego image  $SI$  into  $(n \times n)$  blocks denoted by  $BS_{i,j}$ .
- 2) For each image block  $BS_{i,j}$ , we calculate the standard deviation  $Std(i,j)$  by applying Equation (2).
- 3) Determine the candidate blocks  $\tilde{B}S_{i,j}$  which have a standard deviation  $std \geq Th + \alpha$ .
- 4) For each block we perform the same steps adopted in the embedding process by replacing Equation (4) by Equation (5).

$$2_{Secret\_bits} = 2_{LSB_{stego}} \text{ XOR } 2_{LSB(H(7,4))} \quad (5)$$

## V. EMPIRICAL SECURITY

Initially, we evaluate our method in terms of Peak Signal-to-Noise Ratio (PSNR) (Figure 7). We notice that our proposed method guarantees a high values of PSNR ( $67.6048 \leq PSNR \leq 80.6727$ ) measured over 50 images cover and 50 images stego besides the other steganographic methods. These measures implies that the proposed embedding approach maintains the same perceptual quality of stego comparing to cover images. Machine learning based steganalysis schemes have been widely considered as efficient tools due to the ability of extending them into multi-class detection and payload estimation. To ensure higher detection accuracy in digital steganography, we have to bear in mind the dependency of the feature extractor on the used steganographic approach. Feature-based steganalysis works systematically by adopting a specific model of the cover support, then it ensures the training, the testing and the validation phases using machine

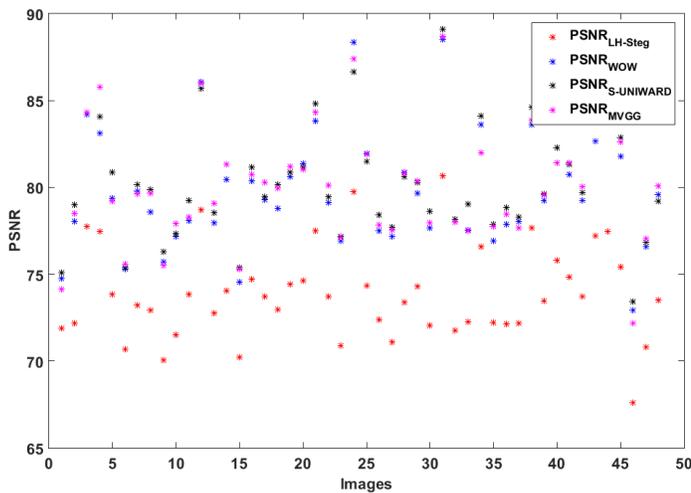


Fig. 7. PSNR measures of LH-Steg, WOW, S-UNIWARD and MVGG

learning algorithms.

We introduce the relative payload  $\theta$  in Equation 6 as the ratio between the number of messages that can be communicated (bits) and the size of the cover medium (pixels).

$$\theta = \frac{\text{Payload}}{M \times N} \quad (\text{bpp}) \quad (6)$$

With *Payload* stands for the number of secret bits. In our experiments  $0.0011 \leq \theta \leq 0.0226$

We present in the next paragraph a statistical way of attacks based on a set of features designed for spatial domain: Spatial Rich Model (SRM) (of dimension 34,671) [17] extracted from both cover and stego objects to evaluate our algorithm LH-Steg in terms of the probability of error  $P_E$  (Equation (7)).

$$P_E = \frac{1}{2} \times (P_{FA} + P_{MD}) \quad (7)$$

We used the Ensemble Classifier (EC) as a machine learning tool [19] built by merging decisions of  $K$  base learners ( $\{B_i\}$ ) implemented as The Fisher Linear Discriminant regarding to its lower computational complexity versus the ability of dealing with high dimensional feature spaces.

The decision of each base learner is regulated to reduce the total detection error where  $P_{FA}$  and  $P_{MD}$  stands respectively for the probabilities of false alarms and missed detection.

The detection errors are computed in terms of averaged  $P_E$  depicted in Equation 7 over 10 diverse random database splits (Table 8). Besides our approach, three steganographic algorithms are used for comparison: Wavelet Obtained Weights (WOW) [20], S-UNIWARD [21] and MVGG [2].

We illustrate in Figure 10, the ROC curves corresponding to the aforementioned steganographic schemes. We notice that the Ensemble Classifier performs weakly for all approaches. It is remarkable as well that the performance of our approach LH-Steg presents the smallest Area Under the

Curve (AUC=0.5072) comparing to WOW (AUC=0.5104), S-UNIWARD (AUC=0.5080) and MVGG (AUC=0.5120).

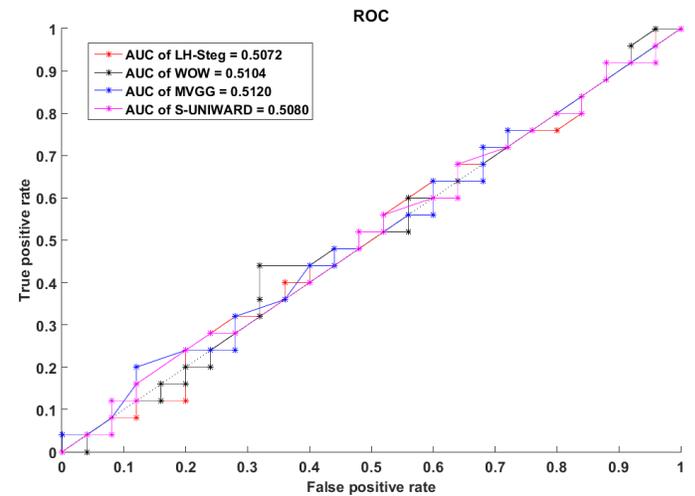


Fig. 9. The ROC curves of LH-Steg, WOW, S-UNIWARD and MVGG

The Out-of-bag (OOB) error, called as well out-of-bag estimate, presents a measure used in evaluating the prediction error of machine learning models which exploit the bootstrap aggregating concept to sub-sample training data. OOB is the mean prediction error on each training sample  $s_i$ , using only the base-learners which did not have  $s_i$  in their bootstrap sample. In general, the Out-of-bag estimates help avoid the need for an independent validation dataset. In Figure ??, we present the measurement of *OOB* as a function of the number of used base-learners. We notice that the *OOB* error is high for the three steganographic approaches which refers to the low performance in predicting the classes of cover vs stego for all training sub-samples implying the Undetectability of the proposed algorithm.

## VI. CONCLUSION

This work presents a novel spatial steganographic scheme based on the concept of restricting the embedding changes to rich zones while avoiding smooth edges which has a great impact over the detection accuracy, thus a better security. The adaptivity criteria was achieved through spotting different regions of the cover object regarding its properties and structure which implies an adaptive payload. Initially, a selecting condition is established by exploiting the standard deviation measures to spot the blocks where diversity of pixels values is presented. Then a set of bits to be proceeded through the Hamming code phase is chosen randomly using the Skew Tent map, later, after generating the Hamming bits, an *XOR* operation is established with the secret bits to generate the stego object. According to the established experiments, LH-Steg, with a fair comparison to another three approaches WOW, S-UNIWARD and MVGG, provides an interesting level of security presented in terms of detection accuracy measured empirically for a given set of images using an Ensemble Classifier (EC) and a set of high-dimensional features(SRM).

Features: SRM	$P_E$										$\overline{P_E}$
WOW	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000 (+/- 0.0000)
S-UNIWARD	0.4800	0.5000	0.5000	0.4800	0.5000	0.5200	0.5000	0.5200	0.5000	0.5000	0.5000 (+/- 0.0133)
MVGG	0.5000	0.5000	0.4800	0.5000	0.5000	0.5000	0.4800	0.5000	0.5000	0.5000	0.4960 (+/- 0.0084)
LH-Steg	0.5000	0.4800	0.5000	0.5000	0.5200	0.5000	0.5000	0.4800	0.5000	0.5000	0.4980 (+/- 0.0114)

Fig. 8. Measured  $P_E$  over 10 splits of the three described spatial domain schemes using SRM features. The mean Probability of error  $\overline{P_E}$  and its standard deviation over those 10 splits is also given.

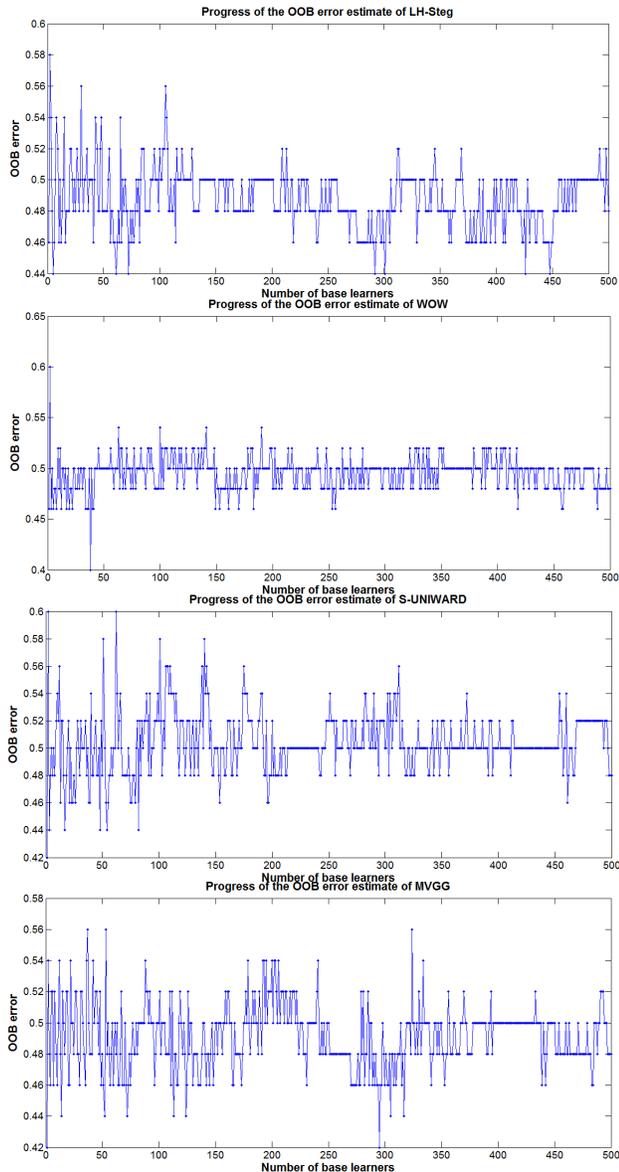


Fig. 10. OOB estimates for LH-Steg, WOW, S-UNIWARD and MVGG

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379-423, 1948.  
 [2] Sedighi, Vahid, Jessica J. Fridrich, and Rmi Cogranne. "Content-adaptive pentary steganography using the multivariate generalized Gaussian cover model." In *Media Watermarking, Security, and Forensics*, p. 94090H. 2015.

[3] Fridrich, Jessica. *Steganography in digital media: principles, algorithms, and applications*. Cambridge University Press, 2009.  
 [4] Cheddad, Abbas, Joan Condell, Kevin Curran, and Paul Mc Kevitt. "Digital image steganography: Survey and analysis of current methods." *Signal processing* 90, no. 3 (2010): 727-752.  
 [5] Ker, Andrew D., Patrick Bas, Rainer Bhme, Rmi Cogranne, Scott Craver, Tom Filler, Jessica Fridrich, and Tom Pevn. "Moving steganography and steganalysis from the laboratory into the real world." In *Proceedings of the first ACM workshop on Information hiding and multimedia security*, pp. 45-58. ACM, 2013.  
 [6] Luo, Xiang-Yang, Dao-Shun Wang, Ping Wang, and Fen-Lin Liu. "A review on blind detection for image steganography." *Signal Processing* 88, no. 9 (2008): 2138-2157.  
 [7] Avcibas, Ismail, Nasir Memon, and Blent Sankur. "Steganalysis using image quality metrics." *IEEE transactions on Image Processing* 12, no. 2 (2003): 221-229.  
 [8] Lou, Der-Chyuan, Nan-I. Wu, Chung-Ming Wang, Zong-Han Lin, and Chwei-Shyong Tsai. "A novel adaptive steganography based on local complexity and human vision sensitivity." *Journal of Systems and Software* 83, no. 7 (2010): 1236-1248.  
 [9] Liu, Guangjie, Weiwei Liu, Yuewei Dai, and Shiguo Lian. "Adaptive steganography based on block complexity and matrix embedding." *Multimedia systems* 20, no. 2 (2014): 227-238.  
 [10] Fridrich, Jessica J. "Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes." In *Information Hiding*, vol. 3200, pp. 67-81. 2004.  
 [11] Pevny, Tomas, and Jessica Fridrich. "Merging Markov and DCT features for multi-class JPEG steganalysis." In *Electronic Imaging 2007*, pp. 650503-650503. International Society for Optics and Photonics, 2007.  
 [12] Lyu, Siwei, and Hany Farid. "Detecting hidden messages using higher-order statistics and support vector machines." In *Information Hiding*, pp. 340-354. 2002.  
 [13] Goljan, Miroslav, Jessica Fridrich, and Taras Holotyak. "New blind steganalysis and its implications." *Security, Steganography, and Watermarking of Multimedia Contents* 8, no. 6072 (2006): 1.  
 [14] Sullivan, Kenneth, Upamanyu Madhow, Shivkumar Chandrasekaran, and B. S. Manjunath. "Steganalysis for Markov cover data with applications to images." *IEEE Transactions on Information Forensics and Security* 1, no. 2 (2006): 275-287.  
 [15] Shi, Yun Q., Chunhua Chen, and Wen Chen. "A Markov process based approach to effective attacking JPEG steganography." In *International Workshop on Information Hiding*, pp. 249-264. Springer, Berlin, Heidelberg, 2006.  
 [16] Chen, Chunhua, and Yun Q. Shi. "JPEG image steganalysis utilizing both intrablock and interblock correlations." In *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on*, pp. 3029-3032. IEEE, 2008.  
 [17] Fridrich, Jessica, and Jan Kodovsky. "Rich models for steganalysis of digital images." *IEEE Transactions on Information Forensics and Security* 7, no. 3 (2012): 868-882.  
 [18] Bas, Patrick, Tom Filler, and Tom Pevn. "Break Our Steganographic System: The Ins and Outs of Organizing BOSS." In *Information Hiding*, pp. 59-70. Springer Berlin/Heidelberg, 2011.  
 [19] Kodovsky, Jan, Jessica Fridrich, and Vojtech Holub. "Ensemble classifiers for steganalysis of digital media." *IEEE Transactions on Information Forensics and Security* 7, no. 2 (2012): 432-444.  
 [20] Holub, Vojtech, and Jessica Fridrich. "Designing steganographic distortion using directional filters." In *Information Forensics and Security (WIFS), 2012 IEEE International Workshop on*, pp. 234-239. IEEE, 2012.  
 [21] Holub, Vojtech, Jessica Fridrich, and Tom Denmark. "Universal distortion function for steganography in an arbitrary domain." *EURASIP Journal on Information Security* 2014, no. 1 (2014): 1.

# New Keyed Chaotic Neural Network Hash Function Based on Sponge Construction

Nabil Abdoun\*, Safwan El Assad\*, Khodor Hammoud†, Rima Assaf†, Mohamad Khalil†, Olivier Deforges‡

\*Polytech Nantes, IETR laboratory, Nantes University, Nantes, France  
e-mail: nabil.abdoun@etu.univ-nantes.fr, safwan.lassad@univ-nantes.fr

†Faculty of Engineering, Azm center, EDST, Lebanese University, Beirut, Lebanon  
e-mail: khodorhammoud94@gmail.com, rima.assaf@ul.edu.lb, mohamad.khalil@ul.edu.lb

‡INSA Rennes, IETR laboratory, Rennes, France  
e-mail: Olivier.Deforges@insa-rennes.fr

**Abstract**— This paper presents a new structure for keyed hash function based on chaotic maps, neural network and sponge construction. The structure of proposed Keyed Sponge Chaotic Neural Network *KSCNN* hash function is composed of three phases: the initialization phase pads the message  $M$  and divides it into  $q$  message blocks  $M_i$  of fixed size  $r$ , the absorbing phase hashes the message blocks by using  $CNN - Block_i$  and produces the intermediate hash value  $HM_i$  and the squeezing phase produces, starting from  $HM_q$ , the final hash value  $h$  with desired length. The combining of sponge construction with the  $CNN - Block_i$  improves, on one hand, the security of proposed hash function and makes, on the other hand, the length of hash value more dynamic. Our theoretical analysis and experimental simulations show that the proposed hash function *KSCNN* has good statistical properties, strong collision resistance, high message sensitivity compared with *SHA-3* and immune against *pre-image*, *second pre-image* and *collision* attacks.

**Keywords**—*Cryptography, Hash function, Chaotic map, Neural network, Sponge construction, Performance analysis, Cryptanalysis.*

## I. INTRODUCTION

Known as cryptographers' Swiss Army Knife, cryptographic hash functions can serve many purposes and used in different applications such as *digital signature*, *data integrity* and *message authentication*. Basically, a cryptographic hash function is a special class of hash function that has certain security properties which make it suitable for use in cryptography. It is an one-way function that maps data of arbitrary size to a hash value of a fixed size. The security level of a cryptographic hash function has been defined using the following three properties: *pre-image*, *second pre-image* and *collision resistance* [1]. To achieve these security properties, many researchers used chaotic dynamical systems and neural network structures. The features of these two systems are used to build *CNN* hash functions based on *Merkle-Damgård* structure [2] [3]. Whereas, the *Merkle-Damgård* construction has several weaknesses and also prone to several attacks such as *Herding*, *multicollisions*, *length extension* and *Joux* attacks [4]. To avoid these attacks, NIST selected *KECCAK* as the winner of Secure Hash Algorithm *SHA-3* competition on August 5<sup>th</sup> 2015. *SHA-3* is based on sponge construction that use a hashing process  $H$  operating on intermediate hash values of  $b$  bits. It can be used to model or implement *cryptographic hashes*, *message authentication codes*, *mask*

*generation functions*, *stream ciphers*, *pseudo-random number generators* and *authenticated encryption* [5]. Many other researchers have used sponge construction to build new hash functions. In 2010, *Aumasson* et al. proposed a novel design philosophy for lightweight hash functions *QUARK* in order to minimize memory requirements [6]. In [7], *Guo* et al. proposed a lightweight hash function family *PHOTON* based on *AES* design with new mixing layer method. It achieves excellent area/throughput trade-offs and very acceptable performances with simple software implementation. In [8], *Bogdanov* et al. proposed another family of lightweight hash functions *SPONGENT* with hash sizes vary from 88 (for preimage resistance only) to 256 bits. Interestingly, each of the three lightweight functions has unique characteristics, and none seems to dominate on all aspects. For example, *PHOTON* and *SPONGENT* build the permutation function  $H$  on highly optimized block cipher and have slightly lower memory footprints, whereas *QUARK* is inspired by the stream cipher *GRAIN* and by the block cipher *KATAN*. However, *SPONGENT* has a significantly lower throughput than *QUARK* and *PHOTON*, while *PHOTON* appears to have a lower security margin. Thus, the necessity of a new hash function based on sponge construction with high throughput and level security has arisen. The structure of our proposed *KSCNN* is presented in section 2, showing in detail  $CNN - Block_i$  and the operations performed in the initialization, absorbing and squeezing phases. The performance analysis of obtained results, including collision resistance, sensitivity of hash value to the message and immunity against *pre-image*, *second pre-image* and *collision* attacks are shown in section 3. Finally, our conclusion is presented in section 4.

## II. KSCNN CONSTRUCTION

The general structure of *KSCNN* is shown in Fig. 1. In this figure,  $M_i (i = 1, \dots, q)$  represents the message blocks of  $r$ -bit size where  $r = 1024$  bits,  $Z_i (i = 1, \dots)$  are the parts of the final hash value  $h$  and  $HM_i (i = 1, \dots, q, \dots)$  are the intermediate hash values with size  $b = r + c = 1536$  bits where  $c = 512$  bits ( $b$  is the width,  $r$  is the rate and  $c$  is the capacity). The initial value  $HM_0$  is initialized to  $0^b$ . The sub-keys  $KM_i (i = 1, \dots, q - 1, \dots)$ , of size 160 bits, are used by the *chaotic system* of  $CNN - Block_i$ , where the initial value  $KM_0$  is the secret key. The used *chaotic system* is already presented in our previous works [2], [3],

[9] and the structure of *KSCNN* is composed of three phases: *initialization*, *absorbing* and *squeezing*.

1- *Initialization phase*: This phase prepares the message  $M$ , the initial parameters and the secret key  $KM_0$ . In this phase,  $M$  is first padded with one bit '1' followed by the bit pattern  $00\dots 0$  of length  $v$  bits. Then, the binary value of the message length  $L$ , represented by 64 bits, is appended at the end of the message, as shown in Fig. 2 and given by equation (1):

$$v = 1024 - \text{mod}[(L + 64), 1024] - 1 \quad (1)$$

Note that, if  $L$  exceeds  $2^{64}$ , the 64 bits specified for the length of  $M$  denotes  $L \bmod 2^{64}$  [10]. Finally, the padded message is divided into  $q$  message blocks of  $r$ -bit size.

2- *Absorbing phase*: This phase calculates the intermediate hash value  $HM_i$  ( $i = 1, \dots, q$ ) for every  $r$ -bit input message block  $M_i$  ( $i = 1, \dots, q$ ). Once  $M_i$  is padded by  $c$ -bit of  $\{0\}$ , it is xored with the intermediate hash value  $HM_{i-1}$ . The result  $I_i$  of the xor operation is interleaved to *CNN-Block<sub>i</sub>*: for  $i = 1, \dots, q$  do

$$I_i \leftarrow HM_{i-1} \oplus (M_i \parallel 0^c)$$

$$HM_i \leftarrow \text{CNN} - \text{Block}_i(KM_{i-1}, I_i)$$

return  $HM_q$ .

3- *Squeezing phase*: This phase adapts the length of the intermediate hash value  $HM_q$  with the hash output's desired length  $u$  bits, when  $u$  is greater than  $b$ . In this case,  $HM_q$  is used as input of the squeezing phase and the obtained hash values  $HM_i$  ( $i \geq q$ ) is sequentially applied to *CNN-Block<sub>i</sub>*. The first  $r$  bits of each  $HM_i$  ( $i \geq q$ ) is equal to  $Z_j$  ( $j = 1, \dots$ ) and the concatenation of all obtained  $Z_j$  values constituent the final hash value  $h$ :  $Z_1 \leftarrow r \text{ bits of } HM_q, j \leftarrow 2, h \leftarrow Z_1$  for  $i = q + 1, \dots$  do

$$I_i \leftarrow HM_{i-1}, HM_i \leftarrow \text{CNN} - \text{Block}_i(KM_{i-1}, I_i)$$

$$Z_j \leftarrow r \text{ bits of } HM_i$$

$$h \leftarrow h \parallel Z_j, j \leftarrow j + 1$$

return  $h$  if length of  $h$  is equal to  $u$ .

In our proposed structure *KSCNN*, *CNN-Block<sub>i</sub>* are used in absorbing and squeezing phases. The structure of *CNN-Block<sub>i</sub>* is given in Fig. 3. The input layer has eight neurons  $\mathbf{C} = [C_0, C_1, \dots, C_7]$  and the output layer  $\mathbf{H} = [H_0, H_1, \dots, H_7]$  is composed of a combination of non-linear functions [11].

At the input layer, each neuron has 6 inputs data messages:  $P_i$ , ( $i = 0, \dots, 5$ ) for neuron  $C_0$ ;  $P_i$ , ( $i = 6, \dots, 11$ ) for neuron  $C_1$  and so on  $P_i$ , ( $i = 42, \dots, 47$ ) for neuron  $C_7$ . Each  $P_i$  is weighted by  $WI_i$ , ( $i = 0, \dots, 47$ ) and  $P_i$ ,  $WI_i$  both are samples (integer values) of 32 bits length. As we can see, we iterate the same block  $i$  many times with different input message block. For more explanation, we present in Fig. 4, the structure of neuron  $i$  in the input layer. The first three inputs  $P_j$  ( $j = 6i, \dots, 6i + 2$ ) are weighted by the  $WI_j$  ( $j = 6i, \dots, 6i + 2$ ) and then added together with the bias  $BI_i$  to form the input  $SKsi(n-1)$  of *DSTmap*. For the second three inputs,  $P_j$  ( $j = 6i + 3, \dots, 6i + 5$ ), the data message  $P_j$  are weighted by  $WI_j$  ( $j = 6i + 3, \dots, 6i + 5$ ) and then added together with the same bias  $BI_i$  to form the input  $SKpi(n-1)$

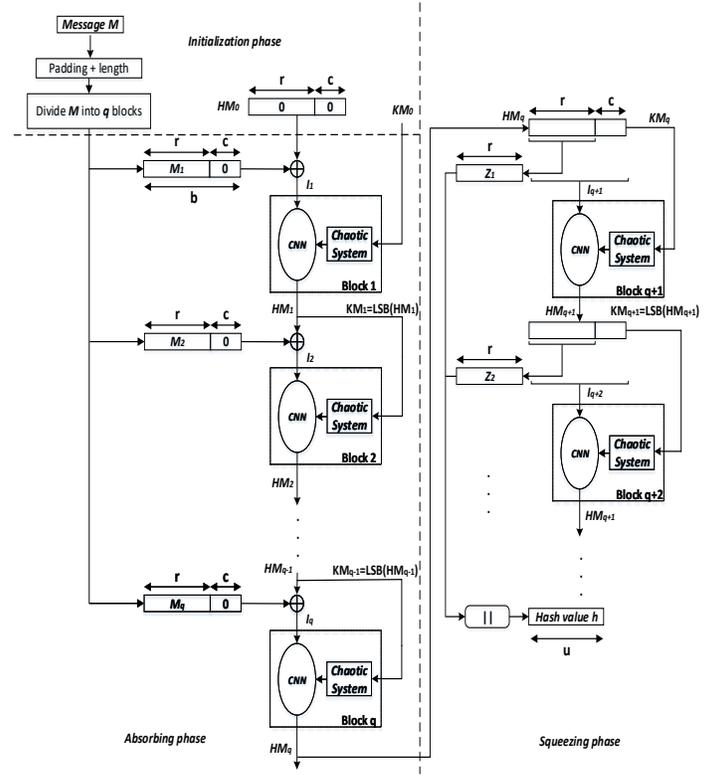


Fig. 1: Keyed Sponge Chaotic Neural Network construction

of *DPWLCmap*. The biases  $BI_i$  are also samples of 32 bits length. After computation, the two outputs of chaotic maps *DSTmap* and *DPWLCmap*  $\{SKsi(n), SKpi(n)\}$  are xored together to generate the output of neuron  $C_i$ . The output  $C_i$ , ( $i = 0, \dots, 7$ ) are weighted by  $WO_{i,i}$ , ( $i = 0, \dots, 7$ ), samples of 32 bits length and then applied to the non-linear output layer to produce the output  $H_i$ , ( $i = 0, \dots, 7$ ). The number of rounds for the output layer is 6. We choose this number of rounds in order to increase the security of *KSCNN* and to obtain the necessary length of intermediate hash value (1536 bits), 256 bits extracted in each round. When  $u > b$ , *KSCNN* enters in the squeezing phase where  $Z_i$  ( $r$  bits) are extracted from  $HM_i$  ( $i \geq q$ ) to form the final hash value  $h$  of length  $u$  bits.  $h$  is given by the following equation:

$$h = Z_1 \parallel Z_2 \parallel Z_3 \parallel \dots \quad (2)$$

### III. PERFORMANCE ANALYSIS

#### A. Statistical tests

1) *Analysis of collision resistance*: It is hard to provide a mathematical proof on the capability of collision resistance of chaotic hash functions. Thus, the following test is usually conducted to quantitative analysis on collision resistance. First, the hash value of a random message with size 1024 Bytes is generated and stored in ASCII format. Next, a bit in the message is selected randomly and toggled, a new hash value is generated and stored in the same format. The two hash values

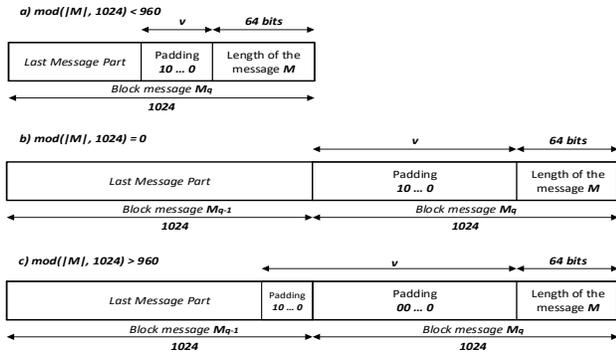


Fig. 2: Padding of input message  $M$

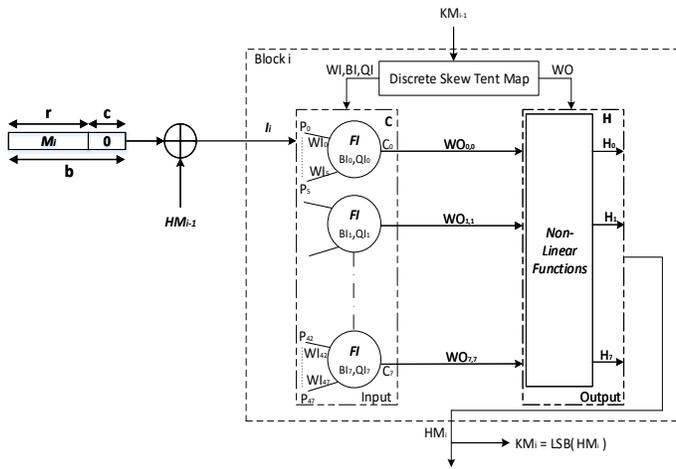


Fig. 3: Structure of  $CNN - Block_i$

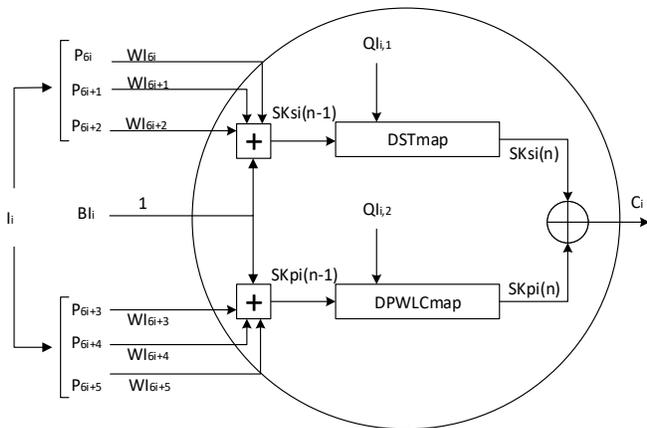


Fig. 4: Structure of neuron in input layer

	Length of hash value	Number of hits			
		0	1	2	3
SHA-3	256	1782	252	13	1
	512	1633	363	51	1
KSCNN	256	1770	260	17	1
	512	1625	370	52	1

TABLE I: Number of hits

are compared with each other, and the number of character with the same value at the same location, namely the number of hits, is counted. This test is repeated 2048-time for  $u$  equal to 256, 512 bits and the obtained results are shown in table I. As we can see, for SHA3-256, the number of 3 hits is 1, the number of 2 hits is 13, the number of 1 hit is 252 and the number of 0 hit is 1782. For our proposed KSCNN-256, the number of 3 hits is 1, the number of 2 hits is 17, the number of 1 hit is 260 and the number of 0 hit is 1770. For 512 bits, the number of test of 0 hits decreases from 1782 to 1633 for SHA3 and from 1770 to 1625 for our proposed KSCNN. The maximum number (3) of hits is appeared only one time then the collision is very low.

2) *Sensitivity of hash value to the message*: An efficient hash function should be very sensitive to any input message  $M$ , which means that a slight change in the input message should produce a completely different hash value  $h_i$ . To verify this property, for a given secret key  $KM_0$ , we calculate the number of changed bits  $B_i(h, h_i)$  and the sensitivity of hash value to the original message measured by *Hamming Distance*  $HD_i(\%)$  for the given message  $M$ : **With the wide application of internet and computer technique, information security becomes more and more important. As we know, hash function is one of the cores of cryptography and plays an important role in information security. Hash function takes a message as input and produces an output referred to as a hash value. A hash value serves as a compact representative image (sometimes called digital fingerprint) of input string and can be used for data integrity in conjunction with digital signature schemes.**

$$B_i(h, h_i) = \sum_{k=1}^{|h|} [h(k) \oplus h_i(k)] \quad (3)$$

$$HD_i(h, h_i)\% = \frac{B_i(h, h_i)}{|h|} \times 100\% \quad (4)$$

The message variants are obtained under the following conditions:

- Condition 1: The original message is the one given above.
  - Condition 2: We change the first character  $W$  in the original message to  $X$ .
  - Condition 3: We change the word **With** in the original message to **Without**.
  - Condition 4: We change the **dot** at the end of the original message into **comma**.
  - Condition 5: We add a **blank space** at the end of the original message.
  - Condition 6: We exchange the first message block  $M_1$ : **With the wide application ... function takes a mes** with the second message block  $M_2$ : **sage as input ... signature schemes**.
- In table II, we give  $B_i$  and  $HD_i(\%)$  under each condition

Message variants		$B_i$		$HD_i\%$	
		256	512	256	512
SHA-3	1	-	-	-	-
	2	116	244	45.31	47.65
	3	119	254	46.48	49.6
	4	126	252	49.21	49.21
	5	135	242	52.73	47.26
	6	118	243	46.09	47.46
KSCNN	1	-	-	-	-
	2	118	237	46.09	46.28
	3	132	251	51.56	49.02
	4	122	262	47.65	51.17
	5	123	247	48.04	48.24
	6	123	253	48.04	49.41

TABLE II: Sensitivity of hash value to the message

for  $u$  equal to 256 and 512 bits. These results indicate that  $HD_i(\%)$  is very close to ideal value 50% for our proposed *KSCNN* structure in compare to the standard SHA-3. Therefore our proposed hash function has high plain text sensitivity.

### B. Cryptanalysis

1) *Pre-image and second pre-image attacks*: In a *pre-image* attack, given only the hash value  $h$ , the attacker tries to find out the original message  $M$  such that  $H(M) = h$  without attempting to recover the secret key  $KM_0$ . In a *second pre-image* attack, the adversary knows more information. Specifically, he knows the hash value  $h$  for a given message  $M$ . It is necessary that our proposed structure *KSCNN* hash function is resistant to *pre-image* and *second pre-image* attacks. The fastest way to compute a *pre-image* or *second pre-images* is through a *brute force* attack. For a given secret key, the *brute force* method is to pick values of  $M$  at random and try each value until a collision occurs. For the *pre-image* and *second pre-image* attacks, the level of effort is proportional to  $2^u$  where  $u$  is the length of hash, which is considered too high for  $u$  equal to 128 bits. Specifically, the adversary would have to try, on average,  $2^{u-1}$  values of  $M$  to find one that generates the given hash value  $h$ . *KSCNN* produces hash values of length 256 and 512 bits, so the minimum amount of work required by an attacker to violate the *pre-image* or *second pre-image* resistance property should be  $2^{255}$  or  $2^{511}$  operations, which considered very high. Then the proposed hash functions are robust against *pre-image* and *second pre-image* attacks.

2) *Collision resistance attack or Birthday attack*: The minimum amount of work required by an attacker to violate the collision resistance property should be approximately  $2^{u/2}$  operations. The required effort is proven by a mathematical result referred to as the *birthday paradox*. *Birthday* attack, type of strongly collision resistant attack, is widely exploited for finding any two messages  $M$  and  $M'$  such that  $H(M) = H(M')$ , namely  $(M, M')$  is a collision. If the hash values are random with uniform distribution, an adversary can expect to find a collision  $(M, M')$  with probability 50 % within  $\sqrt{2^u} = 2^{u/2}$  attempts. The *collision resistance* attack requires considerably less effort than a *pre-image* or *second pre-image* attacks, as any two messages can be used to collide. In keyed hash functions, this kind of attack is only valid for an unknown given secret key  $KM_0$ . According to the above analysis, our proposed *KSCNN* hash function is one-way and secure against *pre-image* and *second pre-image* attacks. Thus, for *collision*

*resistance* attack, the length of hash value determines the security. In *KSCNN*, the hash value is 256 or 512 bits. So, the attack difficulty is  $2^{127}$  and  $2^{255}$ , on average. This keeps our proposed hash function secure against this kind of attack.

## IV. CONCLUSION

In this paper, an efficient keyed hash function based on chaotic maps, neural network and sponge construction, called Keyed Sponge Chaotic Neural Network *KSCNN*, was designed, realized and analyzed. *KSCNN* consists of three phases: initialization, absorbing and squeezing. In our structure, the used *CNN - Block<sub>i</sub>* is composed of *chaotic system* and *CNN*. The *chaotic system* generates the necessary samples to supply the *CNN* which is composed of 8 neurones in the input layer followed by combination of non-linear functions in the output layer. Theoretical analysis and simulation results showed that our proposed hash function *KSCNN* has good statistical properties, strong collision resistance, high message sensitivity compared to SHA-3 and immune against *pre-image*, *second pre-image* and *collision* attacks.

## ACKNOWLEDGMENT

This work is supported by The European Celtic-Plus project 4KREPROSYS-4K ultraHD TV wireless REMote PROduction SYStems.

## REFERENCES

- [1] M. Bellare, R. Canetti, and H. Krawczyk, "Keying hash functions for message authentication," in *Annual International Cryptology Conference*, pp. 1–15, Springer, 1996.
- [2] N. Abdoun, S. El Assad, M. A. Taha, R. Assaf, O. Deforges, and M. Khalil, "Hash function based on efficient chaotic neural network," in *International Conference on Internet Technology and Secured Transactions*, pp. 32–37, 2015.
- [3] N. Abdoun, S. El Assad, M. A. Taha, R. Assaf, O. Déforges, and M. Khalil, "Secure hash algorithm based on efficient chaotic neural network," in *The 11th International Conference on Communications*, p. comm2016, 2016.
- [4] M. M. J. Stevens *et al.*, *Attacks on hash functions and applications*. Mathematical Institute, Faculty of Science, Leiden University, 2012.
- [5] G. Bertoni, J. Daemen, M. Peeters, and G. Van Assche, "Duplexing the sponge: Single-pass authenticated encryption and other applications," in *Selected Areas in Cryptography*, vol. 7118, pp. 320–337, Springer, 2011.
- [6] J.-P. Aumasson, L. Henzen, W. Meier, and M. Naya-Plasencia, "Quark: A lightweight hash," in *CHES*, vol. 6225, pp. 1–15, Springer, 2010.
- [7] J. Guo, T. Peyrin, and A. Poschmann, "The photon family of lightweight hash functions," *Advances in Cryptology—CRYPTO 2011*, pp. 222–239, 2011.
- [8] A. Bogdanov, M. Knežević, G. Leander, D. Töz, K. Varıcı, and I. Verbauwhede, "Spongnet: A lightweight hash function," *Cryptographic Hardware and Embedded Systems—CHES 2011*, pp. 312–325, 2011.
- [9] M. A. Taha, S. E. Assad, A. Queudet, and O. Deforges, "Design and efficient implementation of a chaos-based stream cipher," *International Journal of Internet Technology and Secured Transactions*, vol. 7, no. 2, pp. 89–114, 2017.
- [10] A. J. Menezes, P. C. Van Oorschot, and S. A. Vanstone, *Handbook of applied cryptography*. CRC press, 1996.
- [11] N. FIPS, "180-2: Secure hash standard (shs)," *Information Technology Laboratory, National Institute of Standards and Technology (October 2008)*, [http://csrc.nist.gov/publications/fips/fips180-3/fips180-3\\_final.pdf](http://csrc.nist.gov/publications/fips/fips180-3/fips180-3_final.pdf), 2001.

# Reveal false names of accounts as a result of hackers attacks

Security systems with heightened safety of information

Desislav Andreev

Dep. Computer Systems, Faculty of Computer Systems and Technologies  
Technical University of Sofia  
Sofia, Bulgaria  
dessislav.andreev@gmail.com

Simona Petrakieva

Dep. Theory of Electrical Engineering, Faculty of Automation  
Technical University of Sofia  
Sofia, Bulgaria  
petrakievas-te@tu-sofia.bg

Ina Taralova

Laboratoire des Sciences du Numerique de Nantes  
Ecole Centrale de Nantes  
Nantes, France  
ina.taralova@ec-nantes.fr

**Abstract** - The aim of the paper is to reveal false names of accounts as a result of hackers attacks. The probability a given account to be either false or actual is calculated using methods for machine learning analysis. The suspected account will be used as a pattern and by classification techniques clusters will be formed with a respective probability this name to be false.

The purpose of investigation is to determine if there exists a trend, which arises during the creation of new accounts, to detect the false ones and to discriminate them from the real ones independently of that if two types of accounts are generated with the same speed.

These security systems are applied in different areas, where the security of the data in users' accounts is required. For example, they can be used in on-line voting for balloting, in studying the social opinion by inquiries, in protection of the information in different user accounts of given system etc.

**Keywords** - security systems, chaotic generator, cluster analysis, machine learning, content generation

## I. INTRODUCTION

Nowadays the most expensive thing is the information itself. Some systems contain confidential information that should be kept safe and secure from hackers' attacks at all time. The consequences of unauthorized access can be devastating from *identity theft* problems to loss of irreplaceable data. When some data is changed or deleted on purpose or by chance, it creates chaos, calling into question about the reliability or accuracy of all data in the system. Usually, users'

accounts are defended with encryption on their *username* and *password*, but this is not absolutely reliable and sometimes it needs of additional verification about the account – false or actual. In general, Windows has a build-in User Account Control which helps to the respective organization to configure its security policies [1]. Some software companies as Oracle [2], Microsoft [3], Apple [4] and etc., have developed their own techniques for managing usernames and passwords. Based on the reasons, mentioned above, to increase the reliability and safety of the data in users' accounts in the present paper a new technique for reveal false users' accounts, is proposed. It orders possibly-false user accounts as separate clusters around the respective actual one, which is used as a pattern. The proposed method determines in each cluster the probability given account to be actual or false.

As a direct impact, the proposed algorithm to reveal false names of accounts can lead to increase in voters participation, since voters will be able to securely use all modern tools for communication (phones, tablets, netbooks, etc.) to vote even when they will be far from their home cities (during bank holidays for instance), and also, simultaneously decrease the election costs usually quite high in traditional paper voting. Besides this can facilitate citizen rights for voting for disabled people who could vote from their homes even when they live far away from the voting centers.

The present paper is organized as follows. In the next section II the main requirements for chaotic generator design are proposed. Section III describes the proposed cluster analysis method for determining the probability a considered user's

account to be either false or actual. The applicability of suggested technique, combining chaotic generator and clusters analysis is illustrated on example in section IV. The paper finished with conclusion remarks about advantages and disadvantages of the new proposed security checking technique as well as the most general areas of its applicability.

## II. CHAOTIC ANALYSIS FOR DETERMINING THE PROBABILITY FOR REVEAL A FALSE ACCOUNT

Chaotic systems are defined by their extreme sensitivity to small variations in the initial conditions and parameters (known as the “butterfly effect”). There are very good candidate whenever security is required, since they can exhibit excellent statistical (i.e. “random”) features, while being quite simple for analytical design. Indeed, complex (e.g. chaotic) behavior can be exhibited also by some apparently simple piece-wise linear systems. At the same time, these systems could be considered – and analyzed - as switched systems from control theory point of view. So, switched systems in chaotic regime show up as perfect candidates able to generate chaotic sequences. Under some conditions, the latter can be considered as independent and secure pseudo-random sequences and used as information carriers.

Nonlinear chaotic functions have already been proposed [5, 6, 7, 8, 9] as interesting alternative solutions to improve security, since small deviations from the initial conditions - or the parameters - may lead to exponential divergence of the corresponding (generated) trajectories. Since, complex (e.g. chaotic) behavior can be exhibited even by some piece-wise linear (PWL) maps [8, 9]. It makes them particularly attractive for different real life applications such as Secure Communications, Information Encryption or Secure Electronic Transactions. The above maps, when used in chaotic regime, show up as perfect candidates to generate uncorrelated (independent) output signals. Under some conditions, the latter can be considered as pseudo-random sequences, and applied as chaotic pseudo-random generators or carriers for secure information transmission. The main difficulty arises from the fact that the majority of chaotic maps that are widely spread in the literature are not naturally suitable for encryption, and most of them don't exhibit even satisfactory properties for encryption, mainly because of their weak chaoticity. However, a mix of them they still can be used for random signal generation purposes.

Here many practical problems arise, from the choice of the structure and parameters of the particular chaotic generator, and its parameters, to the best coupling while satisfying predefined criteria for randomness and security. The theoretical foundations of chaotic dynamics involve also the notion of “robust” chaos, which means generally preserving chaotic features within a given parameter “window”. The latter is to

be investigated in connection with different pattern analysis algorithms. A comparative analysis of the chaotic functions is to be carried out to designate the most suitable ones according to predefined criteria from system theory point of view: uniform probability density function; chaotic attractor versus basin of attraction; positive Lyapunov exponents; correlation function decay.

Obviously, the selected chaotic functions must be tested for their capability to be (successfully) used as pseudo-random numbers generators (PRNGs) to generate seeds, or key streams. For application purposes, after binarization has been carried out, the aim is to generate sequences of quasi-random binary values. If so, it should not be possible to predict the future bit streams from the current ones, neither past bits could be retrieved out of the current and observable ones.

For that reason, besides the well-known NIST tests for randomness, different correlation analysis have been carried out to prove that past and future bits are indeed uncorrelated. In addition the quasi-random sequence should not lose its pseudo-random properties in case of small parameter modification, which should remain qualitatively the same.

The task therefore for this part of the study has been to design robust chaotic generator that could be used successfully as a pseudo-random number generator, while satisfying the above mentioned properties.

## III. CLUSTER ANALYSIS TECHNIQUE FOR FORMING SEPARATE CLUSTERS AROUND THE PATTERN

An interesting approach for detecting fake content was presented in [12], where the text is being randomly taken from different sources and concatenated in a new and realistically looking article or document. Redesigning this solution at least for account *usernames* provides the following ideas:

If the *usernames* should be created only by the real names of the users - for example an online voting website, than randomly generated names could appear if software techniques for fake content generation, mentioned in [12] are applied. Such a behavior should be detected easily, because the autonomous string creation will lead to unexpected same-letter multiplication, for example: *johnsssmith* instead of *johnsmith*<sup>1</sup>

This is obviously not enough for detecting false names, because the administrator or the MIS should not limit that much the naming convention of the users. Moreover if the fake content detector, based on relative entropy in [12] is able to mark suspicious names, this will not help in the detection of the so-called *gibberish* (also-known-as random text), for example: *fsdfasdf* instead of *fionaalder*<sup>1</sup>

<sup>1</sup>The usernames are randomly created and only suggested for presentation reasons, so they do not represent a 100 % example of user name creation in different online services.

Having these two points analyzed it becomes clear that certain algorithm for detection and representation of possibly-false names is required. With the machine learning techniques of clustering strings and simple probability measuring, the outliers in a given array of names – possibly a database on a web-server, will be detected and marked as possible intruders. The continuously arising number of the outliers in time is an obvious trend, which must be a signal for high security risk. Defining these requirements provides the opportunity for continuing this analysis.

The clustering technique is based on the concepts of similarity and distance, while proximity is determined by a distance function. It allows the generation of clusters where each of these groups consists of individuals who have common features with each other. An important point in this technique is that the groups are hardly defined a priori and the usage of the algorithm has to be verified by the person, using the system. One of the best choices for a clustering algorithms is the **k-means** one, but it's required that the user sets the number of the expected patterns. It provides a quick evaluation with good clustering quality and its implementation is easy to support [10]. In Section II we defined that with the methods of the chaos analysis we could define a number of suspected account patterns, e.g. the number of clusters we required as an input for the **k-means** algorithm. Since this methodology is currently not analyzed throughout, a simplified version of the cluster analysis and the false name probability will be used.

Since the current application is just a proof-of-concept we could limit ourselves in using the traditional technique, although in future systems an automated approach could become a better solution. Let's say we discover more patterns on the run - in that case we would require re-evaluation of the algorithms with updated number of *k*-s. A study providing solid results is already conducted and we see that such approach is meaningful, but will not provide any asset to the current research [11].

IV. EXAMPLE FOR REVEAL OF FALSE USERS' ACCOUNTS

Following the points from above the algorithm at first will check the probability of a *name* to be random generated (written with the English alphabet). This could be evaluated easily with some basic calculation of the relationship between the number of vowels and consonants in a given input. With simple observations over the names in Bulgaria, we could simply parameterize as this: The unique letters are 35 – 45 % in a name; The names are averagely 10-15 letters long; The vowels are between 45 and 55 percent in average. This will be used as the first dimension of our future clusters.

In addition to these probabilities, we will assign a possible number of close names to a given one, using a similarity measurement algorithm [13].

By *Jaro* algorithm the distance between two compared strings *s*<sub>1</sub> and *s*<sub>2</sub> determines by the following formulae:

$$sim_{jaro}(s_1, s_2) = \frac{1}{3} \cdot \left( \frac{c}{|s_1|} + \frac{c}{|s_2|} + \frac{c-t}{c} \right) \quad (1)$$

where: *sim* is the similarity function between two given strings *s*<sub>1</sub> and *s*<sub>2</sub>, measured firstly with *Jaro*'s formulae.

*c* is the number of common (matching) characters.

*t* is the minimum number of single-character transpositions required to change one string to another.

The *Jaro*'s formula (1) is used in computer science and mathematics to measure the distance between two text strings or sequences - *s*<sub>1</sub> and *s*<sub>2</sub>. Each character of *s*<sub>1</sub> is compared with all its matching characters in *s*<sub>2</sub>. A possible example for the values of *c* and *t* could be observed below:

S1	P	A	U	L
	↓	↘	↗	↓
S2	P	U	A	L

Table 1

where: *c* = 4 and *t* =  $\frac{2}{2} = 1$

One improvement of *Jaro*'s algorithm for calculation the distances between two strings *Jaro*'s is so called *Jaro-Winkler* algorithm. The latter is a good approach which is based on the following formulae:

$$sim_{wink}(s_1, s_2) = \begin{cases} sim_{jaro}(s_1, s_2), & \text{when } sim_{jaro}(s_1, s_2) < 0.7 \\ sim_{jaro}(s_1, s_2) + \frac{l}{10} \cdot (1 - sim_{jaro}(s_1, s_2)), & \text{otherwise} \end{cases} \quad (2)$$

where the new parameter *l* is the length of common prefix at the start of the string up to a maximum of four characters.

It is obvious that *Winkler*'s formulae (2) is based on the *sim* function from *Jaro*'s formula. *Winkler*'s will produce more favorable ratings for names that match from the beginning (defined by the additional prefix calculation in the second branch of (2)).

For our purpose of cluster analysis, each *name* will have the number of the similar to it measurements, and this will be the second dimension of our cluster. For the clustering we will use Lloyd's algorithm:

$$C_k = \{x_n : \|x_n - \mu_k\| \leq \text{all } \|x_n - \mu_l\|\} \quad (3)$$

$$\mu_k = \frac{1}{C_k} \sum_{x_n \in C_k} x_n \quad (4)$$

where:  $\mu_k$  and  $\mu_l$  are given sets of centroids;

*C*<sub>*k*</sub> is *k*<sup>th</sup> cluster;

*x*<sub>*n*</sub> is *n*<sup>th</sup> measurement.

The clusters are updated to contain the points that are closest in distance to each centroid. For every set of clusters the centroids are recalculated as the means of all points belonging to a cluster.

For the initial training we will use lists of existing names, where the number of the names will be the  $k$  parameter of the clustering algorithm. In addition to these names we will feed into the algorithm a list of random-name variations for each input, where a chaotic random generator from section II could be used. The parameter with the highest probability to be a real name automatically will be assigned as a centroid for each cluster and its variations will be scattered around it. On cluster would have the values with the highest probability not to be a real name or such, that's not provided with the training data? This will mark the inputs we have to check immediately. Nevertheless, the main idea for this algorithm is to mark suspicious input that will be double-checked by the system's administrator.

In the current research the clustering method - **k-means** - from machine learning theory is applied in order to partition  $n$  data or measurements into  $k$  clusters (number of Voronoi cells). The idea is that each measurement - or observed data - is attributed to a corresponding  $k$  - class, in according to some quantifiers, as the nearest mean. In the considered case,  $n$  is 200 and  $k$  is 3. The following evaluations results are shown on Figure 1.

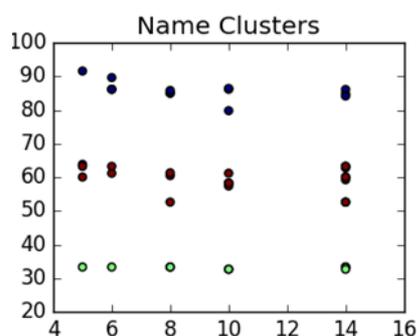


Figure 1

On the X-axis we can observe the number of similarities found for each name and on the Y-axis - the probability of a name to be invalid. The observed clusters are sparse and with possible poor quality, which as expected for few reasons - the number of names that are fed on input is not sufficient and should be bigger. For the experiment 20 names are used. They include the variations, creates a set of 200 inputs.

The numbers of similarities from then-on will be updated accordingly and the clusters should become denser. Nevertheless, this proves easily the original initial idea of the paper.

## V. CONCLUSION

In the present paper a new technique for reveal users' accounts, is suggested. It combines itself chaotic generator design used for undesirable account simulation, and cluster analysis for detection between false and real accounts. If there exist false accounts, then the actual ones shall be used as a pattern and all neighbors will form clusters with the respective probability. This technique detects false accounts from the actual ones even when false and actual accounts are generated with the same speed. On one hand, the advantage of the chaotic generator analysis there consists of the fact that by him it determines the probability for arise the false account

with name similar to this of the some actual account. On the other hand, cluster analysis accelerates the process of forming the clusters of false accounts around each actual pattern.

One of important and significant application of the proposed in this paper security system is in on-line voting for balloting. Other application of them is in studying the social opinion by inquiries. Third application is connected with protection of the information in different user accounts of given system. In this sense all their applications in different areas are pointed to increasing the security of the data in actual users' accounts. So for the applications it can lead to more secure users accounts for: e-banking, medical data exchange, health insurance etc, since users can securely and freely benefit from modern tools for communication; as for voting, which could decrease the election costs.

The latter could result especially in higher voters' participation, to facilitate direct voters involvement in governmental decisions while preserving total privacy.

## REFERENCES

- [1] <https://msdn.microsoft.com/en-us/library/cc505882.aspx>
- [2] <https://support.apple.com/en-us/>
- [3] D. Arroyo, S. Li, J. M. Amigó, G. Alvarez and R. Rhouma, "Comment on "Image encryption with chaotically coupled chaotic maps"," *Physica D*, vol. 239, p. 1002, 2010.
- [4] S. Sam, P. Devaraj and R. S. Bhuvaneshwaran, "An intertwining chaotic maps based image encryption scheme," *Nonlinear Dynamics*, vol. 69, p. 1995, 2012.
- [5] N. Pisarchik and M. Zanin, "Chaotic map cryptography and security", *Int. J. of Computer Research*, vol. 19, no. 1, 2012.
- [6] R. Lozi, I.Taralova "From chaos to randomness via geometric under-sampling", *ESAIM: Proceedings and surveys*, 2014, Vol. 46, pp. 177-195  
<http://dx.doi.org/10.1051/proc/201446015>
- [7] O. Garasym, I. Taralova, R. Lozi, "New nonlinear CPRNG based on tent and logistic maps", *Complex Systems and Networks - Dynamics, Controls and Applications*, 2015, Springer pp. 131-161.
- [8] O. Garasym, I.Taralova, R. Lozi, "Robust PRNG based on homogeneously distributed chaotic dynamics", *Journal of Physics: Conference Series* 692, 2016, 012001.  
<http://iopscience.iop.org/1742-6596/692/1>
- [9] S. El. Assad and M. Farajallah, "A new chaos-based image encryption system", *Signal processing: Image Comm.*, vol. 41, pp. 144, 2016.
- [10] D. Andreev, "Comparison of Clustering Algorithms through External Evaluation", *FDIBA, TU-Sofia*, November 2015.
- [11] K. Pavan, Allam Appa Rao, A.V. Dattatreya Rao, "An Automatic Clustering Technique for Optimal Clusters", Department of Computer Applications, Rayapati Venkata Ranga Rao and Jagarlamudi Chadramouli College of Engineering, Guntur, India, Jawaharlal Nehru Technological University, Kakinada, India, Department of Statistics, Acharya Nagarjuna University, Guntur, India, September 2011.
- [12] Thomas Lavergne, Tanguy Urvoy, Francois Yvon, "Detecting Fake Content with Relative Entropy Scoring", *PAN '08 Proceedings of the 2008 International Conference on Uncovering Plagiarism, Authorship and Social Software Misuse - Volume 377*, pp. 27-31.
- [13] Peter Christen, "A Comparison of Personal Name Matching: Techniques and Practical Issues", *Joint Computer Science Technical Report Series*, Department of Computer Science; Faculty of Engineering and Information Technology, Computer Sciences Laboratory; Research School of Information Sciences and Engineering, September 2006, TR-CS-06-02

# Lightweight Signcryption Scheme Based on Discrete Chebyshev Maps

Ta Thi Kim Hue<sup>1</sup>, Thang Manh Hoang<sup>1</sup>  
<sup>1</sup>School of Electronics and Telecommunications  
 Hanoi University of Science and Technology  
 1 Dai Co Viet, Hai Ba Trung, Hanoi, Vietnam  
 Email: hue.tathikim@hust.edu.vn

An Braeken<sup>2</sup>,  
<sup>2</sup>ETRO and INDI, Vrije Universiteit Brussel  
 Pleinlaan 2, B-1050 Brussel, Belgium  
 Email: an.braeken@vub.ac.be

**Abstract**—Signcryption schemes are cryptographic mechanisms providing both encryption and signing in a very efficient way. This paper presents a lightweight signcryption scheme based on Chebyshev chaotic maps over finite fields. For that, complex dynamic properties of the Chebyshev map are investigated and shown to be adequate for the construction of a signcryption scheme. The proposed signcryption scheme is proven to be secure for both outsider and insider attacks. Moreover, its computational cost is low, and the lower hardware complexity in compared with others based on elliptic curves.

## I. INTRODUCTION

Internet of Things (IoTs), also called “Smart Object network”, is a new computing environment and is becoming more and more popular. There are many constrained devices connected in this network via the internet. These devices are also called IoT devices and is typically lack of protection methods due to their constrained nature. There is the fact that lots of IoT devices are connected to information systems leading to an enormous amount of security vulnerabilities. Recently, scientific research has been focused on several security properties or services which should be provided by these devices, including confidentiality, integrity, authentication, authorization, and freshness in order to keep the IoT devices secure [1].

Lightweight cryptography is one of the security solutions that allow more efficient cryptographic implementations and thus is suitable for the typical hardware constraints of the IoT devices, including RFID tags, sensors, contact-less smart cards, health-care devices and so on. In other words, lightweight cryptographic solutions are suitable for the devices with restricted resources such as computational power, memory, storage space, and limited energy, while maintaining adequate performance [2]. Consequently, lightweight ciphers provide an adequate balance between security and computational performance. Several potential candidates have been announced by ECRYPT such as PRESENT [3], PHOTON [4], LBlock [5], LED [6], and TWINE [7], etc.

In 1997, Zheng introduced a new cryptographic primitive, called “Signcryption”. The scheme simultaneously realizes both a digital signature and an encryption in a logically single step, and it provides more efficient performance in compared with two separate processes, signing and encryption.

There have been many signcryption schemes proposed in recent years for different types of applications with the use of different types of underlying cryptographic operations to satisfy different security features.

*Related work* - The first signcryption scheme was introduced by Yuliang Zheng in 1997 [8] and is based on the Elgamal signature and encryption scheme. In 1998, Zheng [9] proposed an elliptic curve-based variant of this scheme, also called the elliptic curve signcryption scheme (ECSS). It shows that his scheme saves around 58% of computational and 40% of communication costs, in compared with the traditional approach of first signing and then encrypting.

Recently, various signcryption schemes have been introduced by Changji Wang et al. in [10], Youn and Hong in [11], Liu et al. [12] and others in [13], [14] and [15]. Unfortunately, those schemes are insecure and shown to be vulnerable for devastating attacks due to several security flaws. For example, the scheme proposed by Liu et al. [12] was totally broken by concrete attacks in [16] and subtle public key replacement attacks in [17]; the Youn et al.’s scheme [11] does not satisfy the existential unforgeability under chosen-message insider attacks as presented in [18]. Moreover, these schemes require costly pairing operations; those are not practical for constrained nodes in the IoT.

Chebyshev chaotic systems and their possible applications to cryptography such as Elgamal-like, RSA-like public-key encryption algorithms, and digital signature schemes are presented in [19]. Maze in [20] determined that Chebyshev polynomials in finite fields fulfill cryptographic requirements such as mixing property and difficulty of the underlying mathematical problem. Moreover, Chatterjee et al. in [21] presented a new authentication scheme for multi-server environments using the Chebyshev chaotic map suitable for devices with limited battery life and smaller computation power.

*Our work* - In this paper, we demonstrate that the Chebyshev map with chaotic properties is suitable for the use in both encryption and digital signature. A new construction of the signcryption scheme based on the discrete Chebyshev problem in finite fields is proposed. The proposed signcryption scheme satisfies the security features of confidentiality and authenticity. As a consequence, only the receiver of the signcrypted message is able to read the content and can verify the sender’s

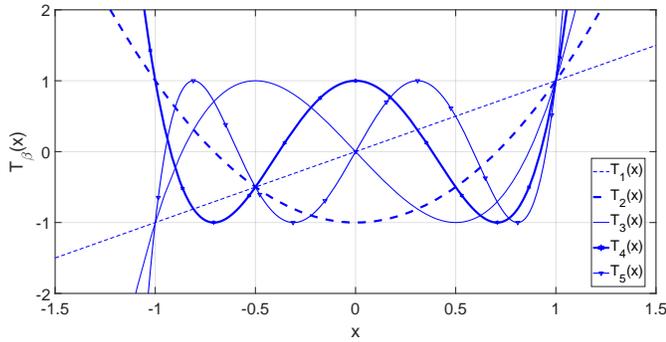


Fig. 1. Attractors of Chebyshev polynomials of the first kind with  $\beta = 1, 2, 3, 4$  and  $5$ .

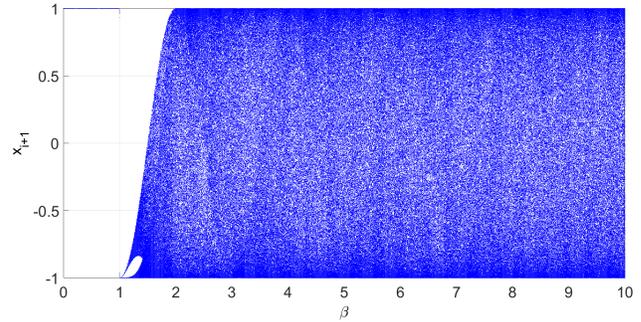


Fig. 2. Bifurcation diagram of Chebyshev map

identity by the intended entity. An attacker intercepting the signcryptured message cannot learn anything about the original message and will not be able to forge the message or claim to be the originator. The proposed scheme is lightweight, as it is shown that the Chebychev operation is slightly more efficient than an elliptic curve multiplication with respect to required hardware resources.

## II. CHEBYSHEV CHAOTIC MAP

### A. Properties of the Chebyshev map

In this section, the characteristics of the Chebyshev map on commutative finite rings  $\mathbb{R}$  are briefly presented. A wealth of information on these polynomials can be found in [20], [22] and [23].

The Chebyshev map is given as follows:

$$x_{i+1} = \cos(\beta \cdot \arccos(x_i)). \quad (1)$$

The following recurrent relation can be used to define Chebyshev polynomials of the first kind

$$T_{\beta}(x) = 2xT_{\beta-1}(x) - T_{\beta-2}(x), \quad (2)$$

which maps the interval  $[-1, 1]$  with  $\beta$  times onto itself in Fig.1.

The first few polynomials are explicitly listed below:

- $T_0(x) = 1;$
- $T_1(x) = x;$
- $T_2(x) = 2x^2 - 1;$
- $T_3(x) = 4x^3 - 3x;$
- $T_4(x) = 8x^4 - 8x^2 + 1;$
- $T_5(x) = 16x^5 - 20x^3 + 5x$  and so on.

Let  $\beta$  be a positive integer and  $x$  be a variable having a value over the interval  $[-1, 1]$ . The Chebyshev map satisfies the following important properties:

- The semi-group property:
  - 1)  $T_{nm}(x) = T_n(T_m(x)).$
  - 2)  $T_n(T_m(x)) = T_m(T_n(x)).$
  - 3)  $T_n(\frac{1}{2}(x + x^{-1})) = \frac{1}{2}(x^n + x^{-n}).$
- The chaotic property: As can be seen from the bifurcation diagram of the Chebyshev map in Fig.2, the dynamic shape is more complex when the parameter  $n$  is greater

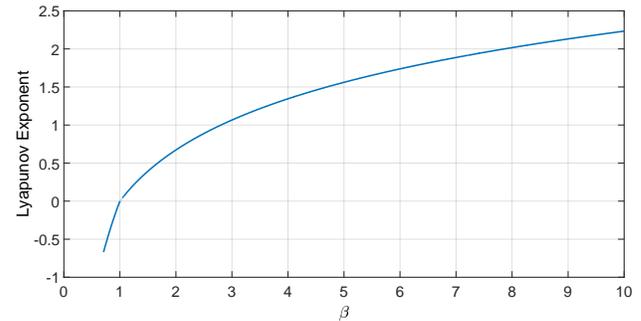


Fig. 3. Lyapunov exponent (LE) of Chebyshev map

than 2. In addition, the complexity of Chebyshev map is measured by the Lyapunov exponent (LE). A positive LE implies that the map is known to be chaotic [24]. The LE of the Chebyshev map is calculated by

$$\begin{aligned} LE &= \frac{1}{n} \ln \left| \frac{df^{(n)}(x_0)}{dx} \right| \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| f'(x_i) \right| \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| \frac{\beta}{\sqrt{1-x_i^2}} \sin(\beta \cdot \arccos(x_i)) \right|. \end{aligned} \quad (3)$$

According to Eq. (2), the LE of the Chebyshev map is shown in Fig.3. Obviously, the Chebyshev map is in chaotic state when  $\beta \geq 2$  as seen in Figs.2-3.

Furthermore, the experimental evaluation and performance analysis in [23] demonstrate that the Chebyshev chaotic map with attributes (e.g. chaos attractor and approximate entropy) has a good randomness and complex dynamic performance for using in the proposed cryptographic algorithms and other potential applications.

### B. Discrete Chebyshev problem in finite fields

To enhance the properties of the Chebyshev chaotic map, a modified Chebyshev polynomial is defined as in Eq. (4) on the domain  $\mathbb{R} = \mathbb{F}_q$ , the finite field with  $q = p^d$  elements.

The map  $T_g : \mathbb{F}_q \rightarrow \mathbb{F}_q$  is defined by

$$y = T_g(x) \text{ mod } p, \quad (4)$$

where  $x$  is an integer and  $p$  is a large prime number.

The following properties hold for the modified Chebyshev polynomials as

$$T_n(x)(T_m(x) \bmod p) \bmod p = T_{nm}(x) \bmod p. \quad (5)$$

The recurrence formula  $T_{g^s}(x)$  denote for  $s$  iteration times of  $T_g(x)$  as

$$T_{g^s}(x) \bmod p = \underbrace{T_g(T_g(T_g \dots T_g(x) \bmod p) \dots)}_{s \text{ iteration times}} \bmod p. \quad (6)$$

From Eq. (5), we have

$$T_{g^a}(T_{g^b}(x) \bmod p) \bmod p = T_{g^{(a+b)}}(x) \bmod p. \quad (7)$$

A strong connection between the discrete Chebyshev problem and the discrete logarithm (DL) problem in  $\mathbb{F}_q$  is proven in [20]. The discrete Chebyshev problem is expressed as a proposition. Let  $\mathbb{R}$  be a finite ring with unity. Given  $x$  and  $y$  such that  $y = T_l(x)$  for some  $l \in \mathbb{N}$ , find an integer  $n > 0$  such that  $y = T_n(x)$ . As a result, the discrete Chebyshev and the DL problems in  $\mathbb{F}_q$  are computationally equivalent.

It is very interesting that the Chebyshev polynomials can be used in the DL problem and the computational Diffie-Hellman (CDH) one as follows

- DL problem: Given two elements  $x$  and  $y$ , it is computationally infeasible to find an integer  $r$  such that the equation  $T_r(x) \equiv y \bmod p$  holds.
- CDH problem: Given  $x, T_r(x), T_s(x)$ , it is computationally infeasible to compute the result of the computations  $t_{rs} \equiv T_{rs}(x) \equiv T_r(T_s(x)) \equiv T_s(T_r(x)) \equiv T_{sr}(x) \equiv t_{sr}(\bmod p)$ .

### C. The computational complexity of the $T_g(x)$

Equation (2) gives

$$\begin{pmatrix} 0 & 1 \\ -1 & 2x \end{pmatrix} \begin{pmatrix} T_{g-2}(x) \\ T_{g-1}(x) \end{pmatrix} = \begin{pmatrix} T_{g-1}(x) \\ T_g(x) \end{pmatrix}, \quad (8)$$

and by induction

$$\begin{pmatrix} 0 & 1 \\ -1 & 2x \end{pmatrix}^{g-1} \begin{pmatrix} 1 \\ x \end{pmatrix} = \begin{pmatrix} T_{g-1}(x) \\ T_g(x) \end{pmatrix}. \quad (9)$$

It is obvious to see from Eq. (9) that the computational complexity of the  $T_g(x)$  is reduced to  $O(\log_2(g))$  [20] and the cost of the evaluation  $T_g(x)$  is 8 multiplications and 4 additions for the matrix multiplication.

## III. THE PROPOSED LIGHTWEIGHT SIGNCRYPTION SCHEME

The signcryption scheme makes parameters public to all users, which are summarized in Tab. I. The architecture of the proposed signcryption scheme is illustrated in Fig.4. The considered scenario contains the following processes

- Key agreement: Generate public/private pair of keys for two parties (Sender - Alice and Recipient - Bob), as given in Algorithm 1.
- Signature generation and encryption: Alice desires to send a message to Bob in a secure manner. Alice obtains

TABLE I  
SUMMARY OF NOTATIONS

$p$	Large prime.
$g$	Integer in $[1, \dots, p-1]$ with order $p-1$ modulo $p$ .
$m$	Length of the message.
$\alpha_a$	Alice's secret key.
$\alpha_b$	Bob's secret key.
$HashK$	Keyed one-way lightweight hash function, at least 128 bits.
$E$	Lightweight encryption algorithm.
$D$	Lightweight decryption algorithm.
ECC	Elliptic Curve Cryptography

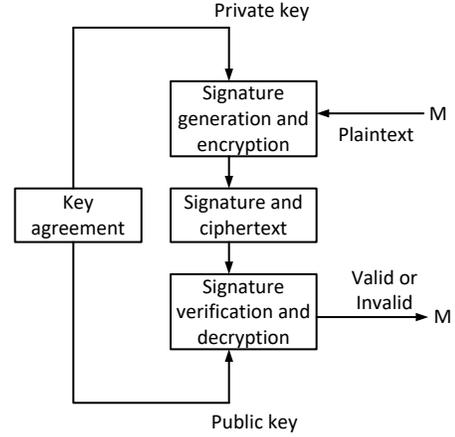


Fig. 4. Signcryption processes

the public key of Bob to signcrypt the message  $M$ , and return a signciphertext  $C$ . The signcryption scheme based on the Chebyshev map is detailed in Algorithm 2.

- Signature verification and decryption: Bob receives the signciphertext  $(r; s; C)$  from Alice to unsigncrypt and to recover the message  $M$ . Bob has to realize the procedure given in Algorithm 3.

### Algorithm 1 Key Generation Algorithm

- 1: **procedure** GENERATE THE KEYS
- 2: Generates two large integers  $\alpha_a, \alpha_b$ .
- 3: Alice has chosen a secret key  $\alpha_a$ . Similarly, Bob's secret key is  $\alpha_b$ .
- 4: Selects a random number  $x$  and compute both the public key of Alice  $T_{g^{\alpha_a}}(x) \bmod p$  and Bob  $T_{g^{\alpha_b}}(x) \bmod p$ .
- 5: Send the tuples  $(\alpha_a; x, T_{g^{\alpha_a}}(x)) \bmod p$ ; and  $(\alpha_b; x, T_{g^{\alpha_b}}(x)) \bmod p$  to Alice and Bob respectively.
- 6: Publish the public keys of Alice  $(x, T_{g^{\alpha_a}}(x))$  and Bob  $(x, T_{g^{\alpha_b}}(x))$ .
- 7: **end procedure**

## IV. SECURITY ANALYSIS

In this section, we give a formal security analysis of our proposal. Let us consider two classes of attacks to discuss the

**Algorithm 2** SignCryption Scheme

- 1: **procedure** SIGNCRYPT
- 2: Obtain Bob's authentic public key  $(x, T_{g^{\alpha_b}}(x) \bmod p)$ ;
- 3: Choose randomly  $\alpha$ , being a positive integer randomly from  $[1, \dots, p-1]$  with  $\alpha$  no divisor of  $(p-1)$ ;
- 4: Compute  $K = T_{g^{\alpha_a - \alpha}}(T_{g^{\alpha_b}}(x)) \bmod p$ ; we have  $K = T_{g^{\alpha_a + \alpha_b - \alpha}}(x) \bmod p$ ;
- 5: Split  $K$  into  $k_1$  and  $k_2$  of appropriate length;
- 6: Compute  $r = HashK_{k_1}(M)$  ;
- 7: Compute  $s = T_{g^{\alpha_a - \alpha - r}}(x) \bmod p$  ;
- 8: Obtain the ciphertext  $C = E_{k_2}(M)$  ;
- 9: Send  $(r; s; C)$  to Bob as Alice's signcryption output for the message  $M$  ;
- 10: **end procedure**

**Algorithm 3** UnsignCryption Scheme

- 1: **procedure** UNSIGNCRYPT
- 2: Recover  $K$  from  $r, s, g, p, \alpha_b$ ;
- 3: Compute  $K = T_{g^{\alpha_b + r}}(s) \bmod p$ ; we have  $K = T_{g^{\alpha_b + r}}(T_{g^{\alpha_a - \alpha - r}}(x) \bmod p) \bmod p$ ; From Eqs.(5) and (7), we obtain  $K = T_{g^{\alpha_a + \alpha_b - \alpha}}(x) \bmod p$ ;
- 4: Split  $K$  into  $k_1$  and  $k_2$  of appropriate length;
- 5:  $M = D_{k_2}(C)$ ;
- 6: The signature is valid if  $HashK_{k_1}(M) = r$ ;
- 7: Accept  $M$  as a valid message originated from Alice only if  $HashK_{k_1}(M)$  is identical to  $r$ .
- 8: **end procedure**

security for the signcryption scheme. The first class is called the outsider attacks in which an attacker accesses only the public information, including the public-keys of the involved entities. In the second class of attacks, also called insider attacks, an attacker has additionally access to the sender's private key.

**A. Outsider attacks**

The attacker is said to be successful if he succeeds in one of the two challenges. The first one is related to confidentiality, where he is able to distinguish different ciphertexts. The second one corresponds with the authenticity of the scheme, in which he achieves existential forgery [25].

Let us for example consider the situation where the attacker forges a valid signature for a given message  $M$ . We assume that the attacker only has access to public information  $pub = (x, r, s, C)$ . It is apparent that the security of the proposed scheme relies on the difficulty of finding  $(K, M)$ .

Because of the semi-group property, we have

$$\begin{aligned} T_n(x) &= T_n\left(\frac{1}{2}(x + x^{-1})\right) \\ &= \frac{1}{2}(x^n + x^{-n}) \\ &= y, \end{aligned} \quad (10)$$

and therefore  $x^n = y \pm \sqrt{y^2 - 1}$ . This requires at most two DLPS in  $\mathbb{F}_q(\sqrt{y^2 - 1})$ . From Algorithm 2, we have  $s =$

$T_{g^{\alpha_a - \alpha - r}}(x) \bmod p = T_{g^\Omega}(x) \bmod p$  with  $\Omega = \alpha_a - \alpha - r$ . It means that the following conditions are valid:

$$x^{g^\Omega} = s \pm \sqrt{s^2 - 1}. \quad (11)$$

Suppose that the attacker wants to find both  $n$  and  $\Omega$  such that  $T_n(x) = s$  and  $n = g^\Omega$ . The most powerful tool to solve the DLP in a finite field has an expected running time of  $O(\exp((\varsigma + O(1))(\ln(n))^{1/3}(\ln(\ln(n)))^{2/3}))$ , where  $\varsigma$  depends on the finite field ( $\varsigma \cong 1.92$  for a prime field) [20]. It means that it is infeasible to find  $\alpha_a$  from  $(x, r, s)$ . Consequently, it is impossible to forge the value of  $K$ .

**B. Insider attacks**

In this attack, a tuple  $(K, r, s)$  is computed from the sender's secret key, either  $\alpha_a$  or  $\alpha_b$ . When distinguishing ciphertexts, we give the attacker access to the sender's secret key,  $\alpha_a$ . That is, the attacker can easily find  $K = T_{g^{\alpha_a - \alpha}}(T_{g^{\alpha_b}}(x)) \bmod p$ , but he also has to recover  $(r, s)$ . Therefore,  $s$  can be recovered by the following equation:

$$\begin{aligned} s &= \left( T_{g^{\alpha_a - \alpha - HashK_{k_1}(D_{k_2}(C))}}(x) \right)^{g^{\alpha_b + HashK_{k_1}(D_{k_2}(C))}} \bmod p \\ &= K \pm \sqrt{K^2 - 1} \end{aligned} \quad (12)$$

However, it is infeasible to find  $k_1, k_2$  from  $s$ . The CDH problem indicates that this is too difficult to solve Eq. (12) and to find  $g^{\alpha_b + HashK_{k_1}(D_{k_2}(C))}$  in the finite field.

**V. THE COMPUTATIONAL COST**

The computational cost is evaluated in terms of the number of expensive operations needed for the signcryption and unsigncryption processes. This cost indicates how much computational effort has to be invested both by the sender and the recipient of the signcrypted message. Typically these operations include encryption and decryption, hashing, modulo addition, multiplication, division (inversion), and exponentiation. Therefore, the number of these operations figures out the computational cost in terms of hardware implementation [26].

We compare the computational complexity of the Chebyshev operation with the operations applied in ECC. Here, three point operations of ECC in the prime field need to be considered, Point Addition (PA) with  $R = P + Q$ , Point Doubling (PD) with  $R = 2P$  and Point Multiplication (PM) with  $R = kP$ . Obviously, multiplication can also be done by using single point doubling and addition [27], which requires only  $O(\log_2(k))$  steps. The inverse of that point multiplication, i.e. the value  $k$  from  $R$  and  $P$ , is known as the discrete logarithm problem. The cost of ECC operations over the prime field in terms of additions, squaring, inversion and multiplication is estimated in Tab. II.

We suppose that the operational times of addition, square, inversion and multiplication are  $t_1, t_2, t_3$  and  $t_4$ , respectively. Then, the operational time of a PM from Tab. II is calculated following:  $Time_{PM} = 12t_1 + 3t_2 + 2t_3 + 4t_4$ . It can be concluded from Eq. (9) that the number of operations to compute  $T_g(x)$  is equal to 8 multiplications and 4 additions, thus the operational time of  $T_g(x)$  is  $Time_{chev} = 4t_1 + 8t_4$ .

TABLE II  
THE COST OF ECC OPERATIONS [27]

ECC operations	Number of single operations			
	Addition	Square	Inversion	Multiplication
PA	8	1	1	2
PD	4	2	1	2
PM	12	3	2	4

From point of view of hardware implementation, we can consider as  $t_2 \approx t_3 \approx t_4$ ; it means that  $Time_{PM} \approx 12t_1 + 9t_4$ . Therefore, the operational time of  $T_g(x)$  is less than one PM. Similarly, the time of  $T_{gn}(x)$  is equal to  $nTime_{chev} = n(4t_1 + 8t_4)$ , see Eq. (6). A lightweight signcryption scheme in [26], called ECKSS+, is required two PMs for signcryption and two PMs and one PA for unsigncryption. Hence, the total operational time of ECKSS+ is  $Time_{ECKSS} = 4Time_{PM} + Time_{PA} \approx 56t_1 + 40t_4$ . In our proposed scheme, we have to take three  $T_{gn}(x)$  operations, so the operational time is given as  $T_{Chev-signc} = 3n(4t_1 + 8t_4)$ . Consequently, the computational complexity of our scheme can be considered similar to ECKSS+ in case of  $n \leq 3$ . By this way, the performance of our signcryption scheme, built up with Chebyshev polynomial operations, always outperforms in compared to similar schemes using underlying ECC operations.

## VI. CONCLUSION

In this study, we have designed a new signcryption which is based on the discrete Chebyshev map in the prime field, and on exploiting the complex dynamic properties of the Chebyshev map. Our scheme is proven to resist against two types of security attacks, depending on whether the adversary is an insider or outsider. In addition, the computational complexity of the proposed scheme outperforms in compared to the schemes based on ECC operations.

To conclude, the proposed signcryption scheme is secure and compact, making it in particularly suitable for smart card based applications, useful in a variety of areas including digital cash payment systems, personal health systems based lightweight electronic transaction protocols.

## ACKNOWLEDGMENT

This research is funded by the Hanoi University of Science and Technology (HUST) under project number T2016-LN-11.

## REFERENCES

[1] K. T. Nguyen, M. Laurent, and N. Oualha, "Survey on secure communication protocols for the internet of things," *Ad Hoc Networks*, vol. 32, pp. 17 – 31, 2015. Internet of Things security and privacy: design methods and optimization.

[2] T. Eisenbarth, S. Kumar, C. Paar, A. Poschmann, and L. Uhsadel, "A survey of lightweight-cryptography implementations," *IEEE Design Test of Computers*, vol. 24, pp. 522–533, Nov 2007.

[3] A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. Robshaw, Y. Seurin, and C. Viskelsoe, "Present: An ultra-lightweight block cipher," in *International Workshop on Cryptographic Hardware and Embedded Systems*, pp. 450–466, Springer, 2007.

[4] J. Guo, T. Peyrin, and A. Poschmann, "The photon family of lightweight hash functions," in *Annual Cryptology Conference*, pp. 222–239, Springer, 2011.

[5] A. Poschmann, G. Leander, K. Schramm, and C. Paar, "New light-weight crypto algorithms for rfid," in *2007 IEEE International Symposium on Circuits and Systems*, pp. 1843–1846, IEEE, 2007.

[6] J. Guo, T. Peyrin, A. Poschmann, and M. Robshaw, "The led block cipher," in *International Workshop on Cryptographic Hardware and Embedded Systems*, pp. 326–341, Springer, 2011.

[7] T. Suzaki, K. Minematsu, S. Morioka, and E. Kobayashi, "Twine: A lightweight block cipher for multiple platforms," in *International Conference on Selected Areas in Cryptography*, pp. 339–354, Springer, 2012.

[8] Y. Zheng, *Digital signcryption or how to achieve cost(signature & encryption)  $\approx$  cost(signature) + cost(encryption)*, pp. 165–179. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997.

[9] Y. Zheng, *Signcryption and its applications in efficient public key solutions*, pp. 291–312. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998.

[10] C. Wang and J. Liu, *Attribute-Based Ring Signcryption Scheme and Its Application in Wireless Body Area Networks*, pp. 521–530. Cham: Springer International Publishing, 2015.

[11] T.-Y. Youn and D. Hong, "Signcryption with fast online signing and short signciphertext for secure and private mobile communication," *Science China Information Sciences*, vol. 55, pp. 2530–2541, Nov 2012.

[12] Z. Liu, Y. Hu, X. Zhang, and H. Ma, "Certificateless signcryption scheme in the standard model," *Information Sciences*, vol. 180, no. 3, pp. 452 – 464, 2010.

[13] T. Matsuda, K. Matsuura, and J. C. N. Schuldt, *Efficient Constructions of Signcryption Schemes and Signcryption Composability*, pp. 321–342. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[14] G. Wang, R. H. Deng, D. Kwak, and S. Moon, *Security Analysis of Two Signcryption Schemes*, pp. 123–133. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[15] R. Steinfeld and Y. Zheng, *A Signcryption Scheme Based on Integer Factorization*, pp. 308–322. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000.

[16] J. Weng, G. Yao, R. H. Deng, M.-R. Chen, and X. Li, "Cryptanalysis of a certificateless signcryption scheme in the standard model," *Information Sciences*, vol. 181, no. 3, pp. 661 – 667, 2011.

[17] S. Miao, F. Zhang, S. Li, and Y. Mu, "On security of a certificateless signcryption scheme," *Information Sciences*, vol. 232, pp. 475 – 481, 2013.

[18] D. Zhou, J. Weng, C. Guan, R. Deng, M. Chen, and K. Chen, "Cryptanalysis of a signcryption scheme with fast online signing and short signciphertext," *Science China Information Sciences*, vol. 57, pp. 1–5, Jul 2014.

[19] L. Kocarev and S. Lian, *Chaos-based Cryptography: Theory, Algorithms and Applications*. Springer Publishing Company, Incorporated, 1st ed., 2011.

[20] G. Maze, *Algebraic methods for constructing one-way trapdoor functions*. PhD thesis, University of Notre Dame Notre Dame, 2003.

[21] S. Chatterjee, S. Roy, A. K. Das, S. Chattopadhyay, N. Kumar, and A. V. Vasilakos, "Secure biometric-based authentication scheme using chebyshev chaotic map for multi-server environment," *IEEE Transactions on Dependable and Secure Computing*, 2016.

[22] R. Adler and T. Rivlin, "Ergodic and mixing properties of chebyshev polynomials," *Proceedings of the American Mathematical Society*, vol. 15, no. 5, pp. 794–796, 1964.

[23] Y. Liu, Y. Luo, S. Song, L. Cao, J. Liu, and J. Harkin, "Counteracting dynamical degradation of digital chaotic chebyshev map via perturbation," *International Journal of Bifurcation and Chaos*, vol. 27, no. 03, p. 1750033, 2017.

[24] L. Kocarev, "Chaos-based cryptography: a brief overview," *IEEE Circuits and Systems Magazine*, vol. 1, no. 3, pp. 6–21, 2001.

[25] K. Chain and W.-C. Kuo, "A new digital signature scheme based on chaotic maps," *Nonlinear Dynamics*, vol. 74, pp. 1003–1012, Dec 2013.

[26] K. T. NGUYEN, *Lightweight Security Protocols for IP-based Wireless Sensor Networks and the Internet of Things*. PhD thesis, Information, Telecommunications and Electronics of Paris, 2016.

[27] D. Hankerson, A. J. Menezes, and S. Vanstone, *Guide to elliptic curve cryptography*. Springer Science & Business Media, 2006.

## **Workshop 2: The 8th International Workshop on Cloud Applications and Security (CAS'17)**

Title: Forward Secure Searchable Symmetric Encryption  
(Authors: Muhammad Saqib Niaz, Gunter Saake)

Title: CPU Workload forecasting of Machines in Data Centers using LSTM Recurrent Neural Networks and ARIMA Models  
(Authors: Deepak Janardhanan, Enda Barrett)

Title: An Advanced Reinforcement Learning Approach for Energy-Aware Virtual Machine Consolidation in Cloud Data Centers  
(Authors: Rachael Shaw, Enda Howley, Enda Barrett)

Title: Predicting Host CPU Utilization in Cloud Computing using Recurrent Neural Networks  
(Authors: Martin Duggan, Karl Mason, Jim Duggan, Enda Howley, Enda Barrett)

# Forward Secure Searchable Symmetric Encryption

Muhammad Saqib Niaz

Databases & Software Engineering  
Otto von Guericke University  
Magdeburg, Germany  
saqib.niaz@ovgu.de

Gunter Saake

Databases & Software Engineering  
Otto von Guericke University  
Magdeburg, Germany  
gunter.saake@ovgu.de

**Abstract**—Data outsourcing to third party clouds poses numerous data security threats. Access by unauthorized users is one of the security threat to the outsourced data. Unauthorized access can be avoided by encrypting the data before outsourcing. However, encrypting data before outsourcing renders it unsearchable to the data owner. Searchable encryption schemes are developed to specifically target this problem. A dynamic searchable encryption is the one that allows the data owner to add or delete a file after data outsourcing. Dynamic searchable encryption schemes are vulnerable to two specific security threats that are not applicable to the static searchable encryption schemes namely forward privacy and backward privacy. Forward privacy requires that the addition of a file should not reveal the presence of a previously searched keyword. Backward privacy requires that a search should not return the file identifier of a previously deleted file. In this paper, we propose a dynamic searchable scheme that guarantees forward privacy. It only uses the symmetric key algorithms hence reducing the requirements for storage and processing power on the client side. Furthermore, our proposed scheme is space reclaiming. After the deletion of a file, the redundant data nodes are also deleted from the secure index in the subsequent searches. Because of this space reclaiming capability of the scheme, the scheme is also partially backward private.

**Keywords**-component; *Dynamic Searchable Symmetric Encryption; Dynamic Searchable Encryption; Searchable Symmetric Encryption; Forward Privacy; Backward Privacy; Space Reclaiming*

## I. INTRODUCTION

Data outsourcing to the third-party clouds has seen a tremendous growth both in the personal and business setting. Despite the fact that the use of the third-party cloud storage is growing incredibly, one cannot deny the security threats involved in using a third-party cloud storage. One of the security threats is the unauthorized access by the malicious users. If the user data is stored without any encryption, the cloud administrators can simply access the data. Furthermore, in the case of a hack, the hackers could also access the user's plaintext data. In order to curb this problem, one simple solution is to encrypt all the files and then store them on a third-party cloud. However, encrypting the files with the conventional encryption schemes renders the data unsearchable. In order to search in such a setting, the user would have to download all the encrypted files to the local storage, decrypt all the files and then run the search query. This is infeasible and impractical as it eliminates almost all the benefits of a cloud storage.

Searchable encryption (SE) gives a user the ability to run a search query without decrypting the file contents. Basic idea behind an SE is to extract the keywords from the plaintext files. Based on the extracted keywords, create an index. Encrypt the files along with the associated index and send both to the cloud storage. Later on, when the user wants to search a keyword, he creates a trapdoor based on the search keyword so that the server does not learn the keyword being searched. This trapdoor is then sent to the server; server searches the trapdoor in the secure index and returns the corresponding encrypted files.

Initially proposed SE schemes were static in nature. Once the secure index was generated, user was unable to add or delete a file without regenerating the whole index. Later on, dynamic searchable encryption (DSE) schemes are developed that gives the data owner the ability to add or delete a file after creation of the secure index. Update of a file in these schemes does not require the regeneration of the secure index. There are some established security criteria for SEs. These criteria are also applicable to DSEs. Furthermore, two new security criteria namely forward privacy and backward privacy are introduced that are only applicable to DSEs [1]. Forward privacy requires that the addition of a new file should not reveal the presence of an already searched keyword. Backward privacy requires that a search should not return any file identifier of a previously deleted file.

In this paper, we present a forward secure searchable symmetric encryption abbreviated as FSE. The basic idea of how to make a forward privacy scheme is inspired from the [2]'s scheme. The idea is to insert the new nodes in a way that the server is unable to relate the newly inserted nodes with any of the already searched keywords. Later on, during the search, the server should be able to retrieve all the files corresponding to the search keyword including the newly inserted nodes. Our scheme presents numerous improvements over [2]'s scheme. Our scheme does not use any public key algorithms both on the server and the client-side. As the scheme only uses the symmetric key cryptography therefore the client-side storage and processing requirements are minimized. Our proposed scheme is also space reclaiming. In case of a file deletion, the redundant data nodes are deleted from the secure index in the subsequent searches. The space reclaiming feature does not require regeneration of the whole secure index. Due to the space reclaiming feature, our scheme partially fulfills the criteria of backward privacy.

## II. RELATED WORK

First searchable encryption scheme was introduced in 2000 [3]. Every word of each document was separately encrypted hence making the scheme vulnerable to various statistical attacks. Encrypting every word separately also makes the files incompressible.

Goh proposed a bloom filters based scheme [4] but it presented weaker security guarantees. By design, bloom filters produce false positives which is considered a weakness in a searchable encryption scheme. The author also presented the first security definitions for a searchable scheme.

Chang et al. proposed a scheme in 2005 [5]. This scheme required generation of one index per document in the document collection. The main drawback of the scheme was the search overhead on the server side. Search algorithm had to go through each index to search a single keyword.

Curtmola et al. proposed an inverted index based scheme [6]. The authors achieved a sub-linear searching efficiency by storing all the file identifiers related to a keyword in a linked list. The authors also presented new security definitions for SEs.

Van Liesdonk et al. proposed a searchable encryption scheme based on public key encryption [7]. It was a multi-writer and single-reader scheme. This scheme allowed wildcard searches over any alphabet.

Kamara et al. proposed a scheme in 2010 [8]. It was an improved version of [6] and it was a static scheme. Kamara et al. proposed a dynamic scheme in 2012 [9]. This scheme had a sub-linear search complexity. It leaked the hashes of the keywords contained in the updated documents. Kamara et al. proposed an improvement in 2013 [10] resolving the hash leaking problem at the cost of extra storage at server-side.

Schemes proposed by Goh [4] and Van Liesdonk et al. [7] were dynamic in nature but Kamara et al. were the first one to propose a DSE with sub-linear search complexity [9].

Cash et al. introduced a scheme that was capable of running boolean queries on outsourced data [11]. Later on, they extended this work and proposed a DSE claiming to handle very large databases [12].

Yavuz et al. presented a DSE with efficient update complexity [13]. However, the search speed of the scheme came down to linear from sub-linear. Search algorithm has to process nodes equal to the total number of documents in the file database instead of only processing the nodes equal to the frequency of the search keyword.

Islam et al. [14] and Cash et al. [15] presented new attacks that become possible due to the leakage in SEs.

Stefanov et al. presented new security definitions for DSEs [1]. The authors also proposed a scheme that was forward privacy secure but the scheme suffered from computational and storage inefficiencies.

In 2016, Bost presented a scheme [2] that was also forward privacy secure. Bost's scheme offered some improvements over the [1]'s scheme.

## III. SECURITY CRITERIA FOR DSEs

Goh was the first one to introduce the security criteria for a searchable encryption scheme in 2003 [4]. According to his security definition, an adversary should not be able to deduce the document contents from the secure index. However, the security criteria do not require any security for the search trapdoors. The security definition presented two security criteria i.e. IND1-CKA and IND2-CKA. IND1-CKA requires that the indices of the equal sized documents should be indistinguishable to an adversary. IND2-CKA requires that the indices of the unequal sized documents should be indistinguishable to an adversary. These security criteria specifically deal with the schemes that produced one index per document.

Curtmola et al. proposed the new and improved security definitions for the searchable encryption schemes in 2006 [6]. The security definitions require that the remotely stored index and files should not leak any information to an adversary. Additionally, these definitions also require the security of the trapdoors. They presented two security criteria i.e. IND-CKA1 and IND-CKA2. IND-CKA1 requires the security against a non-adaptive attacker. Non-adaptive adversary is the one who generates all the queries at the beginning of the protocol. IND-CKA2 requires the security against an adaptive attacker. An adaptive adversary is the one who generates the subsequent queries dynamically based on the results of the previous queries.

Shi et al. proposed the two new security criteria especially for dynamic searchable encryption schemes in 2014 [1]. These security criteria are named as forward privacy and backward privacy. Forward privacy requires that the addition of a new file should not reveal the presence of an already searched keyword. Backward privacy requires that after a deletion operation, a search query should not return the file identifiers of the already deleted files.

## IV. FSE CONSTRUCTION

FSE is a forward secure dynamic searchable symmetric encryption scheme. As the name describes it, our scheme is a forward secure scheme. In case of a file addition operation, the server is unable to link the nodes of the newly inserted file with any of the already searched keywords. It is also a space reclaiming scheme. After a file deletion operation, the redundant nodes are removed from the secure index in the subsequent searches. Due to the space reclaiming capability of the FSE, it is also a partially backward private scheme. Search complexity of the FSE is sub-linear. Sub-linear search means that the number of nodes processed by the server search algorithm is equal to the frequency of the keyword being searched. Addition and deletion operation for a file is executed in a single step on server side. FSE only uses the symmetric key cryptosystem and a keyed hash function. No public key crypto-algorithm is used at any stage on the client or the server side.

### A. General Ideas

Generally, in the searchable encryption schemes, the nodes related to one keyword are linked. In search operation, the server receives a trapdoor and its associated decryption key from the client. The search algorithm at the server-side goes to the first node, decrypts it, gets the address of the next node and a file

identifier. The search algorithm keeps on decrypting and collecting the file identifiers until the end of the list is reached. After completion of the search algorithm, the server knows all the nodes/file identifiers related to a trapdoor. Now, if the client performs a file addition operation. Client sends the encrypted file along with its associated trapdoors to the server. Server inserts those nodes at specific locations as dictated by the client's addition algorithm. At this point, if the server is able to relate the new nodes with the previously searched keyword/trapdoor lists then the scheme is not forward secure. In contrast, if server is unable to relate the newly added nodes to any of the already searched keyword/trapdoor lists then the scheme is forward secure.

Our proposed scheme is a forward privacy secure searchable encryption scheme. The technique is to generate the addition token in a way that the server cannot relate the newly added nodes to the list of nodes of an already searched keyword/trapdoor. But at a later time when that keyword is searched again, the server gets a new trapdoor and a new decryption key for the same keyword and at that point server is able to get all the related file identifiers to the searched keyword.

In FSE, the addition token contains a new node and an address. The server addition algorithm simply stores the new node data at the associated address. The node data contains the file identifier of the newly added file, the encryption key to the next node in the list and the address of the next node, here the next node refers to the first node for the keyword in question. Every new node is inserted before the first node of the link list for that specific keyword. The contents of the node data are encrypted with a newly generated secret key that stays private at the client-side. At the time of the search, the client retrieves the address and the secret key associated with the search keyword. Client sends this address and encryption key to the server, server goes to the address and decrypts the node to get the file identifier and from the node data server also gets the address and the decryption key of the next node in the list. In this way server navigates through the whole list. But after the addition and before a search, it is not possible for the server to know the linkage of the newly added node to any of the old lists of nodes for a keyword/trapdoor.

### B. Preliminaries

File identifiers and the keywords are considered to be strings. Let  $\lambda$  be the security parameter.  $\text{Poly}(\lambda)$  and  $\text{negl}(\lambda)$  denote the polynomial and the negligible function in  $\lambda$ . The set of all binary strings of length  $\lambda$  is denoted as  $\{0,1\}^\lambda$ , and the set of all finite binary strings as  $\{0,1\}^*$ .  $x||y$  refers to the concatenation of the two strings  $x$  and  $y$ .  $x \leftarrow X$  means that the  $x$  is uniformly sampled from a finite set  $X$ . A *database*  $DB = (ind_i, W_i)_{i=1}^D$  is a set of file-identifier/keyword pair with  $ind_i \in \{0,1\}^\lambda$  and  $W_i \subseteq \{0,1\}^l$ . The set of keywords in the whole database can be represented by  $W = \cup_{i=1}^D W_i$ .

Our searchable encryption scheme is comprised of following algorithms:

- Setup is an algorithm that takes the parameter  $\lambda$ . It outputs a secret key of length  $\lambda$ , an encrypted database EDB and  $\sigma$  the client's state.
- AddFile is an algorithm that runs on the client-side. It takes as input the client state  $\sigma$ , the client secret key  $K_s$  and the new file to be inserted in the database. It updates the client state and outputs the encrypted file and the corresponding addition tokens.
- AddToken is an algorithm that runs on the server-side. It takes input the EDB, the encrypted file and the addition tokens. It outputs the updated EDB.
- DeleteFile is an algorithm that runs on the client-side. It takes as input the file identifier of the file to be deleted. It outputs a deletion token.
- DelToken is an algorithm that runs on the server-side. It takes input the EDB and the deletion token. Deletes the file and outputs the updated EDB.
- Trapdoor is an algorithm that runs on the client-side. It takes the keyword  $w$  and the client state  $\sigma$  as input. If the keyword exists in the client state, it outputs the address and the encryption key for the keyword.
- Search is an algorithm that runs on the server-side. It takes the encrypted database EDB and the trapdoor (address & encryption) for the search keyword as the input. It outputs the file identifiers of the files matching the trapdoor's criteria.

### C. Basic Construction

Basic construction of FSE is forward private secure but it is not space reclaiming. After a file deletion operation, the extra node data corresponding to deleted file stays in the secure index.

```

Setup
 $K_s \xleftarrow{\$} \{0,1\}^\lambda$ 
 $W, T \leftarrow \text{empty map}$ 
return  $(K_s, T, W)$ 
T for server,  $K_s$  & W for client
    
```

#### 1) Setup

Setup of the FSE is straightforward. A secret key is generated. This key always stays at the client-side. This key is used for encryption/decryption of the files. Two empty maps are generated. T can be any data structure that provides a functionality of mapping a key to a value. A key is associated with every node in the secure index. We call this key the 'address'. Each node in this T is inserted at a specific address. Later on, in search algorithm, each node is accessed using its corresponding address. Client-side map W is a structure that stores data based on the keywords. Each element of this map can be accessed using a keyword. For each keyword, two values are stored in the W array, first is the address (in the T map on server-side) of the first node containing this keyword and second is the corresponding decryption key.

```

AddFile:
Extract keywords from the file
Encrypt the file with  $K_s$ 
Do following for each keyword:
Generate  $K_w \leftarrow \{0,1\}^\lambda$ 
Generate address  $Add_w \leftarrow H(K_s, F_{identifier} || W)$ 
Search  $w$  in  $W$  array
If value found:
    Create node containing  $F_{identifier}, Add_{w-1}, K_{w-1}$ 
endif
If value not found
    Create node containing  $F_{identifier}, null, null$ 
endif
Put  $K_w$  and  $Add_w$  in  $W$  array at location  $w$ 
Encrypt the node data with  $K_w$ 
Output the encrypted node data and  $Add_w$ 
    
```

### 2) AddFile

AddFile algorithm extracts the keywords from the file being added. It encrypts the file with client-side secret key. For each extracted keyword, it checks the existence of that keyword in client-side  $W$  array. If the keyword does not already exist in client state, it generates a random key for the keyword. It generates an address using a keyed hash function based on the file identifier and the keyword. The algorithm creates a data node with an empty next node pointer and encrypts this node data with the newly generated random key. It creates a data node for the client state containing the address for the server-side map and the encryption key for the node data. If the keyword already exists in the client state, then the AddFile algorithm creates the new node for an already existing keyword in such a way that server is unable to link the newly added data nodes with any of the already searched lists of nodes. AddFile algorithm gets the address and the encryption key of the previously stored first node of the list. It creates a new node with the new file identifier, puts the address and the encryption key of the already existing first node. Encrypts this node with a new randomly generated key. Keeps the new secret key with client and sends the encrypted node data along with the newly calculated address where server is supposed to store this data. In this way, link is created but the link information is kept secret from the server. Server is unable to link this newly generated node with any of the already existing lists of the searched keywords.

```

AddToken:
Goto location  $Add_w$  and store encrypted node data
    
```

### 3) AddToken

AddToken algorithm runs on the server-side. It receives an encrypted file and an address-node pair for each keyword contained in that file. Server stores the encrypted file in file storage. For each address-node pair, AddToken algorithm goes to the specific address and stores the node data on this location.

```

DeleteFile:
Outputs  $F_{identifier}$  (to be sent to server)
    
```

```

DelToken:
Server deletes the corresponding file to  $F_{identifier}$ 
    
```

### 4) DeleteFile & DelToken

File deletion is straight forward. Client sends the file identifier of the file to the server. Server simply deletes the encrypted file from the file storage. This does not affect the secure index.

```

Trapdoor:
 $K_w, Add_w \leftarrow W[w]$ 
if  $(K_w, Add_w)$  is null
    return
Output  $K_w$  and  $Add_w$  (to be sent to server)
    
```

### 5) Trapdoor

Trapdoor algorithm runs on the client side. It takes the search keyword as input. Searches the client-side  $W$  array, if the keyword is found it outputs the address and the corresponding decryption key.

```

Search:
goto  $Add_w$ 
decrypt data with  $K_w$ 
add  $F_{identifier}$  to  $list_{ident}$ 
 $Add_{next} \leftarrow Add_{w-1}$ 
 $Key_{next} \leftarrow K_{w-1}$ 
while  $(Add_{next} \neq null)$ 
    goto  $Add_{next}$ 
    decrypt data node with  $K_{next}$ 
    add  $F_{identifier}$  to  $list_{ident}$ 
     $Add_{next} \leftarrow Add_{w-1}$ 
     $Key_{next} \leftarrow K_{w-1}$ 
end while
Output files corresponding to  $list_{ident}$ 
    
```

### 6) Search

Search algorithm in the basic construction is simple as it does not care about the deleted files. That's why the basic construction is not space reclaiming. Search algorithm runs on server-side, it receives the address and the secret key for the search keyword. The search algorithm simply goes to the address and decrypts the data with the secret key. Upon every decryption, the server gets a file identifier, an address of the next node and the secret key for the next node. At the end of the algorithm a list of identifiers of the files is generated that contains the keyword being searched. Server sends back all the existing files corresponding to that list of identifiers.

### D. Space reclaiming construction

This is an enhanced version of the basic construction. The only difference between the two constructions is the search algorithm. If a file is deleted from the server, the index stays unchanged. Later on, when a keyword is searched for which a file has already been deleted, server removes the deleted nodes from the secure index. It does not matter how many nodes for a

specific keyword are deleted. On the first subsequent search after the deletions, all the extra nodes are removed from the server index.

```

Trapdoor:
 $K_w, Add_w \leftarrow W[w]$ 
if ( $K_w, Add_w$ ) is null
  return
Output  $K_w$  and  $Add_w$ 

```

### 1) Trapdoor

Trapdoor algorithm of enhanced version is identical to the basic version. It takes a keyword and if the keyword exists in the secure index, it outputs the address and the secret key for the first node.

```

Search:
create empty listident
goto Addw
decrypt with  $K_w$ 
extract  $F_{ident-w}$ 
check file existence for  $F_{ident-w}$ 
if  $F_{ident-w}$  doesn't exist
  keep decrypting next nodes until valid  $F_{ident}$  is found
  encrypt all value of this node with  $K_w$ 
  store this newly encrypted node at  $Add_w$ 
   $F_{ident-w} \leftarrow F_{ident}$ 
  if no node with valid  $F_{ident}$  is found
    delete first node from  $Add_w$ 
    return null to client
  endif
end
add  $F_{ident-w}$  to listident
goto next address and start decrypting node data
if file exists add to listident
if encounters a non – existent file
  keep decrypting nodes until a valid node found
  update last node with valid identifier
  only update next node address and encrypt again
  if no node found with valid  $F_{ident}$ 
    update last node with valid  $F_{ident}$ 
    replace next node address with null
    encrypt again with node key
  endif
endif
return files from listident

```

### 2) Search

In the space reclaiming version, the search algorithm at the server-side starts decrypting the linked nodes and while decrypting the nodes and collecting the file identifiers, it also checks the existence of the corresponding files. If the server gets such a file identifier from a decrypted node for which the corresponding file does not exist, the search algorithm deletes that node from the list and reconnects the list again. This search and reclaim space algorithm can encounter three kind of file

deletion scenarios. First possibility is that the first file in the list is deleted. In this case, the server would have to find such a node further in the list for which the corresponding file is still existing in the server file database. Search algorithm would take all the data from this valid node including the next node address and its corresponding encryption key. It would encrypt this node's data with the key that server has received from the client for decryption of first node in the list. It would store this newly encrypted node on the address that it has received from the client i.e. address of the first node as stored in the client's state. By doing so, no change is required on the client side. The initial address and the corresponding encryption key already stored at the client side remains valid even if the first node in the list is deleted. Second possible scenario is that the node stored at the initial address sent by the client has an existing file in the file database but some other files from the list are missing. In this case, the search algorithm would delete the nodes for which the corresponding files are deleted and would relink the list. It would update the last node in the list with a valid file identifier and would have to replace its next node address and its corresponding encryption key with the address and key of the next node in the list that contains a valid file identifier. The search algorithm already has the secret key for each node that makes it possible for it to update the node data and re-encrypt it again. Third possibility is a less likely scenario in which case all the files corresponding to a keyword are deleted. In this case, the server would delete all nodes from the list and would send a null response to the client.

```

Search (Client End):
Client gets the encrypted files
if client gets a null response
  client deletes the corresponding data from W array
endif

```

### 3) Search (Client-side)

Upon receiving a null response from the server, the client would have to update its state. Client would remove the data stored in its W array corresponding to the search keyword.

### E. Storage

Storage on the client side is proportionate to the number of keywords. For each keyword, one address and one encryption key are stored. An additional secret key is also stored on the client side. This key is used to encrypt the files. It is also used in the keyed hash function for the address generation for a new keyword & file-identifier combination.

On server side, one encrypted data node is stored for each keyword & file-identifier combination. Each node contains three elements, a file-identifier, an address of the next node and an encryption key of the next node.

The basic version does not remove the extra nodes for the deleted files from the secure index. In space reclaiming version, the extra nodes are removed from the secure index. Removal of extra nodes is not done at the deletion of files. Extra nodes are removed from the secure index in subsequent searches.

### F. Performance

Our proposed scheme is a non-interactive scheme, as all the operations are performed in one round.

Our scheme only uses symmetric key algorithms and a keyed hash function at the client side. On the server side, only symmetric key algorithm is used in the search algorithm. By avoiding the public key cryptosystem especially at the client side, the scheme is easily deployable on low-resourced (both in processing and memory) mobile devices.

Setup is simple and straight-forward, it does not require any cryptographic calculations.

Addition of a new file requires extraction of the keywords at the client side. One keyed hash and one encryption for each keyword is done. On the sever side, the addition algorithm works in  $O(1)$ . It requires storing of the nodes' data at specific addresses.

Deletion of a file is straight forward. It does not require any processing on the client side. Client just sends the file identifier to the server for deletion. On the server side, the deletion also works in  $O(1)$ . Server just deletes the corresponding file from the file storage. No change is required in the secure index.

Search is sub-linear, the search algorithm on the server-side only needs to process the nodes equal to the frequency of the search keyword. But there could be an exception to this rule, if a file has been deleted and then a search runs first time after the deletion then the search algorithm would have to process the one extra node along with the valid nodes. But after this search algorithm execution, the extra node for the deleted file is removed from the secure index. Later on, in all subsequent searches, the search algorithm would run in the sub-linear time complexity.

As our search algorithm removes the deleted nodes during the search that's why our scheme does not require to maintain a separate deleted file index. Maintaining a separate deletion index also entails regeneration of the whole index at regular intervals to merge the deletion index into the main index.

### G. Security

Our scheme is forward secure and partially backward secure. Forward security is achieved by generation of the insertion nodes in a way that the server is unable to link the newly inserted nodes with any of the already searched keywords. The scheme is partially backward secure as the search of a keyword after a deletion gives out the deleted file identifier but during this search, server deletes the extra nodes and relinks the list again. In all subsequent searches, the search algorithm does not get any deleted file identifiers. Our scheme is indistinguishable against the chosen keyword attack i.e. IND1-CKA & IND2-CKA as outlined in [4]. Our scheme is also indistinguishable against non-adaptive attackers i.e. IND-CKA1 as outlined in [6].

## V. CONCLUSIONS

FSE fulfills the criteria of the forward privacy. The scheme is space reclaiming in case of file deletions. Space reclaiming feature makes it partially backward secure. Search is sub-linear. Addition and deletion of a file requires a single step execution

on the server side. Our scheme does not use any public key cryptosystem algorithms thus making it feasible to be used in the resource restrained environments.

## VI. FUTURE WORK

Security of the scheme needs to be outlined in detail. Detailed analysis is required to check whether the proposed scheme is indistinguishable against an adaptive attacker i.e. IND-CKA2 as proposed in [6]. Implementation of the scheme is also required to check the practicality of the design in different deployment scenarios.

## REFERENCES

- [1] E. Stefanov, C. Papamanthou and E. Shi, "Practical Dynamic Searchable Encryption with Small Leakage," in *21st Annual Network and Distributed System Security Symposium*, San Diego, 2014.
- [2] R. Bost, "Forward Secure Searchable Encryption," in *Conference on Computer and Communications Security*, Vienna, 2016.
- [3] D. X. Song, D. Wagner and A. Perrig, "Practical techniques for searches on encrypted data," in *IEEE Symposium on Security and Privacy*, 2000.
- [4] E.-J. Goh, "Secure Indexes," *IACR Cryptology ePrint Archive*, vol. 2003, p. 216, 2003.
- [5] Y.-C. Chang and M. Mitzenmacher, "Privacy Preserving Keyword Searches on Remote Encrypted Data," in *Applied Cryptography and Network Security*, New York, 2005.
- [6] R. Curtmola, J. Garay, S. Kamara and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," in *13th ACM Conference on Computer and Communications Security*, 2006.
- [7] P. Van Liesdonk, S. Sedghi, J. Doumen, P. H. Hartel and W. Jonker, "Computationally Efficient Searchable Symmetric Encryption," *Secure data management*, vol. 6358, pp. 87-100, 2010.
- [8] M. Chase and S. Kamara, "Structured encryption and controlled disclosure," in *International Conference on the Theory and Application of Cryptology and Information Security*, Berlin, Heidelberg, 2010.
- [9] S. Kamara, C. Papamanthou and T. Roeder, "Dynamic searchable symmetric encryption," in *Proceedings of the 2012 ACM conference on Computer and communications security*, 2012.
- [10] S. Kamara and C. Papamanthou, "Parallel and dynamic searchable symmetric encryption," in *International Conference on Financial Cryptography and Data Security*, Berlin, Heidelberg, 2013.
- [11] D. Cash, S. Jarecki, C. Jutla, H. Krawczyk, M. Rosu and M. Steiner, "Highly-scalable searchable symmetric encryption with support for boolean queries," in *Advances in cryptology--CRYPTO 2013*, Berlin, Heidelberg, Springer, 2013, pp. 353-37.
- [12] D. Cash, J. Jaeger, S. Jarecki, C. S. Jutla, H. Krawczyk, M.-C. Rosu and M. Steiner, "Dynamic Searchable Encryption in Very-Large Databases: Data Structures and Implementation," in *21st Annual Network and Distributed System Security Symposium--NDSS 2014*, 2014.
- [13] A. A. Yavuz and J. Guajardo, "Dynamic Searchable Symmetric Encryption with Minimal Leakage and Efficient Updates on Commodity Hardware," in *International Conference on Selected Areas in Cryptography*, 2015.
- [14] M. S. Islam, M. Kuzu and M. Kantarcioglu, "Access Pattern disclosure on Searchable Encryption: Ramification, Attack and Mitigation," in *Ndss*, 2012.
- [15] D. Cash, P. Grubbs, J. Perry and T. Ristenpart, "Leakage-abuse attacks against searchable encryption," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, 2015.

# CPU Workload forecasting of Machines in Data Centers using LSTM Recurrent Neural Networks and ARIMA Models

Deepak Janardhanan

College of Engineering and Informatics  
National University of Ireland, Galway  
Galway City, Ireland  
Email: djdeepakjana72@gmail.com

Enda Barrett

College of Engineering and Informatics  
National University of Ireland, Galway  
Galway City, Ireland  
Email: enda.barrett@nuigalway.ie

*Abstract*—The advent of Data Science has led to data being evermore useful for an increasing number of organizations who want to extract knowledge from it for financial and research purposes. This has triggered data to be mined at an even faster pace causing the rise of Data Centers that host over thousands of machines together with thousands of jobs running in each of those machines. The growing complexities associated with managing such a huge infrastructure has caused the scheduling management systems to be inefficient at resource allocation across these machines. Hence, resource usage forecasting of machines in data centers is a growing area for research. This study focuses on the Time Series forecasting of CPU usage of machines in data centers using Long Short-Term Memory (LSTM) Network and evaluating it against the widely used and traditional autoregressive integrated moving average (ARIMA) models for forecasting. The final LSTM model had a forecasting error in the range of 17-23% compared to ARIMA model's 3742%. The results clearly show that LSTM models performed more consistently due to their ability to learn non-linear data much better than ARIMA models.

*Keywords*—LSTM Network, ARIMA Model, CPU Usage forecasting, Data Center

## I. INTRODUCTION

One of the major reasons for service disruptions in data centers is the inefficient allocation of work to machines, leading to resources not being utilized fully. A typical modern day machine in a data center can run hundreds to thousands of jobs per day with fluctuating resource requirements. Every data center has a management system that schedules units of work to these machines, but such systems fail to prioritize and adapt to the growing scheduling complexities in data centers. The main reason being their inability to take into consideration a machine's current state (CPU and Memory usage, number of running jobs etc.) before scheduling.

It was noted from previous work [1] that machines in Google's data center use on average only 45% of its CPU resources. The same number is around 60% for memory usage. On another note, Google's data centers all around the world use up 0.01% of the global power which is enough to light up 200,000 homes [2]. Even a 1% increase in efficiency of Google data centers can power over a thousand homes. These statistics are of just Google's Data Centers. Energy efficiency can be maximized with the use of Scheduling Systems that model a machines behavior before scheduling work on them. This work aims to provide a research contribution towards building scheduling systems that manage resources more efficiently and dynamically in data centers.

This study focuses on forecasting CPU usage of machines in Data Centers. In our approach, separate time series detailing CPU usage of machines are extracted from Google's cluster dataset. Later, exploratory time series analyses are carried out to understand the nature of the time series and two models of varying architecture are used to solve the forecasting problem. Firstly, the ARIMA model which is a proven traditional method for time series forecasting is used. Secondly, LSTM Recurrent Neural Network whose architectural design is very much suited for time series forecasting is used. The models are then evaluated and compared detailing the differences in the two models. The study used the cluster data trace released by Google that details the monitoring and resource usage information of a cluster of machines for a 29-day trace period. There isn't any other publicly available dataset that details scheduling information of a cluster as complex and large as this. The dataset is over 300GB in size and consists of scheduling information of over 12000 machines scattered across 5 tables, some of which span over a billion rows.

The rest of the paper is organized as follows. Section 2 describes the related work in the field. In Section 3, the background required to understand time series forecasting is explained including ARIMA models and LSTM Recurrent Neural Network. Section 4 and 5 details the feature extraction steps carried out and the exploratory analysis respectively. The modelling process using ARIMA and LSTM Models are detailed in Sections 6 and 7 respectively. Section 8 explains the evaluation of the two models with final conclusions of the research explained in Section 9 and possible future work described in Section 10.

## II. RELATED WORK

Di et al. [3] presents the design of a prediction algorithm based on a Bayes Model to estimate mean load in machines over a future time interval (up to 16 hours). The data is first reduced to a set of mean load predictions each starting from the current time and using Bayes model with 9 features detailing the current state of the machine. They mention that their model improved load prediction accuracy by 5% compared to other state of the art methods that are based on moving averages, auto-regression and noise filters. The time series method used in this approach is the aggregation of CPU usage into larger time windows to facilitate better forecasting. Since this leads to significant loss

of information, our study aims to forecast CPU usage for every 5 minute windows and incorporate more granular forecasts.

Ismeel and Miri [4] discuss developing a prediction model using k-means clustering and Extreme Learning Machines (ELMs) to estimate future Virtual Machine (VM) requests in data centers. The methodology first creates various categories of VM clusters with an end goal of developing a prediction model for each of these clusters. K-means clustering is used to create clusters of similar VM. Then for each cluster an ELM workload forecasting model is designed to estimate the number of requests for each cluster that may arrive in future time periods. The study evaluates its model on the Google's cluster-usage traces with their results suggesting that their model outperforms other similar models. While this approach aims to forecast VM requests, our study focuses on CPU usage estimation that are more dynamic and granular in nature.

Cao et al. [5] showcase their work using an Ensemble method to gain better performance in CPU load predictions in a cloud environment. The proposed model comprises of 2 layers, one for predictor optimization used for incorporating new predictor instances and removes instances with inferior performance. Another set of ensemble layers are used to facilitate providing feedback to the previous layer to adopt appropriate optimization strategies. This is achieved using a scoring method that assigns each predictor with a score representing its performance which is periodically updated. The study concludes that an ensemble model exhibits better forecasts and stability than individual models like Autoregression and Exponential smoothing. Our study uses the LSTM network to understand its forecasting prowess when compared to the models in this study.

Many of these previous works studied workload prediction of machines for large time windows. As mentioned earlier, this study aims to understand the growing complexities in data center resource allocation by building forecasting models around a granular time series dataset wherein CPU usage forecasts are generated for 5-minute time windows. More generally, previous work [6] [7][8][9][10] has also examined scheduling approaches using methods such as Reinforcement Learning and Genetic Algorithms, focussing not strictly on predicting CPU but also on optimising application response time, workflow scheduling and predicting network bandwidth. The approach detailed in this work can be applied in these more general settings and we hope will form the basis for further additional work.

### III. TIME SERIES FORECASTING

A time series refers to a sequence of observations that make a record of a particular activity over a period of time [11]. Due to the observations being sampled across time, there is an introduction of correlation between these points. Forecasting an activity or an observation usually involves collating its historical time series and finding patterns in them. Defining properties of time series data like trends and seasonal patterns can be identified from a time series plot and amount to the first step in any time series forecasting problem. Generally, transforming a time series into a smoothed-out version and then plotting it can reveal patterns that were previously unknown. One such commonly used smoothing technique is the Moving

Average. The moving average method is used for measuring seasonal variations in a time series by computing the arithmetic mean of values for time intervals throughout the time series and reducing any volatility in the data.

To further understand randomness in the data, the concept of Stationarity is introduced. A time series is said to be strictly stationary if the probability distribution of a set of values in that time series remains the same even when the set is shifted in time. Many of the real-world time series data are non-stationary in nature and hence require data adjustments to make them stationary by removing trends that are increasing or decreasing in nature. One such commonly used approach is differencing the time series. The difference operation is carried out by subtracting the current values with its past values. This is effectively called first order differencing. Similarly, differencing can be applied successively if required to further remove trends completely albeit at the cost of losing more and more information. Autocorrelation Function (ACF) of a time series is the amount of correlation that current values have with past values. ACF and Partial-ACF (PACF) plots help identify any seasonal patterns in the data.

#### A. ARIMA MODELS

White noise is a time series where each observation is randomly drawn from a population with variance and mean equal to zero. Typically, time series are required to follow such a pattern for better forecasting and any deviation from this amounts to a violation. Autoregressive (AR) and Moving Average (MA) are 2 models that help rectify these violations if identified. A stationary series is said to have no trend with constant variations in mean and fluctuates with a consistent pattern. This means that the autocorrelation would remain constant and form a consolidation of signal and noise. ARIMA models can be used to act as a filter that separates the signal and noise after which the signal is used to forecast values in the future. Forecasting using ARIMA is via an equation that is linear, where the predictors are lags of dependent variables and/or the errors (forecast errors). The model with only lag values of the variable becomes an autoregressive model. Hence, coefficients used in ARIMA consist of lagged errors that are optimized using nonlinear optimization [12]. In short an ARIMA model is denoted by ARIMA(p,d,q) where: p is the total autoregressive terms i.e AR terms, d is the total differences needed for stationarity and q is the total lagged forecast errors while forecasting i.e MA terms. Identifying the relevant ARIMA model for a given time series involves determining the values of p, d and q. To determine if AR and MA terms are necessary, ACF and PACF plots are used.

#### B. LSTM NETWORKS

Artificial neural networks are a computational paradigm that mimic biological neural networks through a network of connected computational units called neurons that are organized as layers. A Recurrent Neural Network (RNN) follows such a paradigm that is designed specifically for analyzing information where there is dependence between the current and previous values. LSTM networks are a class of RNN that are used to interpret and learn even long-term dependencies [13]. The major feature associated with LSTM networks are their ability to retain and persist information for long sequences or periods of time.

LSTM networks have 4 processing units within a single repeating unit as shown in Fig. 1. The line that runs through the repeating modules in the diagram is the cell state that runs through the entire LSTM network modifying information as it traverses. The pink blocks in the figure represent modules called gates that modify the cell state optionally.

For every input  $X$  the LSTM decides the amount of information that needs to be removed from the previous cell state using a sigmoid layer, also called the Forget layer. Here,  $C$  represents the cell state of the current repeating module which is passed as input to the next repeating module. For each cell state,  $C(t-1)$  received by the current cell, the forget layer outputs a value between 0 and 1 representing the amount of information to be removed from the cell state.

The LSTM network then decides the amount of new information that needs to be included in the cell state. This has two parts: a *sigmoid* layer and a *tanh* layer. The output of these two layers are multiplied and added to the cell state computed by the forget layer. Finally, each cell needs to output a value. This is done using another sigmoid and tanh layer which outputs a filtered down version of the cell state. LSTM networks as a result are well suited for time series forecasting as it retains information that govern the current state of an activity.

#### IV. FEATURE EXTRACTION: GOOGLE'S CLUSTER DATASET

The cluster data released by Google consists of 29-day worth of workload information of 12,453 machines scattered across 5 different tables. The *Machine Events* table describes machine states and hardware specifications over the trace period. The *Machine Attributes* table describe machine properties such as kernel version, clock speed etc. over the trace period. Life cycle of a job from start to finish i.e whether it's waiting for resource, failed, finished, lost and evicted is detailed in the *Job Events* table. Similarly *Task Events* table (over 100 million records) details the life cycle of a task from start to finish and finally the *Task Resource Usage* table (over 1 billion records) contains the resource (CPU, Memory, Disk I/O etc.) consumption of tasks over the trace period at every 5-minute intervals.

BigQuery is a Big Data querying tool in the Google Cloud Platform that is used to facilitate SQL querying on large datasets. Once imported, large aggregation such as cross joins across tables can be executed in very less time. Using the BigQuery platform, a time series dataset is created detailing CPU Usage of machines for every 5-minute interval, creating a total of 8352 data points for the 29-day period. The dataset is processed from the Machine Events, Task Events and Task Usage tables of Google's cluster dataset. To extract the CPU workload of machines in each time window accurately, we take into account the fact that some tasks run only partially in some 5-minute windows. Hence for each window, CPU is computed by summing all the task's CPU readings and then multiplying each reading with a weight equal to the amount of overlap that it has in the window. This ensures that a task with a one second

run time in a 5-minute window is not counted as the task running for entire window. There were instances where a period of inactivity in some machines were observed wherein N/A values were present in the CPU usage feature. Replacing them with 0 value is not characteristic of a machine. Hence, a linear interpolation was used that filled the N/A values with replacements which linearly interpolated neighboring values to maintain continuity in the series.

#### V. EXPLORATORY ANALYSIS

Its impractical to manually build a model for all 12000+ machines in the dataset. Explained here in detail are the time series analysis, modelling and forecasting approaches conducted for one machine (ID: 979583) and the same was followed for other machines too, though just briefly generalized and mentioned. The most active machines in the dataset were chosen for this study as they have a CPU usage value throughout the trace period and minimal N/A values. Henceforth, the machine with ID 979583 is referred to as Machine A. Time series plots of Machine A (see Fig. 2) exhibit daily seasonal patterns with sinusoidal fluctuations every day. A major difference between different machines modelled in our study is the maximum level of CPU usage reached. Machine A has a maximum usage value of 0.4546, while the same for other machines were significantly different.

#### VI. ARIMA MODELLING

In order to assess stability in the data, moving average summary points were computed for multiple values of time period which potentially smoothed out the raw data. After various considerations, a 2.5 hour moving average smoothing was chosen which brought in stability to the time series without losing too much information. This transformation on the time series was same for ARIMA models for all machines. The Augmented Dickey Fuller (ADF) test were used to determine whether a change in future values in a time series is dependent on previous lagged data and a linear trend. The Null hypothesis of ADF test represents a non-stationary time series. Other tests like ACF and PACF plots were conducted to visually inspect stationarity in a series. From the ACF plots it's concluded that there are significant autocorrelations with numerous lags for machine A. As a result, first order differencing was carried out. After differencing, same tests were again carried out and found Machine A time series to be stationary.

After careful considerations and combination of all parameters, a final ARIMA (2,1,2) model was fit on the Machine A Time series with seasonal components included in our model. The final ARIMA model for Machine-A had an RMSE of 0.0078 on the training data and 0.091 on the test data.

The forecasts (see Fig. 3) were able to follow patterns similar to the original time series till time 200 with clear mirroring of the slightest variations in the actual data. But after time 200, although it keeps up with the sinusoidal variations in the data, it can't mimic it completely. Hence, ARIMA modelling is a good forecasting model for near time predictions but fails to predict

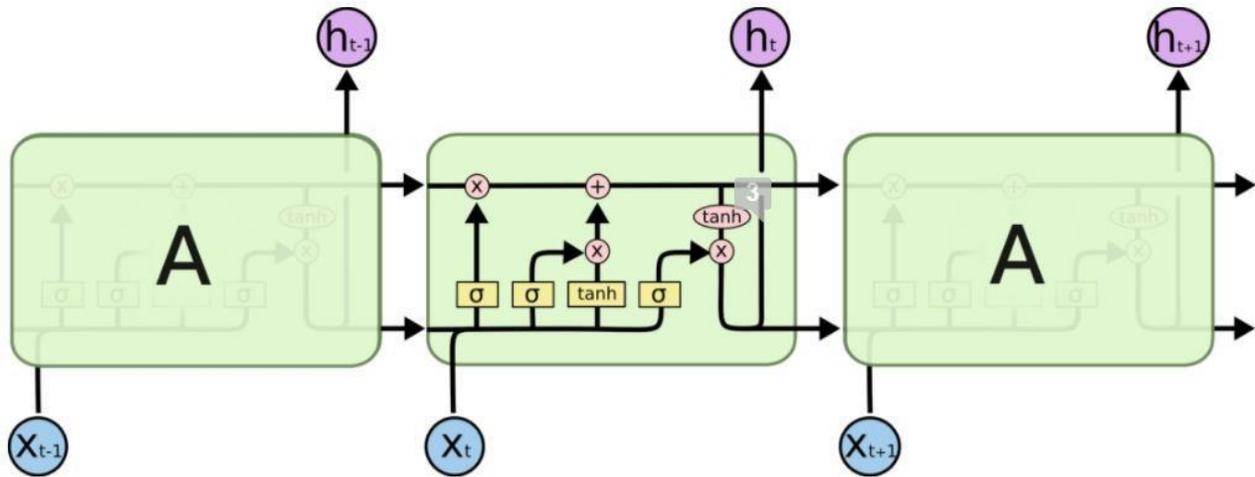


Figure 1. Internal structure of a Long Short-Term Memory Network

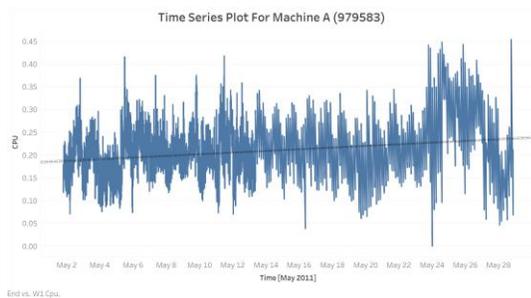


Figure 2. Time Series plot of Machine A

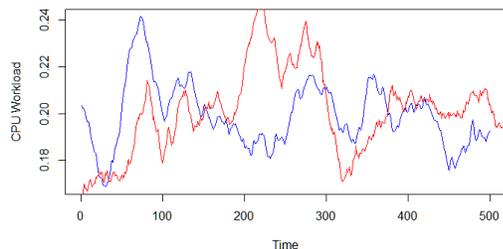


Figure 3. Forecasts (Red) of ARIMA (2,1,2) model with seasonal components included VS Actual (Blue) Time Series of Machine-A

long term CPU usage values. The results of ARIMA models for other machines are also similar where the forecast follows the same pattern as the actual data for near time predictions. But the models lose credibility in making long term predictions. Note that the y-axis scale of the plot in Fig. 3 is in the range 0.2 and 0.5, meaning the forecasts are accurate but due to granularity of the data, it's cannot be interpreted visually with full scale plots.

## VII. LSTM NETWORK MODELLING

Firstly, the time series were transformed to a stationary time series. From the previous section, it was found that performing a first order differencing was sufficient to stabilize the data

without losing much information. Hence, first order differencing was applied. The activation function of LSTM is a tanh function which can output a value between -1 to 1. Hence the time series is transformed by normalizing it within this range before splitting it into training and testing sets. Hyperparameters were tuned after training multiple LSTM Networks with different combinations of training epoch and number of neurons in the hidden layer. Firstly, the effects on training losses were assessed for different training epochs numbers. Then, LSTM models were fit with varying number of neurons and their RMSE and MAPE values were calculated. After tuning the LSTM Network's hyperparameters, a final LSTM model configuration that yielded the best results when configured to 1 hidden layer with 5 neurons and training epochs set to 3000. The models were then fit to the training data and the forecasting approach was tweaked to facilitate better forecasts. Instead of forecasting each time step in the future from the test data, a more dynamic approach was used wherein we updated the model after forecasting each value in the series to include the new forecasted value too. The final LSTM Model for Machine A had a training RMSE value of 0.0317 and test RMSE of 0.0381. Fig. 4 shows the initial few forecasts made by the model.

To look closely at the forecasts, plots of initial sets of forecasts for Machine A were plotted. Specifically, CPU forecasts of the first 2, 12 and 24 Hours in the test data were plotted (see Fig. 4) against their actual CPU values to see how closely the model understood recent data (i.e training data). From plots (ii) and (iii) in Fig. 4, it can be noted that forecasts closely followed the actual CPU usage reading, displaying similar seasonal patterns to an extent. Plot(i) shows the first 2-hour forecasts, and it's evident that the model has trained on the training dataset well. As mentioned earlier, fine tuning the hyperparameters not only increased the performance with respect to metrics RMSE and MAPE, it also has been able to pick up seasonal components in the time series.

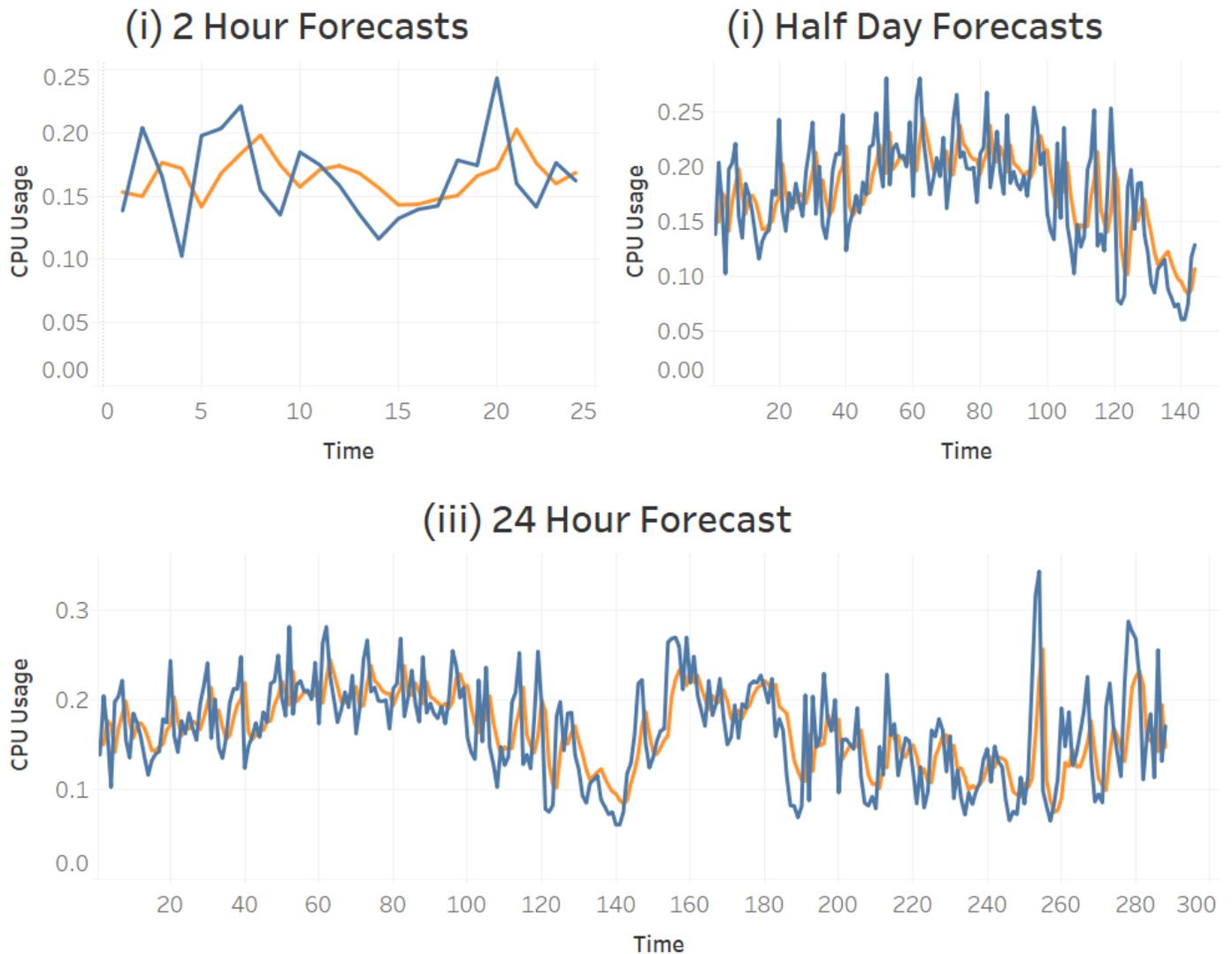


Figure 4. First (i)2 Hour, (ii)Half Day (iii)24 Hour Forecasts of LSTM Model for Machine A test data. Blue: Actual CPU vs Red: Forecasted CPU

The results of the LSTM models for machines other than Machine A exhibited similar model features. Time series of many machines were similar in nature hence didn't require much individual attention. For machines whose CPU usage time series were significantly different were modelled accordingly but exhibited just slightly fluctuating results. But one aspect that was found similar for all the models was the hyperparameters that were required to train them. As noticed in Machines A, there wasn't any notable change in the training RMSE scores when trained on 2000 and 3000 epochs. The number of neurons too were limited to 5 which gave the best results during training of models with a single hidden layer.

## VIII. MODEL EVALUATION

Both the models were trained differently. Differencing and statistical moving average data transformations were used while modelling using ARIMA models. The final ARIMA (2,1,2) models for Machine A had good RMSE scores but significantly

different test RMSE. The Akaike information criterion and Bayesian information criterion scores obtained in the ADF tests were also used to find the best ARIMA models. Both these metrics are found to have very high possibility of overfitting for AR processes [14]. CPU usage of a machine in data centers always have dependence between the previous and current values. This fundamentally makes a CPU usage time series an AR process. LSTM Network essentially being a non-linear time series model performed better than ARIMA model for almost all combinations of its hyperparameter.

ARIMA model was observed to be prone to overfitting for multiple CPU time series of machines. They performed well in terms of training RMSE values but poorly on the test data. On the other hand, LSTM models for all machines were consistently performing better without any significant fluctuations in model performance.

## IX. CONCLUSION

The forecasting of CPU workload of Machines in Data Centers using LSTM Networks has shown significant improvement over the ARIMA models. To the author's knowledge, this is the first time LSTM networks has been used for forecasting CPU workload of machines in Data Centers. More specifically, this is the first piece of work that evaluated LSTM Network in its efficiency to forecast multiple highly volatile CPU usage time series extracted from Google's Cluster Data.

The problems associated with instability and fluctuations in CPU usage of machines in Google's data center and how they affect a time series forecasting model's accuracy are highlighted. The ARIMA modelling approach to forecasting CPU workload is detailed along with its limitations. More generally, its inability to model an unstable and volatile time series is explained taking the CPU usage time series as an example to showcase the level of transformations required to build a decent forecasting model out of it. The initial ARIMA models built couldn't fit a time series with high CPU fluctuations observed on an hourly basis. Significant data transformations have added little to no major improvements in the forecasting performance of the ARIMA model. The ARIMA models were evaluated and found to be overfitting the time series test data with significant difference in results between the RMSE scores for the training and testing phases of the model. The final ARIMA(2,1,2) model had a forecasting error in the range of 37.331 to 42.881% across multiple machines.

On the other hand, the LSTM Network forecasted the CPU workload with a forecasting error rate in the range of 17.566 to 23.65% for multiple machines. The use of LSTM Networks in this study is a new step towards estimation of complex workload of machines in data centers due to its ability to maintain memory of patterns and trends. In the end, it forecasted CPU workload better than many traditional forecasting models as well as some simpler versions of Recurrent Neural network. Hence, a case for LSTM Networks as an excellent time series forecasting tool for workload estimation of resources in data centers has been put forth in this study.

## X. FUTURE WORK

This study detailed the time series forecasting process employed for univariate time series forecasting of CPU workload of machines in data center. As future work, improving the forecasting performance of the LSTM model can be carried out. One such approach can be the use of multiple predictor variables in learning and fitting a LSTM model around it. Some of the predictor variables that can be used are: number of tasks running in machines, the memory usage etc. Other features like the number of tasks failed, completed, restarted and killed in every 5-minute window could be extracted from the dataset too. There are other unconventional features that aren't used like resource usage metrics such as CPI (cycles per instructions), mean disk I/O time and memory accesses per instruction.

Another approach to forecasting the same data is by considering the time series as a sequence. Till now almost all Time Series forecasting approaches that have modelled this data have fit their models considering observations at each time step

as a single input. CPU usage in this data had clear daily seasonal patterns. Hence considering an entire day's observations as a single data point i.e 288 features (1 day has 288 5-minute windows), where each time window of the day is considered as a feature. Such a dataset would basically be a sequence to sequence Time series forecasting problem where the model forecasts the CPU usage for the next day. Such an approach to time series forecasting especially on this data can be promising.

## REFERENCES

- [1] Charles Reiss et al. "Towards understanding heterogeneous clouds at scale: Google trace analysis". In: *Intel Science and Technology Center for Cloud Computing, Tech. Rep 84* (2012).
- [2] *Big Data and Data Center Fun Facts - AIS*. <http://www.americanicis.net/2014/big-data-data-center-fun-facts/>. (Accessed on 09/30/2017).
- [3] Sheng Di, Derrick Kondo, and Walfredo Cirne. "Host load prediction in a Google compute cloud with a Bayesian model". In: *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*. IEEE Computer Society Press. 2012, p. 21.
- [4] Salam Ismaeel and Ali Miri. "Using ELM techniques to predict data centre VM requests". In: *Cyber Security and Cloud Computing (CSCloud), 2015 IEEE 2nd International Conference on*. IEEE. 2015, pp. 80–86.
- [5] Jian Cao et al. "CPU load prediction for cloud environment based on a dynamic ensemble model". In: *Software: Practice and Experience* 44.7 (2014), pp. 793–804.
- [6] Martin Duggan et al. "A network aware approach for the scheduling of virtual machine migration during peak loads". In: *Cluster Computing* (2017), pp. 1–12.
- [7] Martin Duggan et al. "An Autonomous Network Aware VM Migration Strategy in Cloud Data Centres". In: *Cloud and Autonomic Computing (ICCAC), 2016 International Conference on*. IEEE. 2016, pp. 24–32.
- [8] Enda Barrett, Enda Howley, and Jim Duggan. "Applying reinforcement learning towards automating resource allocation and application scalability in the cloud". In: *Concurrency and Computation: Practice and Experience* 25.12 (2013), pp. 1656–1674.
- [9] Martin Duggan et al. "A reinforcement learning approach for the scheduling of live migration from under utilised hosts". In: *Memetic Computing* (2016), pp. 1– 11.
- [10] Enda Barrett, Enda Howley, and Jim Duggan. "A learning architecture for scheduling workflow applications in the cloud". In: *Web Services (ECOWS), 2011 Ninth IEEE European Conference on*. IEEE. 2011, pp. 83–90.
- [11] Rainer Schlittgen. "Robert H. Shumway and David S. Stoffer: Time series analysis and its applications with R examples, 2nd edn." In: *AStA Advances in Statistical Analysis* 92.2 (2008), pp. 233–234.
- [12] Robert Nau. "Statistical forecasting: notes on regression and time series analysis". In: *Durham: Fuqua School of Business, Duke University* (2015).
- [13] *Understanding LSTM Networks – colah's blog*. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. (Accessed on 09/30/2017).
- [14] Yamei Liu. "Overfitting and forecasting: linear versus non-linear time series models". In: (2000).

# An Advanced Reinforcement Learning Approach for Energy-Aware Virtual Machine Consolidation in Cloud Data Centers

Rachael Shaw, Enda Howley, Enda Barrett

Department of Information Technology  
National University of Ireland Galway  
Galway, Ireland  
r.shaw4@nuigalway.ie

**Abstract**— Energy awareness presents an immense challenge for cloud computing infrastructure and the development of next generation data centers. Inefficient resource utilization is one of the greatest causes of energy consumption in data center operations. To address this problem we introduce an Advanced Reinforcement Learning Consolidation Agent (ARLCA) capable of optimizing the distribution of virtual machines across the data center for improved resource management. Determining efficient policies in dynamic environments can be a difficult task, however the proposed Reinforcement Learning (RL) approach learns optimal behaviour in the absence of complete knowledge due to its innate ability to reason under uncertainty. Using real workload data we evaluate our algorithm against a state-of-the-art heuristic, our model shows a significant improvement in energy consumption while also reducing the number of service violations.

**Keywords**- energy efficiency; cloud computing; resource management; reinforcement learning

## I. INTRODUCTION

Cloud computing services offered by companies such as Amazon and Google deliver on-demand virtualized resources which can be accessed over the internet and charged using a pay-as-you-go pricing model [1]. The ability to scale up or down computing resources in response to current demand has led to the tremendous growth and wider adoption of cloud computing across several domains. Despite the advancements in energy efficient computing devices and data center equipment, excessively high energy consumption and carbon dioxide emissions continue to soar in the operation of large scale data centers today. Recent studies in cloud computing have highlighted the environmental impact of data centers in terms of electricity costs and associated CO<sup>2</sup> emissions. In 2013, U.S data centers consumed a total of 91 billion kilowatt-hours of electricity. By 2020 the level of consumption is estimated to increase to approximately 140 billion kilowatt-hours annually, costing 13 billion per year in electricity bills and furthermore, pollution of 150 million metric tons of carbon dioxide [2]. According to GeSI SMARTer 2020 report data center emissions are projected to increase by 7% annually [3]. In addition, a study by Petty et al. highlighted that the Information Communication Technology (ICT) industry contributes to 2% of global CO<sup>2</sup>

emission annually which is the equivalent to that produced by the aviation industry [4].

The era of on-demand computing is powered by the concept of virtualization which was pioneered by IBM in the 1960's [5]. Virtualization aids in the promotion of increased host utilization by apportioning the resources of large physical hosts into smaller independent machines each of which is known as a Virtual Machine (VM). Each VM runs in apparent isolation equipped with its own Operating System (OS) and applications which allows for the simultaneous execution of multiple tasks on a physical host. This overall results in increased efficiency and resource utilization.

However, despite the benefits gained from the advancements in virtualization technologies one of the major inefficiencies in data center deployments is caused by poorly managed and idle resources [6]. Current studies have revealed that on average hosts operate at a mere 12-15% of their full capacity resulting in the wastage of valuable resources while underutilized hosts have been proven to use up to 60% of their maximum power resulting in significant draws on energy consumption [2]. VM consolidation is one approach that can significantly improve the management of resources by strategically reallocating VMs on to a reduced number of hosts in an effort to conserve energy. Consolidating a larger number of VM instances on an already loaded host can however, cause a surge in energy consumption while also potentially causing the host to become overloaded incurring service violations and further requiring VMs to be migrated to additional hosts in the datacenter [7]. Conversely, placing a VM on an underutilized host promotes the continuation of poor resource utilization. As a result, striking a balance between both energy efficiency and performance is essential to achieving high performance while reducing overall energy consumption.

To address this issue we present a self-optimizing Reinforcement Learning (RL) VM consolidation model to optimize the allocation of VM instances while also adhering to strict Service Level Agreements (SLA). The agent continues to learn an optimal resource allocation policy depending on the current state of the system through repeated interactions with the environment. In addition, our model employs an advanced reward shaping technique known as Potential Based Reward

Shaping (PBRS). One of the more profound limitations of standard RL algorithms is the slow rate at which they converge to an optimal policy [8]. PBRS allows expert advice to be included in the learning model to assist the agent to learn more rapidly and as a result encouraging more optimal decision making in the earlier stages of learning. Our results show the RL agents ability to learn and adapt in a highly volatile cloud environment while also delivering an intelligent energy-performance tradeoff capability resulting in improved energy efficiency and performance.

The contributions of this paper are the following:

- Using a state-of-the-art RL technique we present an autonomous VM consolidation model capable of optimizing the distribution of VMs across the data center in order to achieve greater energy efficiency while also delivering the required performance.
- We apply our model to a large scale simulated data center using CloudSim. Using real workload traces we show the advantages of our approach over a state-of-the-art heuristic.

The remainder of this paper is structured as follows: Section II discusses related work, Section III introduces RL, Section IV presents the design of our proposed cloud resource management model, Section V describes our experimental setup and performance metrics, Section VI presents the results and Section VIII concludes the paper.

## II. RELATED WORK

In recent years the pervasiveness of cloud computing has urged research initiatives to tackle the challenging problems of inefficient resource management policies. In the literature two types of approaches are used, these can be classified as Heuristic/Threshold and AI based approaches.

### A. Heuristic/Threshold Based Approaches

Heuristic/Threshold based approaches are the most widely used methodologies for resource management in cloud infrastructure. These types of approaches trigger static resource allocation decisions on reaching a predefined threshold. Verma et al. presented pMapper an application placement controller which aims to minimize energy consumption and migration costs while maintaining SLA [9]. The underlying architecture is composed of three distinct management entities, namely a performance manager which monitors current performance and resizes VM instances based on the SLA, A power manager which manages the power state of underlying hardware and a migration manager which estimates the cost of live migration. Their overall results showed a savings in power consumption of hosts of 25%. Cardosa et al. investigated the impact on VM resource allocation and power consumption by leveraging min, max and share parameters analogous to those used in commercial virtualization technologies [10]. These parameters define the upper and lower bounds for resource utilization for each VM while the shares parameter denotes the priority of each VM during the distribution of spare resources. The authors use such parameters to drive their VM placement and consolidation strategy which achieved a significant improvement in the

overall utility of the data center. Lee et al. proposed the implementation of two task consolidation heuristics known as ECTC and MaxUtil in order to curtail energy consumption of underutilized resources [6]. One of the most important research contributions in the field of VM consolidation is accredited to the work of Beloglazov et al. as outlined in two of their more highly cited papers [7][11]. They introduced a three stage VM resource optimization approach consisting of Host overutilization/underutilization detection, VM selection and VM placement. In particular, they proposed the Lr-Mmt algorithm which manages host utilization detection and VM selection. In addition, this algorithm uses a VM placement heuristic known as Power Aware Best Fit Decreasing (PABFD). This heuristic considers the heterogeneity of cloud resources by selecting the most energy efficient hosts first in order to allocate VMs. Their experimental results concluded that the composition of Lr-Mmt in conjunction with PABFD significantly outperforms all other VM consolidation procedures resulting in a profound reduction in energy consumption and SLA violations.

### B. Reinforcement Learning Based Approaches

Alternatively, research initiatives continue to explore the application of AI methodologies in order to evolve a new generation of autonomic resource management strategies. Duggan et al. introduced an RL network-aware live migration strategy, their model enables an agent to learn the optimal times to schedule VM migrations to improve the usage of limited network resources [12]. Tesauro et al. presented an RL approach to discover optimal control policies for managing Central Processing Unit (CPU) power consumption and performance in application servers [13]. Their proposed method consists of an RL based power manager which optimizes the power performance tradeoff over discrete workloads. Barrett et al. applied a parallel RL learning approach to optimize resource allocation in the cloud [14], while other work introduced an agent based learning architecture for scheduling workflow applications [15]. Rao et al. introduced VCONF an RL based VM auto configuration agent to dynamically re-configure VM resource allocations in order to respond effectively to variations in application demands [16]. VCONF operates in the control domain of virtualized software, it leverages model based RL techniques to speed up the rate of convergence in nondeterministic environments. Das et al. introduced a multiagent based approach to manage the power performance tradeoff by specifically focusing on powering down underutilized hosts [17]. Dutreil et al. also showed how RL methodologies are a promising approach for achieving autonomic resource allocation in the cloud. They proposed an RL controller for dynamically allocating and deallocating resources to applications in response to workload variations [18]. Using convergence speed up techniques, appropriate Q function initializations and model change detection mechanisms they were able to expedite the learning process. Farahnakian et al. [19] implemented RL to learn the optimal time to power on or switch off a host based on future resource demands.

Compared to related research our work is different in that:

- 1) Using a PBRS inspired RL technique we present Advanced Reinforcement Learning Consolidation

---

The primary author would like to acknowledge the ongoing financial support provided by the Irish Research Council.

Agent (ARLCA) a self-optimizing VM consolidation approach that allocates the optimal number of VMs to hosts such that each host operates at an optimized resource usage rate. We show how this approach optimizes three key performance metrics namely energy consumption, VM migrations and service violations.

- 2) Unlike static threshold based approaches this approach is more suitable for decision making in highly dynamic environments while also considering allocation decisions that could possibly suffer from delayed consequences. We show using real workload traces how our model outperforms a state-of-the-art heuristic approach across all performance metrics in order to deliver a more sustainable green cloud infrastructure.

### III. REINFORCEMENT LEARNING

Reinforcement Learning enables an agent to learn optimal behaviour through repeated trial and error interactions with its environment at discrete time steps  $t$  without any prior knowledge. The agent receives a reward depending on the action selected. The objective of the agent is to discover overtime which actions yield the greatest rewards [20].

RL control problems can be intuitively modelled as Markov Decision Processes (MDP) which provide a model for sequential decision making problems faced with adverse uncertainty [14]. The learning process for an RL agent is composed of:

- 1) State space: A set of environment states. At the end of each time step  $t$  the learning agent occupies a state denoted  $s_t \in S$ .
- 2) Action space: The agent selects a possible action  $a_t \in A(s_t)$  where  $A(s_t)$  refers to the set of all possible actions in the current state  $s_t$ .
- 3) Reward signal: Once the selected action is executed it results in a state transition  $s_{t+1}$ , the agent is then allocated a positive or negative reward signal  $R(s_t, a_t)$  depending on the state of the environment after an action has occurred. The overall goal of MDP is to generate a mapping of states to associated actions which maximize the accumulated reward.

Sarsa is a popular RL Temporal Difference (TD) learning algorithm which can be used to discover an optimal policy. TD methodologies implement prediction based learning by incrementally updating current estimates of state-action pairs  $Q(s_t, a_t)$  based on the outcome of previous estimates. Sarsa is an acronym for state, action, reward, state, action. Its name is derived from the sequence of events that must occur in order to transition from one state to the next and update the current model estimates. The update rule is as follows:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]. \quad (1)$$

where  $Q(s_t, a_t)$  denotes the expected reward of selecting action  $a_t$  in state  $s_t$ .  $\alpha$  is the learning rate, a value set close to 1 promotes continuous updates to estimates while  $\alpha$  defined

close to 0 reduces learning.  $\gamma$  is a discount factor which determines the degree to which an agent favors long term rewards over short term gains. A value closer to 1 results in an agent that is more forward looking and strives to maximize future rewards while a rate closer to 0 results in an agent that assigns a greater weight to short term rewards.  $r_{t+1}$  defines the reward signal allocated for selecting an action in a given state while  $Q(s_{t+1}, a_{t+1})$  is the Q-value estimate of the resulting state and the action selected in the next time step  $t$ .

A policy  $\pi$  guides the agent's decision making process when selecting an appropriate action for any given state. In order to discover an optimal policy there is a tradeoff between exploration and exploitation. An agent that invariably exploits the best action fails to discover potentially more lucrative actions by choosing to explore its environment. In order to manage such a tradeoff we implement a softmax action selection strategy [21]. Softmax assigns action probabilities according to the expected utility, thus ensuring higher rewarding actions are more likely to be explored. Illustrated below is the standard Sarsa learning algorithm.

---

#### Algorithm 1: Sarsa

---

**Initialize**  $Q(s, a)$  arbitrarily

**Repeat (for each episode):**

**Initialize**  $s$

**Choose**  $a$  from  $s$  using  $\pi$

**Repeat (for each step of episode):**

Take action  $a$ , observe  $r, s'$

Choose  $a'$  from  $s'$  using  $\pi$

$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma Q(s', a') - Q(s, a)]$

$s \leftarrow s'; a \leftarrow a';$

**until**  $s$  is terminal

---

One of the limitations of RL methodologies is the rate at which an agent converges to an optimum policy. In order to expedite the learning process we implement PBRS which has shown to be a powerful technique to improve the convergence rate of RL agents by using domain knowledge to assist the agent to learn more rapidly [22]. In particular, PBRS provides the learner with an additional reward through the mapping of states to associated potentials using the following function:

$$F(s, a, s') = \gamma \Phi(s') - \Phi(s). \quad (2)$$

where  $\Phi$  is the potential function which maps states to potentials and  $\gamma$  is defined as the same discount factor applied in the update rule (1). The PBRS reward is concatenated to the standard reward received from the environment. Using the results of previous experiments we fine-tuned the potentials with more lucrative rewards for allocation decisions resulting in host utilization states between 40-70% as this resulted in the best performance according to our experimental analysis.

---

The primary author would like to acknowledge the ongoing financial support provided by the Irish Research Council.

#### IV. DYNAMIC VM CONSOLIDATION MODEL

In order to reach new frontiers in energy efficient cloud infrastructure we propose an Advanced RL Consolidation Agent known as ARLCA which is capable of driving both efficiency and Quality of Service (QoS) by dynamically adjusting its behaviour in response to changes in workload variability.

More specifically, ARLCA is presented with a list of VMs that require allocation to suitable hosts in the datacenter. Through repeated interactions with the environment ARLCA discovers the optimal balance in the dispersal of VMs across the data center so as to prevent hosts becoming overloaded too quickly but also ensuring that resources are operating efficiently.

##### A. State-Action Space

The application of RL in more complex problem domains such as the cloud requires careful definition of the state-action space in order for it to operate effectively. We design a novel state-action space by defining all possible states and actions as a percentage ranging from 0-100%. We use increments of 1% which provided the best performance according to an experimental parameter sweep. The state space  $S$  denoted below in (3) represents the global state of the environment. It can be defined as the number of active hosts  $a_h$  in the environment as a percentage of the total number of hosts  $t_h$ .

$$S = \frac{\sum_{i=1}^n a_h}{t_h} \times 100. \quad (3)$$

An action  $A$  is a combined variable composed of the utilization rate of any given host coupled with the size of the VM to be placed. As denoted below in (4) the host utilization rate  $h_u$  is calculated as the sum of the total requested resources  $trr$  for each VM residing on the host as a percentage of the hosts' capacity  $h_c$ . Additionally, VM utilization  $vmu$  is computed as the VMs requested resources  $r$  returned as a percentage of the total host capacity  $hc$ . This determines the size (CPU resource requirements) of the VM to be placed. Defining the state-action space as percentages from 0-100% significantly reduces the size of the state-action space thus preventing the agent from engaging in an exhaustive search. This type of approach allows for the deployment of a more agile and efficient agent capable of pursuing its design objectives.

$$A = \left[ h_u = \frac{\sum_{j=1}^n trr}{h_c} \times 100 + vmu = \frac{r}{h_c} \times 100 \right]. \quad (4)$$

##### B. ARLCA Learning Model

The Sarsa driven learning algorithm coupled with the advanced PBRS component used to train our agent is presented above. When invoked the ARLCA learning algorithm calculates the global state of the environment using (3). The first VM to be placed is selected from the placement list and the CPU host utilization rate for each host is calculated and returned as a percent ranging between 0-100%. Next the size of the VM is computed and a list of possible actions is generated using the combined action variable (4). The agent then selects a host based on the softmax action selection strategy and the host is placed on

a migration list which keep a record of allocation decisions. The global state is recalculated and both the MDP and PBRS rewards are generated. Sarsa updates Q-value estimates based on the action that will be implemented in the subsequent state, in order to do so the utilization rates of the hosts and the size of the next VM is recalculated and used to update the Q-value estimate. The result is stored in the Q-value matrix which effectively stores the mapping of states to associated actions representing the agents current knowledge. Lastly, the global state is updated and the last action selected is implemented in the subsequent iteration. ARLCA continues to execute until all VMs are mapped on to various hosts in the environment.

---

##### Algorithm 2: ARLCA Learning Algorithm

---

```

Input : VM Placement List
calculate globalState
foreach host → hostList do
    calculate hostUtil
end
calculate vmSize
calculate possibleActions ← vmSize + hostUtil
select host from possibleActions using  $\pi$ 
foreach vm → vmPlacementList do
    allocate vm
    observe globalState + 1, rewards
    foreach host → hostList do
        calculate hostUtil
    end
    calculate nextVmSize
    calculate possibleActions ← nextVmSize + hostUtil
    select host from possibleActions using  $\pi$ 
    calculate  $Q(s,a) \leftarrow Q(s,a) + \alpha[r + \mathbf{F}(s,s') + \gamma Q(s',a) - Q(s,a)]$ 
    update QValueMatrix
    globalState ← globalState + 1
    action ← host
end
Output: mapping of VMs to hosts
    
```

---

#### V. EXPERIMENTAL SETUP

To evaluate the performance of the proposed energy-aware RL learning agent ARLCA we have selected the state-of-the-art PABFD heuristic as a benchmark. We develop an RL framework as an extension of the CloudSim simulator as used in the studies of Beloglazov et al. [7][11]. CloudSim supports the management of cloud resources and contains the necessary components to enable the empirical evaluation of energy-aware cloud based simulations. In order to simulate a large scale cloud

environment 800 HP ProLiant ML110 G5 hosts were configured in the data center. These hosts consisted of two cores with the capacity to process 2660 Million Instructions Per Second (MIPS). Our experiments leveraged the 10 day CPU workload traces provided by CloudSim which were generated from real hosts deployed in over 500 locations globally. In order to measure the robustness of the proposed approach we used these traces to generate a randomized 30 day workload which models more precisely the complexity and dynamic nature of the cloud environment overtime while also evaluating more thoroughly the capability of the proposed agent to learn in such an environment.

A. Performance Metrics

The key performance metrics to evaluate the effectiveness of the proposed algorithms are as follows:

- 1) **Energy Consumption:** This is defined as the total energy consumed by the data centers computational resources as a direct result of processing application workloads.
- 2) **VM Migrations:** This is the total number of VM migrations that occur during the simulation process. Each time a VM is migrated it is typically subjected to SLA violations.
- 3) **SLA Violations:** The ability of cloud providers to deliver SLA is critical and a core function of their operation and as a result we also consider the impact on SLA violations.

VI. RESULTS

We evaluate ARLCA against the Lr-Mmt policy which harnesses the state-of-the-art PABFD consolidation heuristic using the stochastic 30 day workload in order to demonstrate the benefits of a more adaptive and intelligent methodology.

Fig. 1 illustrates the behaviour of both policies in relation to energy consumption over the 30 day workload. As shown, the implementation of ARLCA resulted overall in a considerable reduction in energy by a total of 25.35% with an average energy savings of 39.7 kWh per day (Std Dev 20.49). Furthermore, it is also apparent that on day 28 ARLCA failed to outperform Lr-Mmt which resulted in a slight increase in energy of 0.5%. To analyze whether the overall energy reduction achieved is statistically significant a two tailed t-test was performed which resulted in a p-value of <0.0001 with a 96% confidence interval (32.068, 47.372). These results reveal that the energy savings achieved over the Lr-Mmt policy are extremely significant.

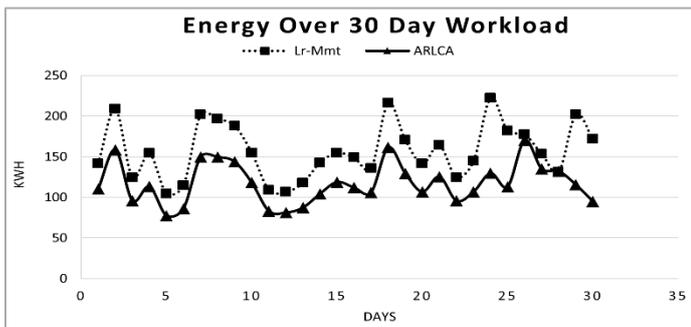


Figure 1. Energy Consumption

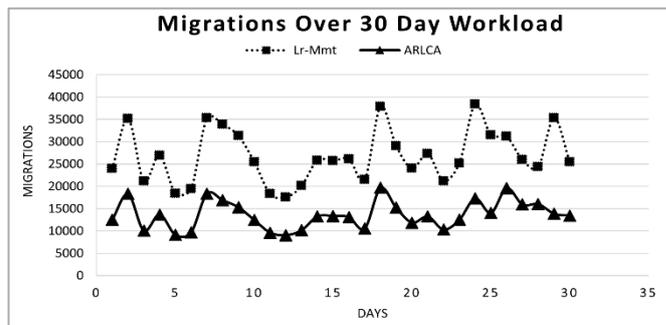


Figure 2. Migrations

Fig. 2 displays the number of migrations incurred by both policies over the 30 day workload. As illustrated, the application of ARLCA also had a positive impact on the number of migrations. ARLCA reduced migrations by 49.17% in total (394,587 migrations). In addition it reduced the mean number of migrations per day by 13,153 (Std Dev 3536.21). The results were also statistically significant with a p-value of <0.0001 with a 95% confidence interval (11832.46, 14473.34).

Fig. 3 presents the number of SLA violations incurred by both policies during the simulation. The Lr-Mmt policy resulted in a surge in the number of service violations while ARLCA showed an overall 63% decrease in the mean number of SLA violations. Again these results were also statistically significant with a p-value of 0.00001121 with a 95% confidence interval (0.21654124273, 048922395074).

An important dimension in achieving greater energy efficiency through resource optimization is managing the tradeoff between energy and performance which are inextricably linked. Notably, an interesting observation in Fig. 1 was that on day 28 the Lr-Mmt policy generated a slight improvement in energy consumption of a mere 0.5%. However, Fig. 3 confirms that this marginal improvement was achieved at the cost of increased SLA violations as indicated by the spike generated in the number of violations by the Lr-Mmt policy on day 28. This suggests that Lr-Mmt consolidated VMs more aggressively in an attempt to successfully reduce energy but failed to efficiently manage the energy-performance tradeoff resulting in a surge in the number of service violations. In contrast ARLCA strikes a more precise balance with such a tradeoff by generating a relatively similar energy rating of just .5% in the difference. However, more profoundly the agent also achieves a significant 76.2% decrease in the number of SLA violations on day 28 alone.

Overall the key points arising out of these results are that through the deployment of our advanced energy efficient learning agent ARLCA we introduce a more agile and adaptive solution to consolidate and support the movement of VMs between physical hosts in the data center. More specifically, our empirical results demonstrate the improved efficiency achieved by leveraging a more sophisticated and dynamic RL solution which has the inherent ability to efficiently adapt to a continuously changing cloud environment. As a result, we deliver a significant reduction in energy consumption of 25.35% with a decrease of up to 44.7% in energy per day over the state-of-the-art Lr-Mmt heuristic. Furthermore, we reduce

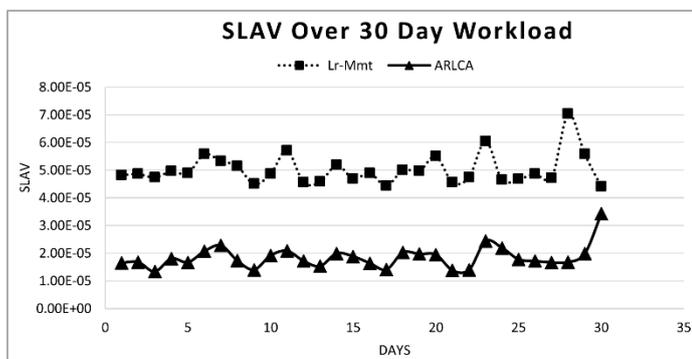


Figure 3. Service Level Agreement Violations

the number of service violations by a significant 63%. Such reductions reduce the overall data center operational costs resulting in a more competitive cloud infrastructure. This solution also has wider implications as it stands to provide a more sustainable green cloud infrastructure in support of global environmental sustainability while also promoting the greater adoption of AI techniques to achieve intelligent decision making in cloud based infrastructure.

## VII. CONCLUSION

Through the innovative application of more sophisticated and advanced methodologies adopted from the field of Artificial Intelligence we developed ARLCA, an intelligent solution to optimize the distribution of VMs across the data center. ARLCA demonstrates the potential of more advanced intelligent solutions capable of reaching new frontiers in data center energy efficiency while also achieving significant improvements in the quality of the service provided.

However, there remains many open challenges that must be addressed in order to provide a more complete solution to this complex problem. One area in particular is energy efficient strategies which take into consideration the utilization of multiple systems resources. As a result, we intend on extending our proposed model to develop a solution which considers resources such as Random Access Memory (RAM) and also network bandwidth.

In addition, our results showed that even through the deployment of an intelligent agent service violations were still evident periodically. We also plan on exploring this further through the implementation of multi-objective optimization techniques where both energy and service violations are optimized simultaneously.

## REFERENCES

- [1] R. Shaw, E. Howley and E. Barrett, (in press) "Predicting the available bandwidth on intra cloud network links for deadline constrained workflow scheduling in public clouds", International Conference on Service-Oriented Computing. Springer International Publishing, 2017.
- [2] J. Whitney and P. Delforge, "Scaling up energy efficiency across the data center industry: evaluating key drivers and barriers", tech. report, Natural Resources Defense Council, August, 2014.
- [3] L. Neves, J. Krajewski, P. Jung and M. Bockemuehl, "GESI SMARTer 2020: The role of ICT in driving a sustainable future", tech. report, Global e-Sustainability Initiative and The Boston Consulting Group, Inc., 2012.
- [4] C. Pettey, "Gartner estimates ICT industry accounts for 2 percent of global co2 emissions", Gartner. April, 2007.
- [5] P. Healy, T. Lynn, E. Barrett and J.P. Morrison, "Single system image: A survey", Journal of Parallel and Distributed Computing, 90, 2016. pp.35-51.
- [6] Y.C. Lee, and A.Y. Zomaya, "Energy efficient utilization of resources in cloud computing systems", The Journal of Supercomputing, vol. 60, 2012, pp. 268–280.
- [7] A. Beloglazov and R. Buyya, "Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers", Concurrency and Computation: Practice and Experience, vol. 24, 2012, pp. 1397-1420.
- [8] M. Grześ, and D. Kudenko. "Multigrid reinforcement learning with reward shaping." Artificial Neural Networks-ICANN 2008, 2008, pp.357–366.
- [9] A. Verma, P. Ahuja and A. Neogi, "pMapper: power and migration cost aware application placement in virtualized systems", Proc. 9th ACM/IFIP/USENIX International Conf. on Middleware, 2008, pp. 243–264.
- [10] M. Cardoso, M.R. Korupolu and A. Singh, "Shares and utilities based power consolidation in virtualized host environments", Proc. 11th IFIP/IEEE International Conf. on Symposium on Integrated Network Management, 2008, pp. 327–334.
- [11] A. Beloglazov, J. Abawajy and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing", Future Generation Computer Systems, vol. 28 , 2012, pp. 755–768.
- [12] M. Duggan, J. Duggan, E. Howley and E. Barrett, "A network aware approach for the scheduling of virtual machine migration during peak loads", Cluster Computing, vol 20, 2017, pp.1-12.
- [13] G. Tesauro, R. Das, H. Chan, J. Kephart, D. Levine, F. Rawson and C. Lefurgy, "Managing power consumption and performance of computing systems using reinforcement learning", Proc. 20th International Conf. on Neural Information Processing Systems, 2007, pp.1497–1504.
- [14] E. Barrett, E. Howley, and J. Duggan. "Applying reinforcement learning towards automating resource allocation and application scalability in the cloud", Concurrency and Computation: Practice and Experience, vol.25, 2013, pp.1656–1674.
- [15] E. Barrett, E. Howley and J. Duggan, "A learning architecture for scheduling workflow applications in the cloud". In Web Services (ECOWS), 2011 Ninth IEEE European Conference on, 2011, pp. 83-90. IEEE.
- [16] J. Rao, X. Bu, C. Z. Xu, L. Wang and G. Yin, "VCONF: A reinforcement learning approach to virtual machines auto-configuration", Proc. 6th International Conf. on Autonomic Computing, 2009, pp. 137–146.
- [17] R. Das, J. O. Kephart, C. Lefurgy, G. Tesauro, D. W. Levine and H. Chan, "Autonomic multi-agent management of power and performance in data centers", Proc. 7th International Joint Conf. on Autonomous Agents and Multiagent Systems: Industrial Track (International Foundation for Autonomous Agents and Multiagent Systems), 2007, pp.107–114.
- [18] X. Dutreilh, S. Kirgizov, O. Melekhova, J. Malenfant, N. Rivierre and I. Truck, "Using reinforcement learning for autonomic resource allocation in clouds: towards a fully automated workflow", Proc. 7th International Conf. on Autonomic and Autonomous Systems, 2011, pp.67–74.
- [19] F. Farahnakian, P. Liljeberg, and J. Plosila, "Energy-efficient virtual machines consolidation in cloud data centers using reinforcement learning." Parallel, Distributed and Network-Based Processing (PDP), 2014 22nd Euromicro International Conference on. IEEE, 2014.
- [20] R.S. Sutton, and A.G. Barto. Reinforcement learning: An introduction. Cambridge: MIT press, 1998.
- [21] R. Shaw. "An artificial intelligence model for autonomous resource allocation in cloud computing environments", Masters thesis. National University of Ireland Galway, 2016.
- [22] A.Y Ng, D. Harada and S. Russell, "Policy invariance under reward transformations: Theory and application to reward shaping", Proc. 16th International Conf. of Machine Learning, 1999, pp.278–287.

# Predicting Host CPU Utilization in Cloud Computing using Recurrent Neural Networks

Martin Duggan\*, Karl Mason, Jim Duggan, Enda Howley, Enda Barrett  
National University of Ireland, Galway  
Email\*: m.duggan1@nuigalway.ie

**Abstract**—One of the major challenges facing cloud computing is to accurately predict future resource usage for future demands. Cloud resource consumption is constantly changing, which makes it difficult for forecasting algorithms to produce accurate predictions. This motivates the research presented in this paper which aims to predict host machines CPU consumption for a single time-step and multiple time-steps into the future. This research implements a Recurrent Neural Network to predict CPU utilisation, due to their ability to retain information and accurately make predictions for time series problems, making it a promising candidate to predict CPU utilization with greater accuracy when compared to traditional approaches.

**Keywords**—Cloud Computing, CPU Prediction, Neural Networks

## I. INTRODUCTION

Armbrust et. al. have listed resource prediction as one of the ten biggest obstacles facing the continued growth of cloud computing [1]. One of the major difficulties for prediction algorithms in cloud computing is that cloud resources are in a constant state of flux. Traditional forecasting techniques such as for as ARIMA rely on patterns in historical data to make future predictions [19]. These approaches are not suitable when the data is not stationary or when there is a significant amount of random variation in the data. This paper uses a recurrent neural network to improve upon traditional forecasting techniques to make accurate time series prediction of host machines CPU utilization as they are much more adaptable and robust than these traditional approaches. CPU is the resource with the highest level of demands in virtualized environments and therefore is a major cause of resource shortages on host machines. CPU is one of the most important metrics for measuring the performance of host machines and is a popular metric for researchers to test when predicting host performance [28], [9], [6]. These studies have examined one-step ahead forecasting using methods such as LOESS and feed-forward Neural Network to predict CPU utilization. However, one step ahead prediction time models (usually 5 minutes ahead) give insufficient time for the cloud resources to be adjusted, when sudden heavy demands occur. Research has shown [5] predicting a workload on a short time scale such as 5 minute intervals is more difficult to produce accurate results than for long-term forecasting (i.e., time steps of days or weeks). This is due to the fact that cloud resources in these short time scales can be extremely unpredictable. The further into the future an algorithm can accurately predict the demand on data centre resources is critical to how well a data centre is able to perform. This is one of the key ideas that has motivated this research.

In recent years machine learning algorithms have received a lot of attention and are becoming popular to use in cloud computing. One of the most effective and diverse machine learning methods is the Neural Network [7], which is inspired by the brain. Neural networks act as function approximators which makes them widely applicable to a broad range of problems from regression to robotics. The Recurrent Neural Networks are of interest in this research due to their ability to retain information making it a promising candidate to predict CPU utilization with greater accuracy when compared to traditional approaches.

In this paper, we predict host CPU utilization using Recurrent Neural Networks. The aims of this research are to:

- 1) Investigate the accuracy of a Recurrent Neural Network for predicting CPU utilization when compared to traditional methods.
- 2) To determine how far into the future the Recurrent Network can accurately predict host CPU utilization.

The outline of the paper is as follows. Section II gives an overview of forecasting in cloud computing, and neural networks. The experimental procedure will be explained in Section III. Section IV will present the experimental results. These results will then be discussed in Section V. Finally, Section VI will conclude the paper.

## II. RELATED WORK

Cloud resources such as CPU are in a constant state of flux and are difficult to predict on a short time scale (e.g. 20-30 minutes). Approaches such as one-step-ahead prediction give very little time for the data centre to re-adjust resource required when bursts of high traffic occur. An algorithm that can produce accurate prediction 20 to 30 minutes into the future could inform the data centre management systems to perform suitable actions such as turn host machines on/off to deal to deal with future demands. The objective of this paper is to use a recurrent neural network to predict host machines CPU utilization with a high degree of accuracy.

### A. Forecasting in Cloud Computing

Host machine CPU is one of the most studied metrics when it comes to performance, as it is a major cause of resource shortage. Dinda and O Hallaron used different linear forecasting models to predict tasks running times, based on CPU load predictions [9]. Zhang et al. employed a multi-step ahead CPU load prediction approach for grid tasks to predict the future performance of the resources [28].

Recently there has been a move towards integrating Artificial Intelligence (AI) and Machine Learning (ML) techniques to improve the overall efficiency of a cloud data centre. Several works show how AI and ML algorithms can provide cloud systems with the abilities to better adapt to the changes in cloud resource consumption to improve resource scaling, VM live migration [13], [11], [10], [12], [25] and resource allocation [3], [2] in cloud computing. Neural networks are one of the most effective and versatile machine learning algorithms and have been successfully applied to areas of cloud computing such as scheduling [14], intrusion detection [26], DDoS attack defence [18] and load forecasting [24]. Neural networks have previously been used to forecast resource demands in cloud computing. Duy et al. employ a neural network predictor for optimising server power consumption in a data centre [14]. They use a feed-forward neural network to predict future load demands based on historical demands to turn on/off servers to minimise the energy usage. Prevost et al. implement neural networks and a linear predictor algorithms to forecast future workloads [24]. Bey et al. use several different models for time series prediction. They use an adaptive network to estimate the future value of CPU load for distributed computing [6]. However, their hybrid predictors were designed to perform for one-step-ahead prediction and the work presented in this paper builds on this work by predicting both one-step and multi-steps ahead. All of the research highlighted above outlines how neural networks are effective at addressing many of the problems in cloud computing, in particular, CPU forecasting. The research presented in this paper makes the novel contribution of applying Recurrent Neural Networks for CPU forecasting.

**B. Neural Networks**

Neural Networks are function approximators that are inspired by the biological neural networks that constitute the human brain [7]. Some of the applications of neural networks include: power generation [21], control [22] and watershed management [23]. Fig. 1 illustrates the architecture for a neural network which is arranged in a number of layers. The input layer is responsible for taking in the inputs to the model, the hidden layer is where the vast majority of the computation is done and the output layer produces the output of the model.

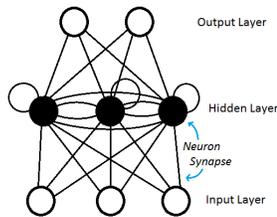


Fig. 1. Recurrent Neural Network [21]. This figure illustrates a Recurrent Neural Network. Neurons are connected weighted synapses that pass signals between neurons. The recurrent synapses can be seen in the hidden layer of neurons. This gives the recurrent network the abilities to retain information.

The standard feed-forward network consists of an input layer of neurons, one or multiple hidden layers of neurons and an output layer. The neural networks receive information in a form of a signal (normalised between 0 and 1) through the input layer neurons and then outputs a signal using the sigmoid function. The signal or input that the network receives in this

paper is in the form of two CPU utilisation values from a host machine (normalised between 0 and 1). This two CPU input is propagated forward through the hidden layers of neurons via synapses (weighted connections). Then the network calculates an output at the output layer neuron or neurons. In this paper, only one output is needed and the output signal corresponds to a future CPU values of a host machine. An error signal is calculated by finding the difference between the actual CPU value and the predicted value. This error is then propagated back through the network, and the weights (synapses) are adjusted to correct the error of the prediction.

Aside from the input layer, a neuron in any other layer will have as input. The sum of the weighted signals that are outputted from other connected neurons. A neurons input signal is described by Equation 1.

$$v_j = \sum_{i=1}^N w_{i,j} a_i \tag{1}$$

Where  $v_j$  is the input to a neuron in the  $j^{th}$  layer, layer  $i$  is the preceding layer to  $j$  that contains  $N$  neurons, each neuron in layer  $i$  has output  $a_i$  and each of these output signals are weighted by the value  $w_{i,j}$  as they are passed to each neuron in layer  $j$ .

Each neuron  $a_i$  outputs a value between 0 and 1. This output value is determined by the activation function of the neuron. The most commonly used activation function is the sigmoid function. This is described by Equation 2

$$a_j = \frac{1}{1 + \exp -v_j} \tag{2}$$

This research will implement a Recurrent Neural Network, illustrated in Figure 1. Recurrent networks are different from the standard feed-forward networks as the hidden layers neurons have recurrent connections. These connections allow the hidden layer neurons to connect to itself. Thus, giving the neural network memory of previous predictions which makes it well suited to the problem of predicting CPU demand. The recurrent network in this paper is trained using the popular Back-Propagation-Through-Time (BPTT) algorithm [27]. The idea of BPTT is the unfolding of the recurrent neural network at a discrete-time into a multilayer feedforward neural network each time a sequence is processed. The BPTT is different to the feed-forward neural networks as it enables the recurrent neural network to store past information, thus suitable for sequential models. The standard feed-forward algorithm is updated as follows:

$$v_j = \sum_{i=1}^N w_{i,j} a_i + \sum_{h=1}^m s_h(t-1) u_{ih} \tag{3}$$

where  $U$  is the recurrent weight matrix and  $s_h(t-1)$  is the previous hidden layer

A normal back propagation algorithm weights are updated by calculation the cost function (the error from the actual answer and predicted answer)

$$C = 1/2 \sum_{p=1}^k \sum_{e=1}^o (d_{pk} - y_{pk})^2 \quad (4)$$

where  $d$  is the desired output,  $k$  is the total number of training samples and  $o$  is the number of output units. Then the change in weights for the output nodes can be calculated as:

$$\delta_{pk} = (d_{pk} - y_{pk})g'(net_{pk}) \quad (5)$$

where  $g$  is the activate function where  $net$  represent inputs. The changes in weights for the hidden layer weight can be represented as:

$$\delta_{pj} = \sum_{k=1}^o \delta_{pk} w_{kj} g'(net_{pj}) \quad (6)$$

Therefore, the recurrent weights can be then back-propagated back through the network:

$$\Delta u_{ih} = \sum_{p=1}^N \delta_{pj} s_{ph}(t-1) \quad (7)$$

### C. Network Parameter Selection

The reason for using the recurrent neural network to predict CPU utilization over a feed-forward neural network is due to their ability to retain information and accurately make predictions for time series problems. This makes it a promising candidate for predicting CPU utilization with greater accuracy when compared to traditional approaches.

The recurrent neural network used in this research has three hidden neuron in the hidden layer and has two inputs from the input layer. The inputs into the network are the current and previous CPU utilization values. When a parameter sweeps was conducted, it showed that that a network with three hidden neurons produced the greatest performance. The parameter sweeps also highlighted that having greater than two inputs of CPU utilisation did not increase the recurrent networks performance. The network had one output that corresponded to the network's prediction of future CPU utilization.

### D. Network Training

The recurrent neural network algorithm will be trained over 10,000 evaluations and will be evaluated on unseen test data. The experiments are repeated over 10 runs to ensure statistically significant results.

## III. EXPERIMENT DETAILS

### A. Data Models

The data-set used to train and test the recurrent neural network comprised of CPU utilization that was generated by the CoMon project, a monitoring infrastructure for PlanetLab. The project contains CPU utilization data, which was obtained from more than a thousand VMs from 500 data centres around the world. There are ten folders worth of workloads, containing CPU utilization values measured every five minutes in VMs. Each file contains 288 values. We ran the CloudSim simulator using the Lr-mmt algorithm to generate hosts CPU values [8]. In the cloudsim simulation, over 800 host are used. We selected

host number 3's CPU values for each of the ten days worth of workloads for our experiments. The first nine workload traces (containing 2296 CPU values data-set) were used to train the recurrent network and the tenth workload (containing 288 CPU values data-set) to test the network. The reason for using these planetlabs files is that they have proven to be useful CPU workflow data-sets when conducting experiment on simulated cloud host machines [12].

### B. Comparative Forecasting Methods

The recurrent neural network will be compared to the following methods:

- 1) Back-propagation (BP).
- 2) Random walk forecasting (RWF).
- 3) Moving Average (MA).

The Back-propagation (BP) algorithm works by calculating the error between the target output and the observed output. This error is then propagated back through the network and is used to update the weights. BP is different to BPTT as it does not store any memory. In this research the BP network has 2 inputs, 3 hidden neurons and 1 output, keeping consistent with the recurrent network's implementation. Random walk forecasting is a basic forecasting method that is implemented as a benchmark algorithm. This approach predicts the next future value as equal to the currently observed value. The moving average method is another commonly using forecasting approach. This method consists of predicting a future value by averaging  $n$  previous values. In this paper, the two previous times steps were averaged to give a future prediction.

### C. Experiments Conducted

There will be three experiments conducted in this paper. The first experiment involves comparing all of the algorithms and methods performances on the training data. The second experiment will evaluate the performance of each algorithm on the testing data. The purpose of this experiment is to examine if the trained recurrent networks (BPTT) are capable of giving a good general performance on data it has not seen. The third experiment will evaluate how far into the future the BPTT network can predict. The results of these experiment will be interesting as it would be beneficial to data centres management systems to in advance how much CPU is used on each host before events such as live migration can occur. This experiment will evaluate the accuracy of the network for predicting CPU utilization further than one step into the future.

## IV. RESULTS

This section presents the results of each of the experiments outlined above followed by a discussion in order to highlight their significance for real world data centre challenges.

### A. Training Data

Figure 2 shows the convergence of both the BPTT and BP trained networks on the training data. This graph highlights the average Mean Absolute Error (MAE) at each time step. The graph shows that BPTT converges to a better solution than BP and highlights that BPTT found a better solution faster also. One reason being that the recurrent network can store memory

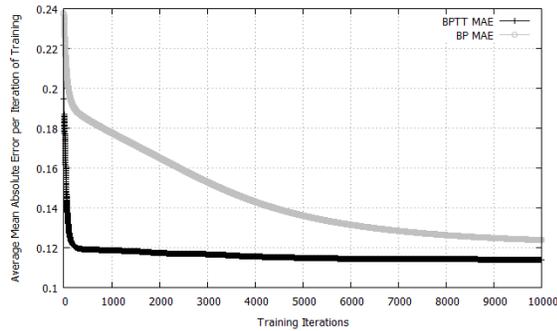


Fig. 2. Convergence of BPTT and BP. This figure illustrates the MAE at each iteration of training for Back-propagation-through-time and Back-propagation neural networks.

from the previous time step, allowing the algorithm to achieve a lower error and produce better predictions than the standard feed-forward neural network.

Table 1 presents the results for all forecasting methods in this paper. The Mean Absolute Error (MAE) and Mean Squared Error (MSE) were used to determine how accurate each forecasting method performed on the training data. From the results in Table I, the random walk, moving average and back-propagation neural network performs significantly worse than BPTT algorithm. This validates the choice of using a recurrent neural network trained for host CPU utilisation prediction. When comparing the performances of each method results reveal that MA was the worst performing for the MAE and MSE on the training data, with BPTT producing the best results.

TABLE I. TRAINING DATA ACCURACY

Algorithm	MAE (Std Dev)	MSE (Std Dev)
BPTT	0.1162 (0.001)	0.0219 (0.0003)
Random Walk	0.1427 (0.00)	0.0354 (0.0000)
Moving Avg	0.1492 (0.00)	0.0367 (0.0000)
Backpropagation	0.1301 (0.01)	0.031 (0.004)

### B. Test Data

The second experiment conducted involved evaluating how well the recurrent neural network can predict unseen test data from the same host machine data that it was previously trained on. Figure 3 plots the prediction of the BPTT and the actual CPU utilisation. The graph shows that BPTT predicts accurately for values between 0.01 and 0.8, however, it struggles to predict values higher than this threshold. One reason for this being that the data has a sudden variation of CPU utilisation. Another reason being that the implementation of the recurrent neural network in this paper only hold the previous step CPU value, if the algorithm stored a longer sequence of data the prediction potentially could improve future predictions.

Table II presents how accurate each of the forecasting methods is when evaluated on the test data. As with the training data, the recurrent BPTT trained network performed the best. Random walk performed significantly worse on the testing data than BP and moving average. The back-propagation algorithm performed worst on both training and testing data when compared to BPTT. This shows that the network trained using BP does not generalise well to time series data.

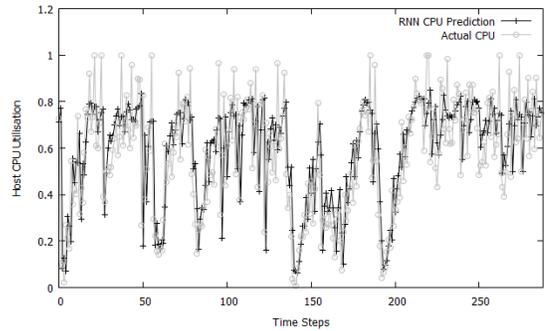


Fig. 3. Host Utilization Predictions for Test Data. This figure illustrates the predicted host utilization of the recurrent neural network on unseen test data.

The results in Table II highlight that the standard deviations for the recurrent network are higher than for the training data. However, this is to be expected as it was being evaluated on previously unseen data so more deviation in the prediction accuracy is to be expected.

TABLE II. TEST DATA ACCURACY

Algorithm	MAE (Std Dev)	MSE (Std Dev)
BPTT	0.1498 (0.001)	0.0468 (0.0008)
Random Walk	0.1716 (0.0000)	0.0513 (0.0000)
Moving Avg	0.1587 (0.0000)	0.0428 (0.0000)
Backpropagation	0.160 (0.010)	0.041 (0.004)

### C. Multi-step Ahead Prediction

The aim of the multiple time steps ahead prediction experiment was to evaluate how far into the future the neural network could predict CPU utilisation and to establish how much the accuracy of the prediction decreases the further into the future the network attempts to predict. Since BPTT trained network had the best accuracy in both the training and testing data sets, this recurrent neural network was implemented to predict CPU utilisation for multiple steps prediction. This experiment involved predicting the CPU utilisation of a host machine at 1, 2 and 3-time steps into the future. Each of these time steps corresponds to 5 minutes (i.e. the aim was to predict 15 minutes into the near future of CPU utilisation with high accuracy).

Table III presents the accuracy of the prediction at each of the first, second and third time steps. As presented in the table the further into that the future the recurrent network predicts, the accuracy decreases linearly. This is true for both the training and testing data as present in Table III.

TABLE III. MULTI-STEP PREDICTION ACCURACY

Number of Steps	Training		Test	
	MAE (Std Dev)	MSE (Std Dev)	MAE (Std Dev)	MSE (Std Dev)
1 Step Ahead	0.116 (0.001)	0.021 (0.0003)	0.153 (0.008)	0.038 (0.003)
2 Step Ahead	0.133 (0.003)	0.029 (0.001)	0.167 (0.015)	0.046 (0.008)
3 Step Ahead	0.146 (0.0008)	0.035 (0.0001)	0.212 (0.018)	0.066 (0.009)

Figure 4 displays the MAE for both the 1 and 3 steps into the future on the test data. The reason for time step 2 been omitted is for clarity and readability of the graph. This graph highlights the time steps the recurrent network's accuracy is performing the greatest and where the error in predictions are

the highest. Figure 4 shows when there are sudden changes in the host machine CPU utilisation the larger the error is in the accuracy of prediction. For instance consider time step 50 in Figure 3. The actual CPU utilisation values show a sudden decrease from 0.9 to 0.2. During this period, Figure 4 shows that there was a sharp increase in the error from the recurrent network prediction at each of the time steps. One reason for the sudden decrease in prediction accuracy is that the network find it difficult to perform well when extreme changes occur in CPU utilisation.

Another observation highlighted from Figure 4 is the difference between the prediction error of the 1 and 3 step ahead predictions on the testing data. Examining time steps 200 to 288 in Figure 3 shows the actual CPU is constant with little sudden changes in CPU utilisation. Figure 4 highlights that the one step ahead predictions produces better results. The third time step ahead prediction errors shows how difficult it is to produce accurate results with a noisy data set used in this experiment. Considering how well the recurrent neural network performed in the multi-time step ahead predicts, from the MAE and MSE results from predicting two-time steps ahead shows that it out performed the random walks results even when that algorithm was only predicting one-time step ahead. The overall average mean squared error for each time step was 0.038, 0.046 and 0.066. These results show a steady increase in the error the further out the network tries to predict.

## V. DISCUSSION

The results of the experiments show that the recurrent neural network has the capabilities to improve upon traditional prediction methods such as random walk, moving average and Back-propagation to predict CPU utilisation with a high degree of accuracy. This is shown both for the results for one-step and multi-step prediction. The first experiment conducted determine how a recurrent neural network could outperform tradition forecasting methods. The results indicate that even with a large amount of noise in the CPU utilisation data the recurrent neural network could produce accurate results on the training data-set compared to the traditional prediction methods. The aim of the second experiment was to test the how well the recurrent neural network could perform on previously unseen data. Again shown in the results the recurrent neural network provided the best prediction accuracy. The third experiment conducted examined how far into the future the recurrent network could predict with a high degree of accuracy. The results indicate that recurrent neural network can produce a reasonable degree of accuracy when predicting multiple time steps into the future. Forecasting multi-time steps ahead for cloud resource has proven to be a difficult area in time series research. The recurrent neural network presented in this research could potentially be integrated with many areas of cloud computing such as host migration and VM scheduling to improve overall performance. For instance, research has shown that instantiating a new virtual machine takes between 5-15 minutes [17]. The results presented in this paper demonstrate that recurrent neural networks are capable of predicting CPU utilisation 15 minutes into the future and still retain a relatively high degree of accuracy. The recurrent neural network could inform the cloud management system when a host is going to become over-utilized so appropriate actions such as live migration or boot up a new VM instances potentially could

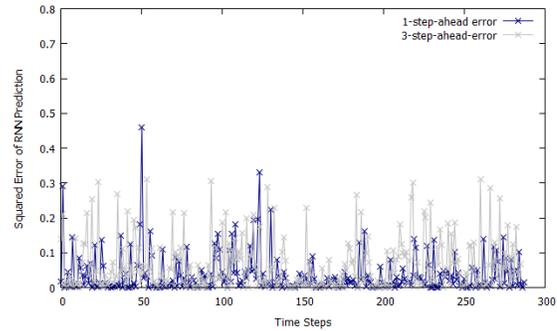


Fig. 4. The MAE of each of the multi-step ahead prediction. This figure illustrates the error of the prediction of CPU utilization for 1-time-step-ahead and 3-time-steps-ahead for the recurrent network.

be initiated prior to a host becoming over-utilized. Power consumption is another key area in cloud computing where accurate forecasting algorithms could enhance the performance of cloud data centre. Gartner et al. highlighted that the ICT industry contributed to about 2% of global CO<sub>2</sub> emitted each year, aligning itself on the same level with the aviation industry [16] Koomey has stated that in 2010 1.3% of all power consumed worldwide was due to data centre usage [20] and this was increasing each year. Cloud data centres integrating more advanced AI and ML algorithm could reduce energy consumption. Energy is constantly being wasted on a substantial portion of the host machines that operate at 10-50% of their full capacity [4], which in-turn results in a significant increase in energy costs. The results presented by Duy et al. have shown how neural networks can be utilized as a predictor to reduce energy consumption in a data centre to turn off host when the traffic load is light [14]. The results in this paper demonstrate that recurrent neural networks can provide more accurate forecasts, therefore having additional benefits of reducing the overall energy consumption of the data centre. For instance, data centres that have a portion of it host machines operating between 0-10% utilization could be predicted by a recurrent neural network for the next 20-30 minutes and shut down these machines to reduce energy consumption and by extension decrease CO<sub>2</sub> emissions from powering the cloud data centres. Recently, companies such as Google have implemented their own DeepMind neural network tool to reduce their data centre energy cost by 40% [15].

## VI. CONCLUSION

The aim of this research was to investigate if recurrent neural networks are capable of accurately predicting CPU utilization for short time periods. The results from this paper indicate that it is possible to predict CPU utilization with a high degree of accuracy for data sets that have sudden extreme changes. The recurrent neural network train with BPTT was able to accurately predict CPU utilization within 10,000 evaluations of the training data. The recurrent neural network performed best on both training and testing data when compared to tradition prediction methods such as random walk, moving average and backpropagation. Results show however that the prediction of the CPU utilization is a difficult task due to the occasional sudden extreme change in CPU utilization. The results also highlight that the recurrent neural

networks prediction accuracy decreases as it predicts further into the future. However, on average the network is capable of predicting with a reasonable level of accuracy 3 steps (15 minutes) into the future. In summary, the contributions of this research are:

- 1) Recurrent neural networks have the capabilities to accurately predicting noisy host CPU utilization.
- 2) The Recurrent neural networks produce relatively high accuracy when predicting 15 minutes into the future. The accuracy of the network predictions decreases in a linearly the further into the future the network attempts to predict.

#### A. Future Work

There are several potential routes for future research that have arisen from this research, which will include using different algorithms such as Long-Short-Term-Memory to train the recurrent neural network to compare and improve the accuracy of the predictions of back-propagation-through-time. Other interesting future work would include predicting other metric such as RAM and disk utilisation of a host machine.

#### REFERENCES

- [1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, et al. A view of cloud computing. *Communications of the ACM*, 53(4):50–58, 2010.
- [2] E. Barrett, E. Howley, and J. Duggan. A learning architecture for scheduling workflow applications in the cloud. In *Web Services (ECOWS), 2011 Ninth IEEE European Conference on*, pages 83–90. IEEE, 2011.
- [3] E. Barrett, E. Howley, and J. Duggan. Applying reinforcement learning towards automating resource allocation and application scalability in the cloud. *Concurrency and Computation: Practice and Experience*, 25(12):1656–1674, 2013.
- [4] L. A. Barroso and U. Hözlze. The case for energy-proportional computing. *Computer*, 40(12), 2007.
- [5] T. Benson, A. Anand, A. Akella, and M. Zhang. Microte: Fine grained traffic engineering for data centers. In *Proceedings of the Seventh Conference on emerging Networking EXperiments and Technologies*, page 8. ACM, 2011.
- [6] K. B. Bey, F. Benhammadi, A. Mokhtari, and Z. Guessoum. Cpu load prediction model for distributed computing. In *Parallel and Distributed Computing, 2009. ISPD'09. Eighth International Symposium on*, pages 39–45. IEEE, 2009.
- [7] C. M. Bishop. *Neural networks for pattern recognition*. Oxford university press, 1995.
- [8] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. De Rose, and R. Buyya. Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience*, 41(1):23–50, 2011.
- [9] P. A. Dinda and D. R. O'hallaron. Host load prediction using linear models. *Cluster Computing*, 3(4):265–280, 2000.
- [10] M. Duggan, J. Duggan, E. Howley, and E. Barrett. A reinforcement learning approach for the scheduling of live migration from under utilised hosts. *Memetic Computing.*, Volume 8.(DOI 10.1007/s12293-016-0218-x.):pp 111., 09 December 2016.
- [11] M. Duggan, J. Duggan, E. Howley, and E. Barrett. An autonomous network aware vm migration strategy in cloud data centres. In *Cloud and Autonomic Computing (ICCAC), 2016 International Conference on*, pages 24–32. IEEE, 2016.
- [12] M. Duggan, J. Duggan, E. Howley, and E. Barrett. A reinforcement learning approach for dynamic selection of virtual machines in cloud data centres. In *Conference: Innovative Computing Technology (INTECH 2016)*, 2016.
- [13] M. Duggan, J. Duggan, E. Howley, and E. Barrett. A network aware approach for the scheduling of virtual machine migration during peak loads. *Cluster Computing*, pages 1–12, 2017.
- [14] T. V. T. Duy, Y. Sato, and Y. Inoguchi. Performance evaluation of a green scheduling algorithm for energy savings in cloud computing. In *Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW), 2010 IEEE International Symposium on*, pages 1–8. IEEE, 2010.
- [15] J. Gao and R. Jamidar. Machine learning applications for data center optimization. *Google White Paper*, 2014.
- [16] G. Inc. Gartner estimates ict industry accounts for 2 percent of global co2 emissions. In <http://www.gartner.com/newsroom/id/503867>, 2007.
- [17] S. Islam, J. Keung, K. Lee, and A. Liu. Empirical prediction models for adaptive resource provisioning in the cloud. *Future Generation Computer Systems*, 28(1):155–162, 2012.
- [18] B. Joshi, A. S. Vijayan, and B. K. Joshi. Securing cloud computing environment against ddos attacks. In *Computer Communication and Informatics (ICCCI), 2012 International Conference on*, pages 1–5. IEEE, 2012.
- [19] M. Khashei, M. Bijari, and S. R. Hejazi. Combining seasonal arima models with computational intelligence techniques for time series forecasting. *Soft Computing*, 16(6):1091–1105, 2012.
- [20] J. G. Koomey. *Estimating total power consumption by servers in the US and the world*. February, 2007.
- [21] K. Mason, J. Duggan, and E. Howley. Evolving multi-objective neural networks using differential evolution for dynamic economic emission dispatch. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 1287–1294. ACM, 2017.
- [22] K. Mason, J. Duggan, and E. Howley. Neural network topology and weight optimization through neuro differential evolution. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pages 213–214. ACM, 2017.
- [23] K. Mason, J. Duggan, and E. Howley. A meta optimisation analysis of particle swarm optimisation velocity update equations for watershed management learning. *Applied Soft Computing*, 2017 (In Press).
- [24] J. J. Prevost, K. Nagothu, B. Kelley, and M. Jamshidi. Prediction of cloud data center networks loads using stochastic and neural models. In *System of Systems Engineering (SoSE), 2011 6th International Conference on*, pages 276–281. IEEE, 2011.
- [25] R. Shaw, E. Howley, and E. Barrett. Predicting the available bandwidth on intra cloud network links for deadline constrained workflow scheduling in public clouds, 01 2017.
- [26] K. Vieira, A. Schultze, C. Westphall, and C. Westphall. Intrusion detection techniques in grid and cloud computing environment. *IT Professional, IEEE Computer Society*, 12(4):38–43, 2010.
- [27] P. J. Werbos. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10):1550–1560, 1990.
- [28] Y. Zhang, W. Sun, and Y. Inoguchi. Predict task running time in grid environments based on cpu load predictions. *Future Generation Computer Systems*, 24(6):489–497, 2008.

## Sessions

## **Session 1: Information Security**

Title: Towards a Security and Privacy Co-Creation Method  
(Authors: Christophe Feltus, Erik HA Proper)

Title: Secured Transactions Technique Based on Smart Contracts for Situational Awareness Tools  
(Authors: Roman Graf, Ross King)

Title: Speech Encryption Based on Hybrid Chaotic Key Generator for AMR-WB G.722.2 Codec  
(Authors: Messaouda Boumaraf, Fatiha Merazka)

Title: A Blind Watermarking Technique based on DCT Psychovisual Threshold for A Robust Copyright Protection  
(Authors: Ferda Ernawan, Muhammad Nomani Kabir, Zuriani Mustafa)

Title: Data Protection by Design in Systems Development From legal requirements to technical solutions  
(Authors: Fredrik Blix, Salah Addin Elshekeil, Saran Laoyookhong)

# Towards a Security and Privacy Co-Creation Method

Christophe Feltus, Erik HA Proper

Luxembourg Institute of Science and Technology,  
5, avenue des Hauts-Fourneaux, L-4362 Esch-sur-Alzette, Luxembourg  
{firstname.name}@list.lu

**Abstract** — Cyber collaboration supports and increases the expansion of value co-creation amongst companies and customers by defining innovative business models and by exploiting new types of infrastructures like those dedicated to social media, collaborative workspaces, or e-supply chains for instance. This proliferation of new types of collaboration generates new types of security and privacy threats to be handled by the companies. The deployment of the appropriate controls to cope with the latter is of great value for the continuity of the day to day business. Therefore, in this paper, we investigate how security and privacy may be regarded as types of value and how they may be considered, in collaborative environments, through the lens of value co-creation. Acknowledging the similarities between security, privacy, and value, we afterwards propose a method to co-create security and privacy and we illustrate how the latter may be deployed in the frame of a financial case-study.

**Keywords:** Security; privacy; method; value co-creation; security co-creation; privacy co-creation; enterprise collaboration.

## I. INTRODUCTION

Cyber collaboration supports and increases the expansion of value co-creation (VCC) amongst companies and customers by exploiting new types of infrastructures like those dedicated to social media, collaborative workspaces, or e-supply chains. The amount and the complexity of these collaborations is at the origin of new types of security breaches which give room to new types of viruses like the ransomware that, according to Kharraz et al. [1], represents forms of cyber-attack hardly resolvable. As cyber collaboration and the resulting VCC is at the origin of new threats, and because the deployment of the appropriate controls to cope with the latter is of great value for the continuity of the business, we propose to investigate, in this paper, how security and privacy may potentially be handled through the lens of value co-creation. Practically, this designed decision is grounded on the motivation that considering security and privacy co-creation (SPCC) may be examined as a specialization of VCC [2]. This assertion is justified by the acknowledgement that value is an abstract concept [3] which expresses a measurable information, of a determined nature, and which represents an *assessment of benefits against sacrifices* [4]. Similarly, the discipline of security and privacy also shared this statement that both represent costs for the company but, in return, generate benefit in terms of protection of their information system (IS).

Unfortunately, despite a plethora of research aiming at

depicting the fundamental of VCC (e.g., VCC concepts, value in use, value in exchange, etc.), few contributions have been poured in the area of methods for V/SPCC design and deployment. Therefore, in this paper, we propose an innovative approach to support the VCC of a security and privacy nature and that is related to assets shared between two partners. This method is a four steps approach which is based on the three dimensions of the value co-creation model from [2]: nature of the value, method of VCC, and object concerned by VCC.

The next section reviews the related works regarding SPCC and reminds previous works related to VCC. In section III we present the security co-creation method and in section IV we illustrate it through a case study in the financial domain. Section V concludes and discusses the proposed approach.

### Running case study

In the financial sector, a retail bank sells assets to its customers and stores and backups the business information in a data center. To monitor the level of privacy, this bank performs regular privacy impact assessments (PIA). In parallel, to monitor the security of the service delivered, the bank's data center performs security GAP analyses (SGAP) that allow estimating the level of compliance between the real level of security and the expected one. Fig. 1, modeled with the e3value language [5], illustrates the exchange of value in and between both stakeholders (blue links). For the time being, the only exchange of value between both consists in the storage of data for the bank and in the money paid to the data center for the storage. The security and privacy co-creation method aims to propose an approach to discover complementary value co-creation in the fields of privacy and security.

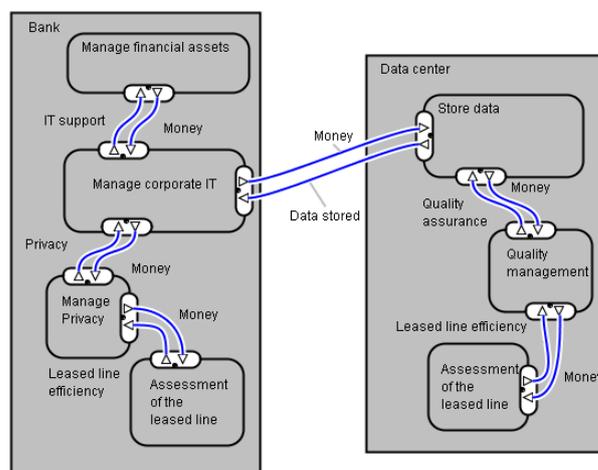


Figure 1. Security and privacy without co-creation - e3value model

The research is supported by the National Research Fund, Luxembourg (<http://www.fnr.lu>) and financed by the ValCoLa (VCC Language) project.

## II. RELATED AND PREVIOUS WORKS

This section first reviews the literature related to the field of security and privacy co-creation and collaborative security, and in the field of VCC more generally, to give an insight on how security, privacy and value co-creation may similarly be handled. Afterwards, the section reminds previous works related to value modeling and to methods of value co-creation.

### A. Literature review

As highlighted by Vicini et al. [6], the challenge of security co-creation is twofold: first, to extract the value of the enormous amount of data available in distributed environment, and second, to improve the perception that these data are handled by a trusted system to store privacy protected content. This challenge is especially important when end-users are directly engaged in the co-creation process [7]. Vicini et al. show how it is possible to integrate practical co-creation processes into security and privacy by design methodologies and propose a methodology and guidelines to translate high-level requirements into verifiable low level and technological ones. In [8], Bennaceur et al. address the support of collaborative security in the field of internet of things and explain how the collaborative security tends to exploit and to compose the capability of the connected device to protect assets from potential harm. The authors propose an approach supported by a dedicated tool to support the above composition using a combination of feature modelling and mediator synthesis. In [9], Martin et al. stress the importance of the collaborative approach to security management in the area of air traffic management, due to the fact that operations and systems become increasingly integrated. Accordingly, they claim that for a successful collaborative approach, security managers need to adopt collaborative leadership skills and approaches. More recently, in [11], Garrido-Pelaz et al. propose a collaborative security approach through the perspective of information sharing which can help to develop early prevention mechanisms. Therefore, they exploit a model for sharing cybersecurity information between dependent organizations that are impacted by different cyber-attacks.

SPCC could be seen as a type of value co-creation. VCC discipline originates from the marketing theory. It aims to define and to explain the mechanisms for the co-generation of value during business exchanges amongst companies [11]-[12]. Vargo et al. [12] [13] formalize it using a framework for defining VCC in the perspective of the service dominant logic (S-DL). According to them, service is the *basis of all exchanges and focuses on the process of value creation rather than on the creation of tangible outputs*. As a result, a service system is a *network of agents and interactions that integrates resources for VCC* [12]. On that basis, value is proposed by a service provider and is determined by a service beneficiary. According to [14], this interaction is defined through situations in which the customer and the provider are involved in each other's practices. Frow et al. [15] propose a framework to assist firms in identifying new opportunities for VCC. Therefore, they provide a strategically important new approach for managers to identify, organize and communicate innovative opportunities. More recently, Chew [16] argues that, in the digital world, service innovation is focused on customer value creation and he proposes an integrated Service Innovation

Method (iSIM) for analyzing the interrelationships between the design process elements. At the IS domains level, Gordijn et al. [5] explain that business modeling is not about process but about value exchange between different actors. Accordingly, Gordijn et al. propose e3value to design models that sustain the communication between business and IT groups. In [17], e3value is extended for considering co-creation. Therefore, the authors define the so called *value encounters* which consist in spaces where groups of actors interact to derive value from the groups' resources. The financial case used to illustrate our method is modelled with this e3value language (Fig. 1 and 9). In the same vein, Razo-Zapata et al. propose visual constructs to describe the value co-creation process [18].

### B. Value modeling

In our previous work [2], one first contribution consisted in a value creation model structured according to three dimensions (Fig. 2): the nature of the value, the method of value creation, the object concerned by the value.

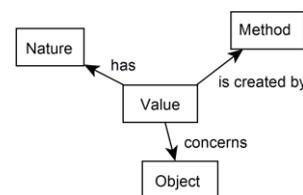


Figure 2. Three value dimensions

These three dimensions of the value creation are defined as:

- **Nature of the value.** The value *has* a nature that expresses a domain of interest and a context that characterize an element of the information system. E.g., security of the IS, actor's responsibility [19], or the data privacy [20].
- **Method (to create value).** The method is an abstract concept that gathers a set of method elements ordered in steps and achieved in order to *create* value. E.g., a process based approach, a risk assessment, the method chunk [21].
- **Object (concerned by the value).** The object *concerned* by the value is the IS element that is better after this value being delivered. E.g., an actor, a process, a data, a server.

In the following, we explain the value creation model and propose three fundamental value co-creation schemas. Based on combinations amongst the latter, more complex VCC schemas may also be designed (e.g., by considering more than one dimension, or for tackling co-creation implying more than two actors). These combinations are not considered in the paper but are available in [2]. The value creation model presented in Fig. 3 includes nine additional concepts which are dedicated to express the three value creation dimensions.

The nature of the value has characteristics that define the value, the latter concerns an object, is created by a method, and is measurable:

- **Characteristics of the Nature of the Value.** This concept expresses the different elements that characterize the nature of the value, or the pillars that found this nature (e.g., availability, confidentiality, portability, etc.).
- **Object.** The object *concerned* by the value is the IS element that is better after this value being delivered. (e.g., an actor, a process, a data).

- **Measure.** The measure corresponds to a property on which calculations can be made for determining the amount of value generated.

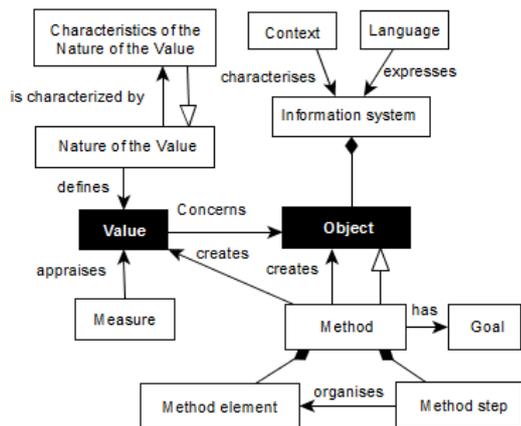


Figure 3: Value creation (VC) model

The method of value creation has a goal, is composed of method elements organized by method steps:

- **Goal.** The goal corresponds to the expected operation on value created by the method (e.g., create value, assess or evaluate value generated, optimize the value).
- **Method element.** The elements of the method correspond to unitary tasks that constitute the method. (e.g., analysis, collection of information, reporting...)
- **Method step.** The method steps consist in the organized and coherent articulations of the method elements (e.g., if-then-else, process elements ordination...)

The objects concerned by the value are impacted by the method. They composed the information system which is characterized by a context and expressed by a language:

- **Information system.** The information system that encompasses the objects concerned by the value.
- **Context.** The context represents the information surrounding of the IS (e.g., the sector of the business entity that is concerned by the IS, the rules and regulations related to this sector, etc.)
- **Language.** The language represents the vocabulary used to express the information system of a specific context.

The second contribution of [2] consists in three generic schemas of VCC, built upon the three dimensions of value creation model: nature of the value, method of value creation, and object concerned by the value. The three generic schemas are: (1) **Method-based VCC:** In this first schema, the method is shared by the companies but the nature of the value and the object of value created are different. In this co-creation case, VCC activities achieved by two companies may generate different types of value nature, concerning different objects evolving in different contexts. As a result, the co-creation described in these first schemas happens because enterprises share and achieve activities together that contribute to value creation (e.g., two companies that create value using a shared process-based approach). (2) **Object-based VCC:** This co-creation concerns a unique object that creates value of different natures in different contexts. It concerns two companies that collaborate to co-create value but this value may be of different nature for each of them (e.g., two

companies that create two different nature of value for the benefit of a joint network). (3) **Nature-based VCC:** In this third schema, the nature of the value co-created is shared by the companies but the object of value created and the value creation method are different. The VCC activities at the level of each company may be achieved by using different methods and may concern different types of objects from different contexts. However, these different activities concern VCC of the same nature (e.g., two companies protecting the privacy of their customers with different methods).

### III. SECURITY AND PRIVACY CO-CREATION METHOD

#### A. Security and privacy co-creation

Security co-creation is an important research topic [6]-[10]. In this paper, we investigate security and privacy co-creation as an instance of value co-creation. Indeed, security and privacy are characteristics of elements of the information system that, when adequately deployed, improve the utilization of the latter. Both security and privacy, according to [2], are themselves defined by the following characteristics: availability, confidentiality, integrity, non-repudiation, etc. (for security) and anonymity, pseudonymity, access to resources, etc. (for privacy). Finally, alike all nature of value, security and privacy are also created by dedicated methods (like risk assessment, cryptography, packet filtering, etc.)

#### B. Security and privacy co-creation method

Based on the three value dimensions and the three co-creation methods presented in Section II.B, the four steps of the SPCC method (Fig. 4) consist, first, in analyzing the value created in each company involved in the SPCC (*Separate assessment*). Afterwards, on the basis of the information collected, the second step consists in searching for potential common opportunities of SPCC regarding one or many of the three value dimensions (method, object and/or nature of the value – *Co-creation analysis*). Thirdly, the method goes on in selecting through the list of opportunities, during an advisory board, which ones of the SPCC the company commits for (*Co-creation commitment*). Finally, step four consists in the deployment of the SPCC within each company’s respective information system (*Co-creation deployment*).

##### 1) Separated assessments

This first step aims to collect, assess and model the company’s assets that are impacted during the interaction between the partners as well as the activities they have in common and the value generated at each partner side by these activities. At this first step, the value considered is not restricted to a security or a privacy nature but may also concern value of other types like the quality or the usability.

During this step, interviews of the key persons from the companies are performed, existing enterprise models (e.g., architecture model, process model, etc.) are collected (independently of the language they are expressed in), and VC instances of the VC model are generated accordingly.

**Input:** The input to start the assessment of the companies’ contexts is simply “the willingness” to be engaged in the process and the commitment of the managers to support it.

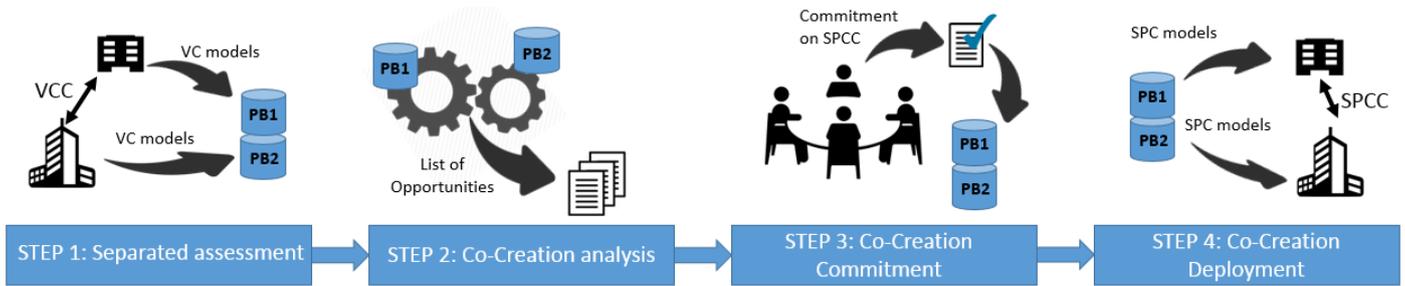


Figure 4: Security Co-Creation method

**Output:** Figures of the companies’ business, data and enterprise architecture models, business process, and instantiations of the VC model.

2) *Co-creation analysis*

During this step, all partners’ instantiated VC models are compared with the other partners’ instantiated VC models. This comparison allows mining the elements that all partner’s VC models have in common and where room exists for co-creation. According to the co-creation methods proposed in II.B.2, the analysis focuses on detecting model-based similarities between the types of nature of the value, the value co-creation methods, or the objects concerned by the value. This step may be supported by mining tools for the automatic model matching detection, as explained by Wille et al. [13].

**Input:** Instantiated VC models from each partners.

**Output:** Integrated opportunities of security and privacy co-creation models.

3) *Co-creation commitment*

This step is an important one in the process. It implies the decision-makers and managers to analyze which security and privacy co-creations they want to commit for based on the co-creation opportunities detected at step 2. To that end, the co-creation opportunity models are considered as relevant material given their capacity to clearly show impacts and benefits of the co-creations. After the decision being made, the integrated SPCC models are exploited to accordingly transform each respective VC models into a companies’ specific security and privacy creation (SPC) models. At this step, two types of company approach may exist: conservative or innovative. In the first case, the company wants to keep working with his ongoing solution but agrees to collaborate in order to support SPCC opportunities. In the second case, the

company is likely to accept changing its way of doing and even to adopt the other company’s approach.

**Input:** List of integrated SPCC opportunities.

**Output:** Security and privacy co-creation commitment from the decision makers and managers, and transformation of the VC models into companies’ specific SPC models.

4) *Co-creation deployment*

This step aims at deploying the co-creation activities in the companies’ running business. Therefore, each companies’ information systems are adapted following the SPC models defined at step 3. These modifications of the companies’ IS models are achieved manually or automatically depending of the models at stake and available tools.

**Input:** Companies’ specific SPC models.

**Output:** Companies’ information system adapted following the SPC models.

Table I provides a summary of the manipulations performed at the modeling level.

I. ILLUSTRATION

This section illustrates the deployment of the security co-creation method along the collaboration between a retail bank and a data center. Each step of the method is illustrated phase by phase.

A. *Step 1: Separate assessment*

This step aims at collecting the value co-creation activities from each company. Therefore, the VC activities (including the PIA and the SGAP) are analyzed and the VC model (Fig. 3) is instantiated accordingly.

TABLE I: MODELS’ CONTRIBUTIONS DURING METHOD STEPS

	Step 1: Separated assessments	Step 2: Co-creation analysis	Step 3: Co-creation commitment	Step 4: Co-creation deployment
Companies’ IS models	Companies’ IS models are used to instantiate VC model			Companies’ IS models are updated based on SPC instances
VC model and instances	VC instances of the VC model are created	VC instances are used for SPCC opportunities mining		
SPCC instances		SPCC opportunities are mined from each partner’s VC instances	SPCC opportunities are proposed and validated by the decision makers	
SPC instances			SPC instances are generated based on selected SPCC	SPC instances are used to update each partners’ IS.

Fig. 6 and 7 illustrate this instantiation respectively for the PIA process at the bank level and the SGAP at the data center level. At the bank, the nature of the value is the privacy of the bank customer’s financial assets. This privacy is generated by a privacy impact assessment method composed of the following elements: assessment of the leased line (that allows the data storage at the data center), assessment of the web portal, assessment of the value of the privacy, and analysis of the value/impact.

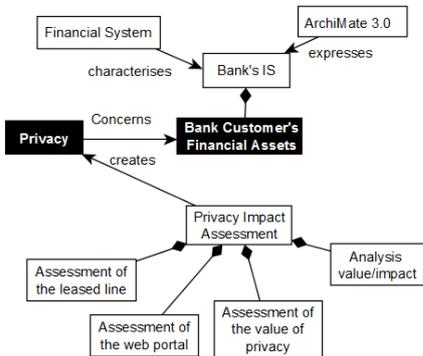


Figure 6: VC instance of the PIA at the bank

The same instantiation of the VC model is afterwards performed at the data center side. At that level, the nature of the value is the security of the data center backup and archiving operations. This security is obtained thanks to a security GAP analysis method which is built on the following four elements: assessment of the leased line, risk analysis, analysis of the cost of the controls, analysis of the business assets and assessment of the impact of a failure.

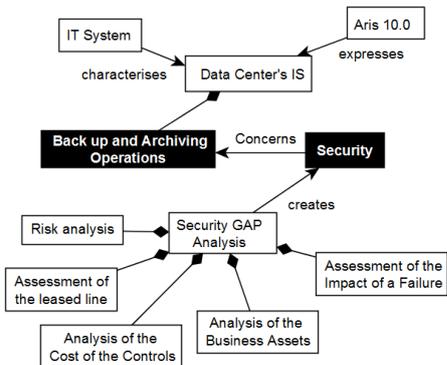


Figure 7: VC instance of the security GAP at the data center

**B. Step 2: Co-creation analysis**

This step aims to analyze and to detect security and privacy co-creation opportunities between companies. Therefore, the instances of the value creation models from both institutions, defined at step 1 (i.e., PIA and SGPA), are systematically compared with each other in order to identify similarities between concepts. As explained in II.B, the similarities may exist at the nature of the value level, at object of value level, or the value creation method level. Fig. 8 illustrates that both the PIA and the SGAP activities need to assess the leased line that allows the transfer of information from the bank to the data center. In that regard, a potential co-creation opportunity could be to assess it once and share the result amongst the partners.

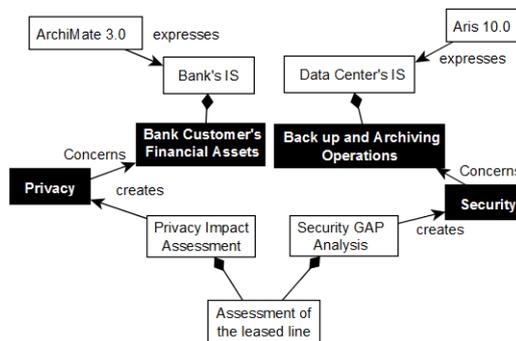


Figure 8: Security and privacy co-creation instance

**C. Step 3: Co-creation Commitment**

This step aims at taking the decision on which co-creation opportunity is relevant and justified for both companies and on adjusting the VC model of each company accordingly. For instance, after the commitment meeting, decision makers and managers of both companies agree to co-create security and privacy by optimizing the assessment of the leased line. More precisely, the assessment is performed by the data center agents and the results are sold to the bank at a good price. This decision needs to be reflected afterwards in the respective VC model. For instance, in Fig. 9, the element of the method “Assessment of the leased line” (Fig. 6) changed in “Receive subcontracted assessment of the leased line results”

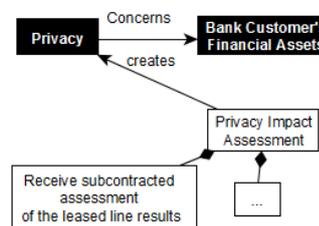


Figure 9: Security and privacy co-creation deployment

**D. Step 4: Co-creation Deployment**

This fourth step concerns the deployment of the co-created value in both ISs of the companies. Only the modeling dimension of this deployment is addressed. Concretely, after pledging commitment for a precise co-creation opportunity, both companies’ information systems needs to be updated accordingly. Therefore, the instantiated VC models (SPC models) are mapped with the respective companies’ ISs. The latter being expressed in their dedicated IS languages (respectively, ArchiMate 3.0 [22, 24] and Aris 10.0 [23]). Fig. 9 illustrates the new security and privacy value co-created after the achievement of each step of the method, respectively: *Leased line information* and *Money* from the bank to the data center, and *Assessment result* the other way round.

**II. DISCUSSION, CONCLUSION AND FUTURE WORKS**

Due to the ongoing developments of collaborative systems, companies are constantly willing to optimize the co-creation of value with their partners. Nowadays, this co-creation that was initially focused on the value of business assets tends to spread over others aspects such as the security and the privacy. This evolution in the creation of security and privacy features calls for new approaches to support companies in investigating

and deploying new security and privacy co-creation opportunities. Based on a value creation model and three co-creation schemas, we propose in this paper a four-steps innovative method for security and privacy co-creation that offers the advantage to be:

- simple to understand and deploy,
- adapted for three types of co-creation, to know: method-based, object-based or nature-based co-creation,
- sensitive to decision makers' and managers' commitment that is largely involved during the commitment step,
- independent of the companies information system architecture language (e.g., in the illustration, the bank used ArchiMate 3.0 and the data center used Aris 10.0).

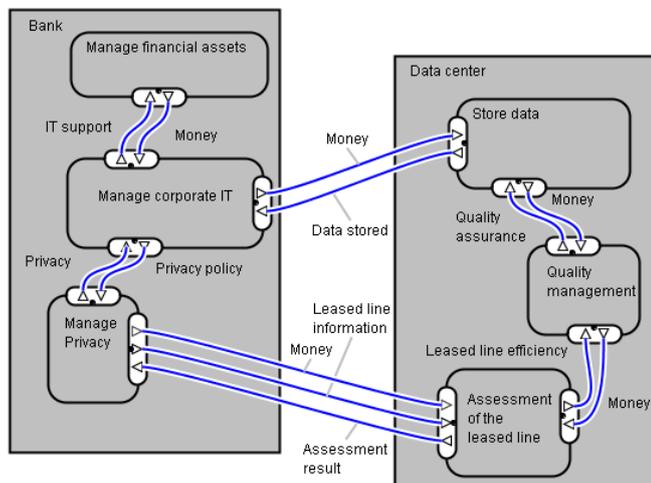


Figure 9: Security and privacy with co-creation - e3value model

As an improvement point, although the illustration happens in a real context setting, further complementary case studies and validations are required to confirm the efficiency of the method. These validations should allow analyzing to what extend the method may also be used to define and deploy security and privacy co-creation in the context of other co-creation schemas (e.g., object- and nature-based co-creation). In parallel, complementary instantiations of the latter should allow verifying to what extend (1) it is adapted and easy to apprehend by security professionals such as consulting companies or other service providers, and (2) it may support the definition of security and privacy co-creation between more than two partners, that is to say in networks of enterprises and business ecosystems.

Another element to be considered in future works consists in equipping the method with the appropriate tools, amongst which a dedicated model mining solution (e.g., [17]). The later would be especially relevant at the co-creation analysis step for detecting the co-creation opportunities by systematical mapping between each value creation instance.

REFERENCES

[1] A. Kharraz, W. Robertson, D. Balzarotti, L. Bilge and E. Kirde, "Cutting the gordian knot: A look under the hood of ransomware attacks," in DIMVA 2015, pp. 3-24, DOI: 10.1007/978-3-319-20550-2\_1.

[2] C. Feltus, E. Proper, "Conceptualization of an Abstract Language to Support Value Co-Creation", 12th Conference on Information Systems Management (ISM'17), IEEE.

[3] A. Smith, "The Wealth of Nations (1776)," *New York: The Modern Library*. 2000.

[4] V. A. Zeithaml, "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence," *The journal of marketing*, pp. 2-22, Jul. 1988. DOI:10.2307/1251446

[5] J. Gordijn, H. Akkermans, and H. Van Vliet, "Business modelling is not process modelling," In *Int. Conf. on Conceptual Modeling*, 2000, pp. 40-51. DOI:10.1007/3-540-45394-6\_5

[6] S. Vicini, F. Alberti, N. Notario, A. Crespo, J. R. T. Pastoriza and A. Sanna, "Co-creating Security-and-Privacy-by-Design Systems." In *ARES 2016*. DOI:10.1109/ARES.2016.74

[7] C. K. Prahalad and R. Venkat, "Co-creating unique value with customers," *Strategy & leadership*, vol. 32, no. 3, 2004, pp. 4-9.

[8] A. Bennaceur, T. T. Tun, A. K. Bandara, Y. Yu and B. Nuseibeh, "Feature-driven Mediator Synthesis: Supporting Collaborative Security in the Internet of Things", 2016, Doc. diss., The Open University.

[9] M. Hawley, P. Howard, R. Koelle and P. Saxton, "Collaborative security management: Developing ideas in security management for air traffic control," in *ARES 2013*, pp. 802-806, DOI: 10.1109/ARES.2013.107

[10] R. Garrido-Pelaz, L. González-Manzano and S. Pastrana, "Shall we collaborate?: A model to analyse the benefits of information sharing," in *2016 WISCS ACM*, pp. 15-24, 2016, DOI:10.1145/2994539.2994543

[11] S. L. Vargo and R. F. Lusch, "Service-dominant logic: continuing the evolution," *Journal of the Academy of marketing Science*, vol. 36, no. 1, pp. 1-10, Mar. 2008. DOI:10.1007/s11747-007-0069-6

[12] S. L. Vargo and R. F. Lusch, "Evolving to a new dominant logic for marketing," *Journal of marketing*, vol. 68, no. 1, pp. 1-17, Jan. 2004. DOI:10.1509/jmkg.68.1.1.24036

[13] D. Wille, S. Holthusen, S. Schulze and I. Schaefer, "Interface variability in family model mining," in *Proceedings of the 17th International Software Product Line Conference co-located workshops*, pp. 44-51, ACM, August 2013, DOI:10.1145/2499777.2500708

[14] C. Grönroos, "Service logic revisited: who creates value? And who co-creates?," *European business review*, vol. 20, no. 4, pp. 298-314, 2008.

[15] P. Frow, S. Nenonen, A. F. Payne, and K. Storbacka, "Managing Co-creation Design: A Strategic Approach to Innovation," *BJM*, vol. 26, no. 3, pp. 463-483, Jul. 2015. DOI: 10.1111/1467-8551.12087

[16] E. K. Chew, "iSIM: An integrated design method for commercializing service innovation," *Information Systems Frontiers*, vol. 18, no. 3, pp. 457-478, Jun. 2016. DOI: 10.1007/s10796-015-9605-y

[17] H. Weigand, "Value encounters—modeling and analyzing co-creation of value," in *Conf. on e-Business, e-Services and e-Society*, 2009, pp. 51-64. DOI:10.1007/978-3-642-04280-5\_5

[18] I. S. Razo-Zapata, E. K. Chew, and E. Proper, "Visual Modeling for Value (Co-) Creation," in *10th Int. Workshop VMBO 2016*.

[19] C. Feltus, M. Petit, and E. Dubois, "Strengthening employee's responsibility to enhance governance of IT: COBIT RACI chart case study," In *Procs of 1st ACM WISG '09*. ACM, New York, NY, USA, 23-32. DOI: http://dx.doi.org/10.1145/1655168.1655174

[20] C. Feltus, E. Grandry, T. Kupper, and J. N. Colin, "Model-Driven Approach for Privacy Management in Business Ecosystem," in *5th Int. Conf. on Model-Driven Eng. and Software Development*, 2017. DOI:10.5220/0006142203920400

[21] J. Ralyté, "Towards situational methods for information systems development: engineering reusable method chunks," in *Procs. of 13th Int. Conf. on Inf. Sys. Development. Advances in Theory, Practice and Education*. 2004.

[22] A. Josey, M. Lankhorst, I. Band, H. Jonkers, and D. Quartel, "An Introduction to the ArchiMate® 3.0 Specification," White Paper from The Open Group, Jun. 2016

[23] S. Ghatrei, "ARIS Enterprise Architecture's Usage Reviews," *Lecture Notes on Software Engineering*, 2015, vol. 3, no 1, p. 57. DOI: 10.7763/LNSE.2015.V3.166

[24] C. Feltus, E. Dubois, E. Proper, I. Band, M. Petit. 2012. Enhancing the ArchiMate® standard with a responsibility modeling language for access rights management. In *Proceedings of the Fifth International Conference on Security of Information and Networks (SIN '12)*. ACM, New York, NY, USA, 12-19.

# Secured Transactions Technique Based on Smart Contracts for Situational Awareness Tools

Roman Graf

Austrian Institute of Technology GmbH  
Vienna Austria  
roman.graf@ait.ac.at

Ross King

Austrian Institute of Technology GmbH  
Vienna Austria  
ross.king@ait.ac.at

## Abstract

*Modern critical infrastructures are increasingly targeted by highly sophisticated cyber attacks and are protected by increasingly complex tools. Cyber analysts face many challenges finding relevant information in large, complex data sets, and require novel distributed detection and reaction methodologies based on secured transaction techniques. These technologies should automatically analyse incident report and share analysis result in secure way between critical infrastructure stakeholders to achieve better situational awareness. Our goal is to provide solutions in real-time that could replace human input for cyber incident analysis tasks (Triage) to remove false positives and to eliminate irrelevant information. The effective and fast warning system should support cyber analyst to establish cyber situational awareness, and allow analysts to promptly respond in case of an attack. In this paper we evaluate the application of so-called “smart contracts” to an incident warning system and assess its accuracy and performance. We demonstrate how the presented techniques can be applied to support incident handling tasks performed by security operation centers. We show that a real-time “smart contracts” solution can replace human input for a large number of threat intelligence analysis tasks.*

## 1. Introduction

Cyber Situational Awareness (SA) [1] provides an overview of a security and threat situation as well as a current and future impact assessment. In recent years, researchers in SA field have created increasingly complex tools across many application domains to protect critical infrastructure (CI). CI comprises valuable assets (documents, software or hardware), which are essential for business or state authorities. Speed of events, data overload, and meaning underload [6] make real-time SA of cyber operations

very difficult to evaluate. Addressing data that are often vague and imprecise, analyst must rely on imperfect information to detect real attacks and to prevent an attack from happening. Cyber SA aggregates raw data at the lower level. While making a decision, a human cyber analyst faces challenges like finding relevant information in large, complex data sets. For humans to be effective in identifying and defeating future cyber-attacks, novel tools that can automatically make obvious or predefined decisions by means of smart contracts and fill the gap between cyber data and SA are highly desired. We hypothesize that the application of “smart contracts” based on existing blockchain technology (Ethereum [15]) can solve some problems associated with SA. The main purpose of designing smart contracts for SA is to enable rapid and trusted cyber incident warnings, without the need for a large centralized authority. We propose that smart contracts based on decentralised assets such as Bitcoin [2] can reduce effort for securing report transfer, manual analysis costs, and increase speed of severe information sharing.

The research presented evaluates a system based on blockchain and smart contract technology that will automatically warn a cyber analyst of high severity cyber incidents that could impact cyber SA reported and analysed by one of the trusted stakeholders. Within our prototype system, smart contracts trigger incident correlation analysis for large amounts of data by means of knowledge base employing one of the incident analysis tools. Smart contracts also compute incident severity using game theory score [8] and generate threat report if rules coded in related smart contract are met.

This paper is structured as follows: Section 2 gives an overview of related work and concepts. Section 3 explains the workflow for estimation of cyber incident priority level using smart contracts and also covers SA issues. Section 4 presents the experimental setup, applied methods and results. Section 5 concludes the paper.

## 2. Related Work

An overview of the concept of blockchain technology and its potential to facilitate smart contracts, automated banking ledgers and digital assets is provided in [11]. We suggest that the core technology of this approach can be reused in the cyber security domain by means of suitable smart contracts. A distributed peer-to-peer network based on blockchain technology where non-trusting members can interact with each other without a trusted intermediary, in a verifiable manner was examined in [3] for the Internet of Things sector. This mechanism should work also for the automation of multi-step processes for cyber incident analysis. The performance of Blockchain, which is a probabilistic proof-of-work (PoW) based consensus fabric [14] has become an important issue for the modern cryptocurrency platforms. PoW-based Blockchains can be replaced by BFT state machine replication, to improve Blockchain scalability limits. A Blockchain platform comparison [7] discusses five general-use Blockchain platforms and looks at how Blockchain technology can be used in applications outside of Bitcoin [10] to build custom applications on top of it. This comparison suggests that Ethereum is currently the most suitable platform, although the others all suffered in various ways from not yet having been developed as much as Ethereum. Therefore, for cyber incident analysis we employ well established Ethereum Blockchain in its Pyethereum fashion, which supports focused smart contracts testing environment without the need of mining.

A basis for smart contracts development in cyber security realm is a solid threat intelligence that is provided by a number of cyber incident analysis tools. The CAESAIR tool [13] introduces the concept of a cyber intelligence analysis system. CAESAIR provides analytical support for security experts carrying out cyber incident handling tasks on a national and international level, and facilitates the identification of implicit relations between available pieces of information. It provides powerful correlation capabilities, which support the tasks carried out by the analysts of a Security Operation Center during the incident handling process. CAESAIR<sup>1</sup> evaluates how the collected documents are connected to one another, and allows the analyst to select the most appropriate correlation method and to flexibly adjust relevance metrics.

In order to achieve the same objective, other approaches ([16], IntelMQ<sup>2</sup>, MISP<sup>3</sup>) have been proposed in recent years. These approaches aimed at parsing and correlating of cyber incidents, but we make use of CAESAIR because it supports various security information correlation techniques [13], differing in the way relevant information

<sup>1</sup><http://caesair.ait.ac.at>

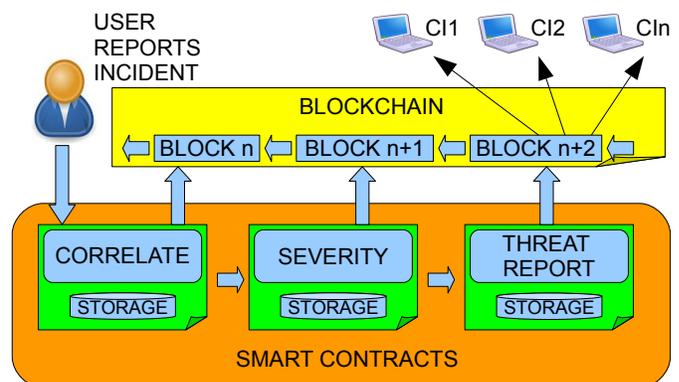
<sup>2</sup><https://github.com/certtools/intelmq>

<sup>3</sup><https://github.com/MISP/MISP>

is extracted from the imported document. Also, CAESAIR provides customizable import capabilities and information is acquired from a multitude of security-relevant sources. These sources include a custom repository, open source intelligence (OSINT) feeds, and IT-security bulletins, as well as a standardized vulnerability library (Common Vulnerabilities and Exposures - CVE). CVEs are especially important for smart contracts with regard to likelihood assessment based on game theory, which is the process of establishing the probability of an attack occurring ([5]) and, together with vulnerability scoring, might be seen as a way to implement risk scoring, which has been suggested to be more effective than vulnerability scoring alone [12].

## 3. Cyber Intelligence Analysis Using Smart Contracts

We evaluate the application of smart contracts to secure transactions (sending incident reports between CIs experts) in order to improve SA. Once a smart contract is triggered, the analysis result is automatically propagated among all participants through inherent blockchain mechanisms. One of the advantages of this approach is that smart contracts cannot be changed or compromised without being detected (through hashed transactions) and that the messages can be verified to originate from a trusted source (through public key encryption).



**Figure 1. The overview of designed system for establishing the Cyber Situational Awareness using Smart Contracts.**

### 3.1. System Design

To design threat intelligence analysis system, we describe incident handling procedure and instructions using smart contract programming language and upload this smart contract on a Blockchain. Smart contract is a source code

that comprises predefined instructions and rules. In this way we created “Correlate”, “Severity” and “Threat report” smart contracts (see Figure 1). The state of smart contracts is stored on the blockchain and is transparent and accessible to all registered community members. The smart contract code is executed in parallel by a network of miners under consensus regarding outcome of the execution. The execution of the smart contract results in an update of the contract’s state (BLOCKn+2) on the blockchain that is synchronized with every participating user (CI1-CIn) through standard peer-to-peer mechanisms. Incident report produced by one of the users (security expert protecting CIs) goes through the smart contracts and is handled automatically, according to the programmed instructions.

the report content. Input data, along with expert profile settings that are specific for an organisation, are passed to the first smart contract “correlation” in the second step. For the correlation computation we employ one of the threat intelligence tools or apply the cosine similarity algorithm [4], [17] for binary vectors that represent incident content. By means of threat intelligence tools we find similar related cyber incidents from the expert knowledge base. In the next step we merge the detected related incidents with institutional settings and using smart contract logic to automatically decide which priority level to assign a given incident. To calculate priority level we employ Equation 1. Priority level is dependent on different evaluation metrics, such as related words, related incidents, significant terms and vulnerability score. Each of these metrics may have its own threshold values and weights. These settings are naturally different by each organization, due to differences in employed software and hardware products, and organization specific business domain. Protected resources are different from organization to organization.

$$P = f(I_r, W_r, W_o, T_s, V_s). \quad (1)$$

Equation 1 shows the priority level  $P$  that returns value - either 0 that corresponds to “Low” or 1 standing for “High”. Priority level is a function of aggregated incident evaluation metrics, which depend on basis indicators, such as “number of related incidents”  $I_r$ , “number of related words”  $W_r$ , “number of original words”  $W_o$ , “detected significant terms”  $T_s$  and “vulnerability score”  $V_s$ .

### 3.3. Application Scenario

We assume that a cyber expert is responsible for a CI and detects suspicious behaviour in her system. She requires more information to select the correct mitigation strategy. In this case it is necessary to collect and analyse all of the available information related to ongoing and previous attacks for a particular use case, and transform it into intelligence. Security information, such as incident reports, vulnerability alerts, advisories, bulletins etc., usually comes in the form of semi-structured text documents. Acquiring cyber threat intelligence from such documents requires extracting the significant information they comprise, and identifying implicit correlations among them, in order to estimate their impact and outline possible mitigation strategies.

To avoid this manual effort, the CI expert can provide an incident report as an input to a contract and receive threat report back if it has sufficient severity. Incident analysis is performed by one of the threat intelligence tools with a solid knowledge base. Such a tool can quickly identify related threats and possible existing solutions by examining numerous Open Source INTelligence (OSINT) feeds.

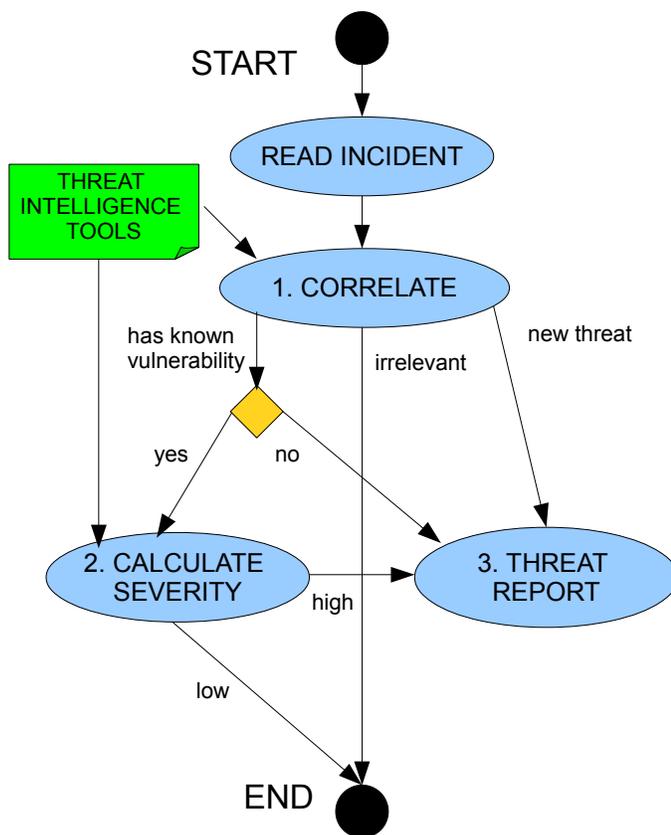


Figure 2. The workflow for estimation of cyber incident priority level using smart contracts.

### 3.2. Cyber Incident Processing

For cyber incident processing we employ three smart contracts as depicted in Figure 2. The workflow execution begins with the reading of an incident report and parsing of

The cyber incident analysis tool correlates the incident report with similar reported security incidents using e.g. the cosine similarity binary method and triggers another smart contract to calculate severity using for instance game theory score [9] in case of detected vulnerability. In the case that the severity is deemed “high,” the next smart contract is triggered and generates a new threat report, which can be automatically shared with all participants by means of storing it in smart contract memory on the blockchain. Other network peers from other CIs can securely query the smart contract data from blockchain by reading data stored in a smart contract. In smart contracts we make use of a state machine and specified rules governing report transfer. An automatic approach delivers a significant improvement in terms of personnel costs when compared to manual cyber incident handling and thanks to Blockchain technology makes all changes transparent to all participants in a trusted way. As a result all participants have access to the up-to-date SA status and we ensure secure information exchange and enrichment inside of trusted CI participant group.

We believe that this automatic smart-contracts-based approach would significantly facilitate analysis and could be used by analysts for the defence of critical infrastructure. The suggested method would make SA analysis less cost-intensive and would perform with higher throughput. However, as is typical in this area, a human-based approach performs with higher accuracy.

## 4. Experimental Evaluation

In this section we reported on measurements of the accuracy of the automated computations, how long it takes for each smart contract to be executed and validated, and how fast participants received SA update message starting from the issue time of original message. We carried out measurements for varying incident report sizes and different volumes of material in the knowledge base.

### 4.1. Evaluation Data Set

One responsibility of the cyber analyst is to prioritise a received cyber incident and to mitigate it or to carry out a selected cyber incident response. For this evaluation, we differentiated between high and low priority. High priority means that the incident has high severity and mitigation steps should be carried out. These types of incidents are tagged with a number 1. Low priority incidents are tagged with a number 0.

In our example scenario, we assume that our CI is a financial organisation that employed Microsoft Office products on Windows OS and used other software products such as Microsoft Internet Explorer, Firefox, Adobe, and so on. This knowledge is important for the correctness validation

of the developed smart contracts. We will check whether evaluation results are relevant for the organisation software landscape suggested in this example.

This evaluation took place on an Intel Core i7-3520M 2.66GHz computer using Python on Ubuntu OS. We evaluated 20 randomly selected cyber incident reports from the Canadian Cyber Incident Response Centre (CCIRC)<sup>4</sup>, the Carnegie Mellon University CERT blog<sup>5</sup> and also included some proprietary test incident reports. We then evaluated the calculation time and computation accuracy.

We assumed that employing of described smart contracts approach should correctly detect high priority cyber incidents among a very large number of incident reports and should not raise too many false positive results. We also expected that smart contracts written in Serpent<sup>6</sup> language, supported by Python 2.7 workflow and subsequent analysis will demonstrate both good performance and sufficient accuracy.

### 4.2. Experimental Results and Interpretation

The basis for the incident priority calculation was provided through incident correlation tool CEASAIR in which 20 incidents from open source intelligence feeds were aggregated, and through cyber experts evaluated configurable metrics. Table 1 represents the cyber incident profiles from the test dataset.

The computed priority level in the second column is compared with the expected priority level from ground truth in column three, which was estimated by a cyber expert.

The risk score in the fourth column demonstrated how many incident related points for different analysis metrics were aggregated in correlation smart contract logic.

The remaining columns are self explanatory.

The smart contract system started the Situational Awareness analysis with incident content retrieval (e.g. incident log report “AL17-006” in Table 1) and parsing. This input queried a threat intelligence tool via e.g. HTTP Rest request and received in response values of basis indicators mentioned in Formula 1. For given example “number of related incidents”  $I_r = 565$ , “number of related words”  $W_r = 42$ , “number of original words”  $W_o = 168$ , “detected significant terms”  $T_s = 5$  and calculated using game theory “vulnerability score”  $V_s = 0$ . High number of related incidents means that there are enough similar incidents in the past and looking on them could provide mitigation steps also for given incident. The number of related words, which is 42 demonstrated that current incident is quite good interconnected with detected related incidents and had many com-

<sup>4</sup><https://www.publicsafety.gc.ca/cnt/ntnl-scrnt/cbr-scrnt/ccirc-ccirc-en.aspx>

<sup>5</sup><http://insights.sei.cmu.edu/cert/atom.xml>

<sup>6</sup><https://github.com/ethereum/wiki/wiki/Serpent>

**Table 1. Priority level calculation for cyber incidents using smart contracts technology**

Incident ID	Priority Level	Priority Ground Truth	Risk Score	Related Incidents Number	Related Words Number	Original Words Number	Detected Significant Terms	Vulnerability Score	Time (sek.)
AL17-006	1	1	7	565	42	168	5	0	1.34
AV17-035	1	1	5	556	19	39	2	1	0.60
AV17-071	1	0	4	556	19	32	1	1	0.83
AL17-005	1	1	7	547	23	49	4	0	0.70
AL17-004	1	1	4	561	71	308	2	0	1.51
AL17-002	None	0	None	554	42	98	1	1	1.51
AL17-003	1	1	7	515	55	167	5	0	1.31
AL17-001	1	0	4	532	31	92	1	0	0.94
AV17-070	0	0	3	529	15	25	0	0	0.65
AV17-069	0	0	2	499	18	108	0	1	0.67
AV17-068	1	1	6	503	20	49	3	0	1.01
AV17-066	1	0	4	563	32	65	1	1	1.21
AV17-065	1	1	4	499	12	26	1	1	0.40
AV17-064	1	1	5	561	22	41	2	1	0.97
AV17-063	1	1	5	519	24	96	3	1	0.79
AV17-062	1	1	5	552	20	49	2	1	0.66
CMU-CERT	1	1	5	531	22	94	3	0	0.82
CMU-CERT2	1	1	5	524	28	110	3	0	0.83
AIT-1	1	1	4	492	37	76	1	0	0.68
AIT-2	0	0	3	478	28	68	0	0	0.51

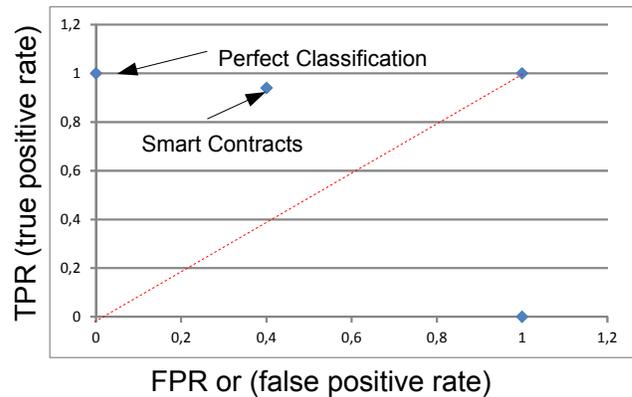
mon words, whereas the total number of words in original document was 168. Therefore, about a quarter of all original words were found in related incidents. And five significant terms were found in incident content. These terms are specific for particular organisation and demonstrate relation to the organisational software, hardware and network. E.g. organisation that make use of Windows OS is not interested in incident reports related to Android OS. Most true positive reports correctly contained significant terms, such as “Microsoft”, “Mozilla”, “CVE”, “Adobe”, “ransomware” etc. These results are good explainable and can be regarded as reliable.

The performance of computation is quite good and ranged between 0,4 to 1,51 seconds depending on incident report size and number of detected related incidents. This performance of the evaluation was achieved using the Pyethereum test environment and is significantly faster than it would be in a real ethereum network. The reason for that is validation procedure where the transaction passes different devices around the world and is waiting for an Ethereum miner to be inserted into a block of transaction on the Blockchain.

### 4.3. Evaluation Effectiveness

We concluded that the examined approach presented reliable results for cyber incident analysis and can be applied as a supporting method for improvement of situational awareness. This method helps to automatically discover incidents with high priority in very large data set. Following this, manual analysis of high priority cyber incident candidates separated real important threats from structural similar reports. The presented method saved time and therefore costs associated with human expert involvement in situational awareness establishment process. Only few false positive detections were made despite the low homogeneity of the experimental collection. Therefore our initial hypothesis is confirmed. But additional research is required to improve

smart contract decision and accuracy metrics of mentioned method, as well as to measure response times under real blockchain network conditions.



**Figure 3. ROC space plot.**

The search effectiveness for high priority incidents could be determined in terms of a Relative Operating Characteristic (ROC). Situational awareness analysis divided the provided incident reports in two groups “high priority and “low priority by associated expert parameters and thresholds. The provided algorithm detected 16 true positive *TP* incidents, three true negative *TN* reports, two false positive *FP* incidents and one false negative *FN* document. The main statistical performance metrics for ROC evaluation are sensitivity or true positive rate *TPR* and false positive rate *FPR* (see Equation 1).

$$TPR = \frac{TP}{(TP + FN)}, FPR = \frac{FP}{(FP + TN)}. \quad (2)$$

Therefore the sensitivity *TPR* of the presented approach for cyber incident prioritisation is 0.94, the *FPR* is 0.4. The associated ROC value for is represented by point (0.4,

0.94). The ROC space (see Figure 3) demonstrates that the calculated  $FPR$  and  $TPR$  value for this point were located very close to the so called perfect classification point (0, 1). These results demonstrated that an automatic approach for cyber incident prioritisation of mentioned method was very effective and it was a significant improvement compared to manual analysis. The distribution of collection points above the red diagonal demonstrated quite good classification results that could be improved by refining of expert settings. Therefore, analysis methods based on smart contracts and blockchain techniques can be suggested as an effective method for high priority incident detection and as a supporting method for establishment of cyber situational awareness. The results of the analysis confirmed our hypothesis that an automated approach is able to detect high priority incidents with reliable quality, thus making analysis of large number of cyber incidents a feasible and affordable process.

## 5. Conclusions

In this work we have presented an automated approach to secure transactions for establishing cyber situational awareness using smart contracts. We have combined expertise gathered during the development of a cyber intelligence tool with the power of the smart contracts approach. The main contribution of this work is a real-time solution that could replace human input for a large number of cyber incident analysis tasks in order to remove false positives and to eliminate irrelevant information. Another contribution is the employment of smart contract techniques to provide a trusted early warning system about severe cyber threats. The presented method employs a domain expert knowledge base collected through a cyber intelligence tools to detect Situational Awareness risks. An additional advantage of this approach is a reduction of human analysis costs. Ultimately, our research will lead to the creation of automated security assessment tools with more precise and more accurate predictions of security attacks.

## References

- [1] P. Barford, , and et al. Cyber sa: Situational awareness for cyber defense. In S. Jajodia, P. Liu, V. Swarup, and C. Wang, editors, *Cyber Situational Awareness*, volume 46 of *Advances in Information Security*, pages 3–13. Springer US, 2010.
- [2] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten. Sok: Research perspectives and challenges for bitcoin and cryptocurrencies. In *2015 IEEE Symposium on Security and Privacy*, pages 104–121, May 2015.
- [3] K. Christidis and M. Devetsikiotis. Blockchains and smart contracts for the internet of things. *IEEE Access*, 4:2292–2303, 2016.
- [4] N. Dehak, R. Dehak, J. Glass, D. Reynolds, and P. Kenny. Cosine similarity scoring without score normalization techniques. in *Proceedings of Odyssey 2010 - The Speaker and Language Recognition Workshop (Odyssey 2010)*, pages 71–75, 2010.
- [5] W. Kanoun and et al. Success likelihood of ongoing attacks for intrusion detection and response systems. volume 3, pages 83–91. IEEE, 2009.
- [6] A. Kott and C. Wang. *Cyber Defense and Situational Awareness*. Springer International Publishing Switzerland, 2014.
- [7] M. Macdonald, L. Liu-Thorrold, and R. Julien. *The blockchain: A comparison of platforms and their uses beyond bitcoin*. The University of Queensland, 2017.
- [8] L. Maghrabi, E. Pfluegel, L. Al-Fagih, R. Graf, G. Settanni, and F. Skopik. Improved software vulnerability patching techniques using cvss and game theory. In *2017 International Conference on Cyber Security And Protection Of Digital Services (Cyber Security)*, pages 1–6, June 2017.
- [9] L. Maghrabi, E. Pfluegel, and S. F. Noorji. Designing utility functions for game-theoretic cloud security assessment: a case for using the common vulnerability scoring system. In *2016 International Conference On Cyber Security And Protection Of Digital Services (Cyber Security)*, pages 1–6, June 2016.
- [10] S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system. [Online] Available: <https://bitcoin.org/bitcoin.pdf>, 2009.
- [11] G. W. Peters and E. Panayi. *Understanding Modern Banking Ledgers Through Blockchain Technologies: Future of Transaction Processing and Smart Contracts on the Internet of Money*, pages 239–278. Springer International Publishing, Cham, 2016.
- [12] T. Reguly. Does anybody really care about vulnerability scoring? [online], Available: <https://www.tripwire.com/stateof-security/risk-based-security-for-executives/risk-management/doesanybody-really-care-about-vulnerability-scoring/>, 2013.
- [13] G. Settanni, Y. Shovgenya, F. Skopik, R. Graf, M. Wurzenberger, and R. Fiedler. Correlating cyber incident information to establish situational awareness in critical infrastructures. In *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, pages 78–81, Auchland, New Zealand, Dec 2016. IEEE.
- [14] M. Vukolić. *The Quest for Scalable Blockchain Fabric: Proof-of-Work vs. BFT Replication*, pages 112–125. Springer International Publishing, Cham, 2016.
- [15] G. Wood. Ethereum: A secure decentralised generalised transaction ledger. <http://gavwood.com/paper.pdf>, 2014.
- [16] S. J. Yang and et al. High level information fusion for tracking and projection of multistage cyber attacks. *Information Fusion*, 10:107–121, 2009.
- [17] J. Ye. Cosine similarity measures for intuitionistic fuzzy sets and their applications. *Mathematical and Computer Modelling*, 53(1-2):91–97, 2011.

# Speech Encryption Based on Hybrid Chaotic Key Generator for AMR-WB G.722.2 Codec

Messaouda Boumaraf

LISIC Laboratory, Telecommunications Department  
USTHB University, Algiers, Algeria  
boumaraf.messa@gmail.com

Fatiha Merazka

LISIC Laboratory, Telecommunications Department  
USTHB University, Algiers, Algeria  
fmerazka@usthb.dz

**Abstract**—In recent years, chaos-based cryptosystems have drawn more and more attention in many disciplines in particular speech encryption. In this paper, we propose speech encryption scheme based on chaos maps for AMR-WB G.722.2 Codec. To increase the security level, we combine chaos maps logistic and Hénon for shuffling and scrambling speech in order to have a hybrid chaotic key generator. The proposed algorithm evaluated with both Perceptual Evaluation of Speech Quality (PESQ) and Enhanced Modified Bark Spectral Distortion (EMBSD) measure confirm the efficiency of our proposed cryptosystem scheme.

**Keywords**—Encryption, Logistic Map, Hénon Map, PESQ, EMBSD, Speech Scrambling, Chaos Cryptosystem.

## I. INTRODUCTION

The emergence of the use of the internet became very apparent, the security of digital information such as voice data is now an important issue for all internet users. Because of this, the data between the legitimate users need to be protected before transmission by using encryption methods. With the significant computers cryptography developments, numerous studies are involved to secure communications [1, 2].

Recently, the amount of research on chaotic cryptography increased more and more in order to improve chaos-based cryptosystems. Chaos-based encryption algorithms are based on different types of chaotic maps whether discrete or continuous maps. Most of them are a combination of two or more chaotic maps to achieve a better security, expanded key space and low complexity [3-5].

A chaotic system is a non-linear, deterministic and presents good properties such as pseudo-randomness, sensitivity to changes in initial conditions, system parameters and aperiodicity which makes it unpredictable. Because of its characteristics, the chaos was used in the encryption system and not allow an adversary to find the outputs without any knowledge of the initial values [6-10].

The introduction of the concepts of cryptosystems is followed by the mapping of the two theories: the chaos and the cryptography. We will highlight the importance of using dynamic chaotic systems in cryptography.

Many chaos-based encryption methods have been introduced during the last decade. The most cited and

important chaos-based structure was presented by Fridrich in [6,7] using substitution and permutation. In [8] Wen has reviewed the dynamic properties of the Hénon map including its fixed points, stability, periodic orbits, and so on, and a physical interpretation of it is discussed. In [9] authors proposed a new image encryption algorithm based on parameter-varied logistic chaotic map and dynamical algorithm. They used the parameter-varied logistic map to shuffle the plain image, and then used a dynamical algorithm to encrypt the image.

In order to use chaos theory effectiveness in cryptography, the chaotic maps should be implemented such that the entropy generated by the map can produce required confusion and diffusion architecture proposed by Shannon [2].

Today, based on some important properties of chaos, such as the unpredictable behavior which can be used in the generation of random numbers.

The chaos-based cryptosystems provide several advantages, such as: very high security level, high speed especially in stream ciphers, computational power, which are easier to be implemented. These features make them more suitable for large scale-data encryption, such as voice and image. However, these characteristics can be refined to simulate the characteristics of a white noise or other "random" signal, which makes chaos a very interesting phenomenon for hiding information signals in order to transmit them in a "secure" manner [11-13]. In other words, the encryption of a data by chaos is done by superposing the initial information at chaotic signal. After, we send the data drowned in chaos at a Receiver which knows the characteristics of the chaos' generator. It then remains for the recipient only to subtract the chaos of his data in order to retrieve the information [14].

The chaos streams are generated by using various chaotic maps.

In this paper, two chaos-based cryptosystems of transmitted speech security are presented and the obtained results are discussed.

The remainder sections of this paper are organized as follows. In section 2, an overview of the AMR-WB G.722.2 is introduced. Section 3, our proposed cryptosystem is

presented. Simulations and interpretation of obtained results are discussed in section 4. Finally, the conclusion is provided in section 5.

II.OVERVIEW OF THE AMR-WB G.722.2

Adaptive Multi-Rate Wideband (AMR-WB) is an apparent wideband speech audio coding standard enhanced and based on Adaptive Multi-Rate encoding, using similar methodology as algebraic code excited linear prediction (ACELP). AMR-WB gives improved speech quality due to a larger speech bandwidth of 50–7000 Hz compared to narrowband speech coders which optimized for POTS (Plain Old Telephone Service) wire line quality of 300–3400 Hz. AMR-WB was upgraded by Nokia and Voice Age and it was first defined by 3GPP [15].

AMR-WB is codified as G.722.2, an ITU-T standard speech codec, formally known as Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB). G.722.2 AMR-WB is the same codec as the 3GPP AMR-WB.

The AMR-WB speech codec contains nine bit rates of 23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85 and 6.6 kbps, these ones are presented by modes 8, 7, 6, 5, 4, 3, 2, 1 and 0 respectively. The bit rate can be changed at any frame boundary of 20 ms. the codec includes Voice Activity Detection(VAD), Discontinuous Transmission (DTX) and Comfort Noise Generation (CNG) features for increased efficiency [15].

The AMR-WB G722.2 uses six parameters to represent the speech and these are shown in Table I for bit rate 8.85kbit/s [15].

TABLE I. G.722.2 – BIT ALLOCATION OF THE AMR-WB CODING ALGORITHM FOR 20-MS FRAME.

Mode 1 (8.85kbit/s)	VAD-flag				1	
	ISP				46	
	Pitch delay	8	5	8	5	26
	Algebraic code	20	20	20	20	256
	Gain	6	6	6	6	24
	<b>Total</b>					<b>177</b>

III.PROPOSED CRYPTOSYSTEM BASED CHAOS

In this work, two chaotic maps are used-2D Hénon map and 1D logistic map, each one has its own property or characteristic and has its own effect on improving the performance of evolutionary algorithm. In General, the information in secure-communication is transmitted through the channel after source encoding, encryption and channel encoding and modulation, then it will be received by reversing these steps, as shown in Fig. 1. In the following we will describe Logistic map, Hénon map and then our proposed algorithm.

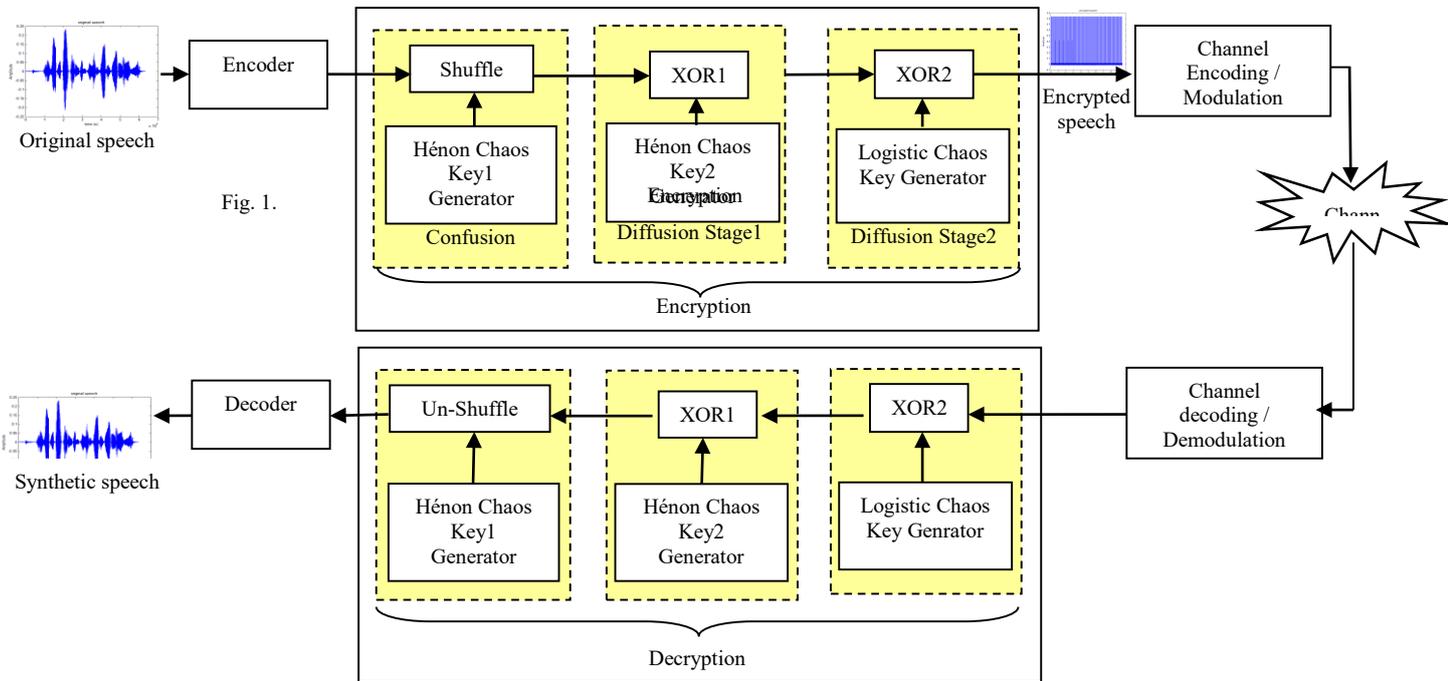


Figure 1. Structure Secure communication based on our proposed scheme

A. Logistic Map

The Logistic map is a prototypical one-dimensional invertible iterated map represented by the state equations with a chaotic attractor that exhibits complicated behavior. This map is introduced by Verhulst [16], it is constituted of a discrete-time dynamical system and its recurrence equation is given by [17]:

$$x_{k+1} = \mu x_k (1 - x_k) \quad x_k \in (0,1) \quad (1)$$

where,  $\mu$  belongs to the interval  $[0,4]$  and this parameter determines the map behavior. When parameter  $\mu$  has the following range ( $3.57 < \mu \leq 4$ ) it becomes a chaotic map, and  $x_k$  belongs to the interval  $[0,1]$ , knowing that any change in initial value or parameter  $\mu$  will give various sequences of random and irregular numbers. In our case, we set the value of the control parameter to the value corresponding to  $\mu = 4$  and  $x_0 = 0.28$ .

Remember that the Chaos can be generated by any non-linear dynamic system. Indeed, simple recurrence equations are capable of creating rich chaotic dynamics, if the parameters are well situated. In many recurrence simple equations, the right choice of these parameters is made by means of the bifurcation diagram and the exponent of Lyapunov. We present the known curves of Bifurcation diagram and the Lyapunov exponent of the map in Figs. 2 and 3 respectively.

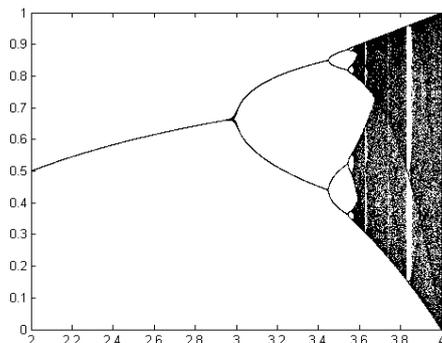


Figure 2. Bifurcation diagram of the logistic map

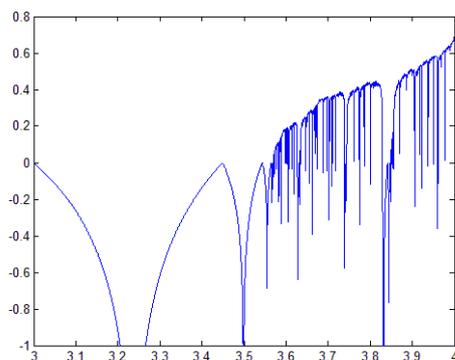


Figure 2. Structure of Lyapunov exponent of the logistic map

B. Hénon Map

The Hénon map is a prototypical two-dimensional invertible iterated map [18]. It has a chaotic behavior, and can be expressed as a recurrence of two chaos signals given by [19]:

$$\begin{cases} x_{n+1} = 1 + y_n - ax_n^2 & (a) \\ y_{n+1} = bx_n & (b) \end{cases} \quad (2)$$

To get a random sequence, we used the for the parameters 'a' and 'b' the following values.  $a = 1.2$ ,  $b = 0.1$ . with these values, the initial point is  $(x_0, y_0) = (0.1, 0.1)$ , the sequence of points is obtained by the mapping's iteration and it tends to a strange attractor. The chaotic Hénon map is shown in Fig. 4.

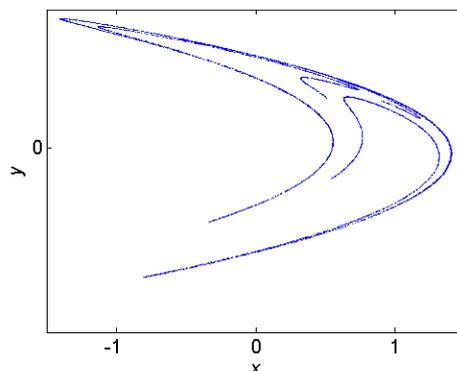


Figure 4. Chaotic behavior of Hénon attractor ( $x_0 = 0.1, y_0 = 0.1, a = 1.2, b = 0.4$ , iteration = 6000)

C. Proposed Cryptosystem speech Encryption's Algorithm

In this subsection, we present the steps of our proposed speech encryption algorithm based on confusion employing Hénon map and diffusion employing Hénon and logistic map. It is divided in three major steps:

**Step1:** In the confusion stage, the parameters of frames are shuffled and permuted by using formula (2-a) of Eq. (2). Because, the key  $x_n$  takes values from interval  $]-1.5, 1.5[$ , we arrange them in an increasing or decreasing order. After sorting, we save the position or the index of each key values then, change the position of speech data according to indexes' keys [18].

**Step2:** In the diffusion stage 1, the permuted parameters of frames are substituted employing XOR operation using formula (2-b) of Hénon Eq. (2). The keys  $y_n$  takes a value from the interval  $]-1, 1. [$ , so we calculate:

$$\text{floor}(\text{abs}(y_n) * 106 \text{ mod } 65536) \quad (3)$$

The diffusion is performed using XOR operation. **Step3:** In the diffusion stage 2, we generate keys by using logistic map. We obtain a series of numbers  $k_1, k_2, k_3, \dots, k_n$  in the range  $[0,1]$ . Where  $n$  is the number of words in the speech to be encrypted. Then, we follow these steps:

- We choose the length of the key with 16 bits because our AMR-WB G.722.2 words are 16 bits. Then, we multiply each number by 65536. Let

$$\text{key}_i = \text{int}(k_i * 65536 + 0.5) \quad (4)$$

- Finally, words of speech data are modified by employing XOR operation.

IV. SIMULATION AND RESULTS

In this section, we present the simulation setup followed by the obtained results. Several experiments are carried out to test the encryption efficiency of the presented wideband speech cryptosystem. The quality of both the encrypted and reconstructed signals is assessed for the standard AMR-WB G.722.2.

The speech file extracted from TIMIT database [20] and sampled at 16 kHz was encoded using AMR-WB G.722.2 CS-ACELP. The resulting bit streams were encrypted using Hénon and Logistic Maps schemes. In experiments, signal inspection in both the time and frequency domains is done to evaluate the changes between the original speech and reconstructed speech and encrypted speech using confusion given in Figs. 5, 6 and 7 respectively. For diffusion and both confusion and diffusion are given in Figs. 8 and 9 respectively.

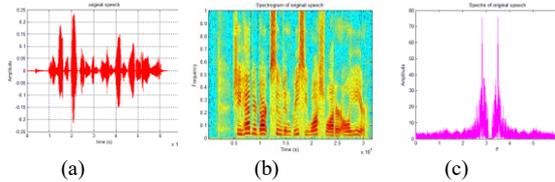


Figure 5. (a) Original speech, (b) its Spectrogram (c) its Spectrum

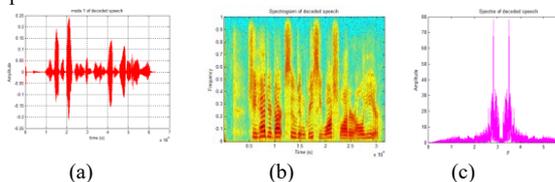


Figure 6. (a) Decoded speech using mode 1(b) its Spectrogram (c) its Spectrum

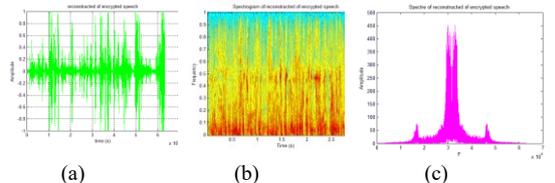


Figure 7. (a) Speech encryption using confusion (Hénon map) (b) its Spectrogram (c) its Spectrum

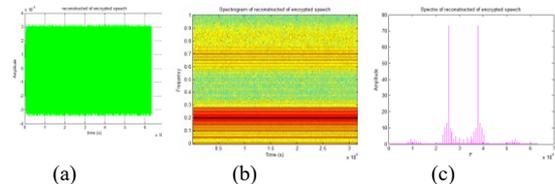


Figure 8. (a) Speech encryption using diffusion (Hénon & Logistic Maps) (b) its Spectrogram (c) its Spectrum

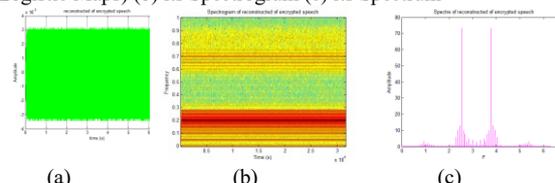


Figure 9. (a) Speech encryption using confusion & diffusion (Hénon & Logistic Maps) (b) its Spectrogram (c) its Spectrum

We can see from Figs. 8-a, and 9-a that the encrypted speech signals by using only diffusion or using confusion and diffusion, are comparable to a white noise signal which indicates that no significant residual intelligibility can be useful for eavesdroppers at the communication channel. Note

that the reconstructed speech signals using the right keys are identical to the original.

The Enhanced Modified Bark Spectral Distortion (EMBSD) [21] is an objective measurement tool, used to evaluate the efficiency of encryption schemes. The obtained results from tests with EMBSD are given in Fig. 10. It is recalled that the EMBSD gives a value of 0 for two identical speech files, and a greater value as the distortion increases. So, we can see that the best values are given for the original speech coder and the worst are given for encryption using both confusion and diffusion i.e. hybrid chaotic generator and for using diffusion only.

We have also evaluated the performance of our cryptosystem using the Perceptual Evaluation of the Speech Quality (PESQ) [22], results are given in Fig. 11 for original speech and encrypted speech using confusion only, encrypted speech using diffusion only and our proposed cryptosystem. It is recalled that the PESQ gives a value of 4.5 for two identical speech files and a less value as the distortion increases. We can see again that the best PESQ is given by the original speech coder since it is not encrypted and the worst PESQ values are obtained for encryption using both confusion and diffusion i.e. hybrid chaotic generator and for using diffusion only.

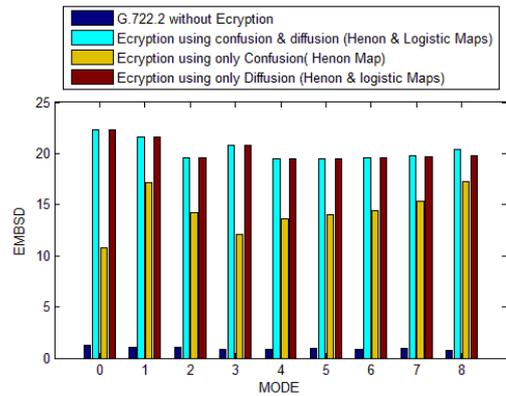


Figure 10. Results with EMBSD evaluation.

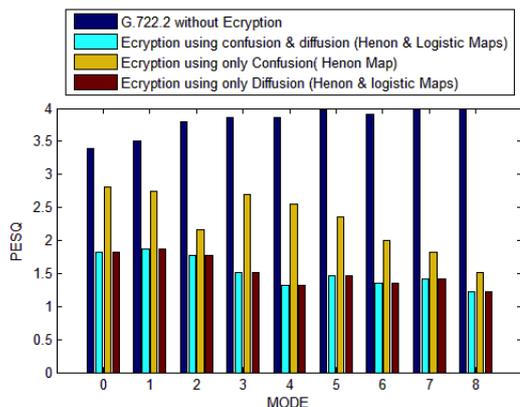


Figure 11. Results with PESQ evaluation

So both metrics confirm the efficiency of our cryptosystem based on hybrid chaotic maps and that confusion does not affect encryption using diffusion with Hénon and logistic maps since results are

comparable for encryption with diffusion or with confusion and diffusion.

Tests have been carried out to compute the time required to execute our proposed cryptosystem scheme. To estimate this time, we repeat the execution of each of the 1000000 times (iterations) and then we divide the duration obtained over 1000000. Results for encryption and decryption using our proposed cryptosystem are depicted in Fig. 12. Results show that our cryptosystem is very fast and does not exceed hundreds of nanoseconds.

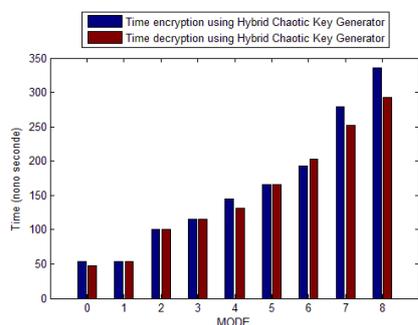


Figure 12. Runtime Per Nano Seconds

## V. CONCLUSION

In this paper, a hybrid chaotic key generator was proposed. It was implemented through two simple chaotic maps to encrypt and decrypt speech for AMR-WB G.722.2 Codec based on the concept on shuffle the position of all parameters of frame in order to produce permutations followed by using two substitutions in order to produce both confusion and diffusion as Shannon's principle requires. The experimental results and analysis show that our cryptosystem provides an efficient and sensitive key space. Additionally, it is effective, fast in terms of execution time and provides a better security compared to using confusion only.

## REFERENCES

- [1] Wei-Der Chang, "Digital secure communication via chaotic systems," *Digital Signal Processing*, 19(4):693–699, 2009.
- [2] Zhang Z and Cao T, "A chaos-based image encryption scheme with confusion- diffusion architecture", *Communications in Computer and Information Science*. 2011;152(1):258–263.
- [3] Alvarez G and Li S, "Some basic cryptographic requirements for chaos-based cryptosystems," *Int. J. Bifurcation Chaos* 16, 2129 (2006).
- [4] Zhu Z, Zhang W, Wong K and Yu H, "A chaos-based symmetric image encryption scheme using a bit-level permutation", *Information Sciences*. 2011;181(6):1171–1186.
- [5] Wong K, Kwok BS and Law W, "A fast image encryption scheme based on chaotic standard map", *Physics Letters, Section A: General, Atomic and Solid State Physics*. 2008;372(15):2645–2652.
- [6] Jiri Fridrich, "Image encryption based on chaotic maps", In *Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation.*, 1997 IEEE International Conference on, volume 2, pages 1105–1110. IEEE, 1997.
- [7] Jiri Fridrich, "Symmetric ciphers based on two-dimensional chaotic maps", *International Journal of Bifurcation and Chaos*, 8(06):1259–1284, 1998.
- [8] Haoran Wen, "A review of the Hénon map and its physical interpretations," *advance online publication*: 30 November 2014.
- [9] Lingfeng Liu, and Suoxia Miao, "A new image encryption algorithm based on logistic chaotic map with varying parameter", *SpringerPlus* (2016) 5:289.

- [10] Sandhya Rani M.H.1 and K.L. Sudha, " Design and Implementation of Image Encryption Algorithm Using Chaos", *International Journal of Advanced Computer Research* (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-4 Number-2 Issue-15 June-2014
- [11] A Akhshani, A Akhavan, S-C Lim, and Z Hassan, "An image encryption scheme based on quantum logistic map," *Communications in Nonlinear Science and Numerical Simulation*, 17(12):4653–4661, 2012.
- [12] Fatiha Merazka, " Wideband Speech Encryption based Arnold Cat Map for AMR-WB G.722.2 codec," *ICISP 2014*: 658-664.
- [13] Alireza Jolfaei and Abdolrasoul Mirghadri, "An image encryption approach using chaos and stream cipher," *Journal of Theoretical and Applied Information Technology*.
- [14] Radu Eugen BORIGA, Ana Cristina DĂSCĂLESCU, and Adrian Viorel DIACONU, "A New Fast Image Encryption Scheme Based on 2D Chaotic Maps," *IAENG International Journal of Computer Science*, 30 November 2014
- [15] "Wideband coding of speech at around 16 kbps using Adaptive Multi-Rate Wideband (AMR-WB)," *ITU-T Standard G.722.2*, 2003.
- [16] Didier Gonze, "The logistic equation," September 30, 2015.
- [17] Kamal Jadidy Aval, Morteza Sabery Kamarposhty and Masumeh Damrudi, "A Simple Method for Image Encryption Using Chaotic Logistic Map," *Journal of Computer Science & Computational Mathematics*, Volume 3, Issue 3, September 2013
- [18] Sattar B. Sadkhan and Hussein Ali, "A proposed Speech Scrambling based on hybrid chaotic key generators," *Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA) – IRAQ* (9-10) May 2016
- [19] Manjunath Prasad and K.L.Sudha, "Chaos image encryption using pixel shuffling with henon map," *Elixir Elec. Engg.* 38 (2011) 4492-4495
- [20] NIST, Timit Speech Corpus, NIST 1990.
- [21] W. Yang, "Enhanced Modified Bark Spectral Distortion (EMBSD): An Objective Speech Quality Measurement Based on Audible Distortion and Cognition Model," Ph.D Dissertation, Temple University, USA, May 1999
- [22] ITU-T Draft Rec P.862, "Perceptual evaluation of speech quality (PESQ), an objective method of end-to-end speech quality assessment of narrowband telephone networks and speech codecs", May 2000.

# A Blind Watermarking Technique based on DCT Psychovisual Threshold for A Robust Copyright Protection

Ferda Ernawan

Faculty of Computer Systems & Software Engineering  
Universiti Malaysia Pahang  
Lebuhraya Tun Razak 26300 Gambang, Kuantan Pahang  
Darul Makmur, Malaysia  
e-mail: ferda@ump.edu.my

Muhammad Nomani Kabir, Zuriani Mustaffa

Faculty of Computer Systems & Software Engineering  
Universiti Malaysia Pahang  
Lebuhraya Tun Razak 26300 Gambang, Kuantan Pahang  
Darul Makmur, Malaysia  
e-mail: nomani@ump.edu.my, zuriani@ump.edu.my

*Abstract*— This work presents a blind watermarking technique based on a psychovisual threshold for a robust copyright protection. Psychovisual threshold can provide a trade-off between imperceptibility and robustness of the watermark. This paper proposes a new watermarking scheme where watermarks are embedded on some selected coefficients of DCT frequencies considering psychovisual threshold. The embedding regions are determined using a modified entropy to indicate less distortion areas. Furthermore, the watermark image is scrambled before embedding to provide additional security. The proposed scheme is tested under different types of signal-processing and geometric attack. The experimental results show that our scheme produces higher imperceptibility and robustness of the watermark than the existing schemes.

*Keywords*-image watermarking; modified entropy; embedding scheme; extraction scheme; psychovisual threshold;

## I. INTRODUCTION

With the rapid growth of multimedia technology, illegal copying, transmission and distribution of digital data, digital copyright become important issues. This issue encourages for developing a solution to overcome this issue. Digital watermarking can be considered as an alternative solution to protect the digital content [1]. Watermark is embedded into the host media (like images, videos, audio, text, etc.) to protect the copyright or ownership of the digital media. Many research works have been conducted in image watermarking techniques. Nowadays, watermarking techniques based on hybrid frequency-based and matrix decomposition have been proposed [2]-[4]. In [5][6], the proposed watermarking scheme adds watermarks into the frequency coefficients with scaling factor. This technique is able to produce sufficient imperceptibility and robustness of the recovered watermark against several

attacks, while the imperceptibility of the watermarked image needs to be improved.

In 2011, Lai scheme [3] presented a novel watermarking method based on human visual system with singular value decomposition, a binary watermark image is embedded into the host image by modifying  $U$  orthogonal matrix. Watermarks were inserted by modifying  $U_{3,1}$  and  $U_{4,1}$  coefficients from the hybrid DCT-SVD. This scheme reportedly provides better performance in terms of robustness under adding noise, histogram equalization, scaling and Gaussian low-pass filter. While the watermark embedding capacity is still insufficient, the robustness of watermarked image needs to be enhanced.

Roy et al. [7] proposed embedding multiple watermarks in the middle band coefficients by a zig-zag order. This method can produce better robustness under noise addition, JPEG compression, and sharpening. This scheme performs sequent embedding process from left-corner to the right, from top to bottom. The diversity of embedded block regions needs to be improved, such that the watermark image is difficult to be recovered. Roy et al. [7] scheme modified 22 coefficients in the middle frequency band of DCT. This scheme provides a large embedding capacity which is about 11 bits for each block. Thus, this scheme provides significant distortion in quality of the watermarked image.

Das et al. [8] presented a blind watermarking based on the inter-block coefficient correlation. This scheme embeds a watermark image to the different adjacent blocks. This technique provides better robustness under JPEG compression than existing techniques. However, robustness of the watermarked image against noise addition, Gaussian filter and sharpening needs to be improved.

This paper proposes a new embedding technique by examining the middle frequency based on psychovisual threshold. Embedding regions are determined based on the lowest modified entropy. The lowest modified entropy value indicates the highest redundant image information. Watermark embedding is performed by modifying selected coefficients of the DCT middle frequencies of the image blocks. The watermark is scrambled before it is embedded into the selected coefficients. The proposed embedding technique is tested under different types of attack. Test results are given in terms of Normalized-Cross Correlation (NC) and Structural SIMilarity (SSIM) index in the watermarked image. This paper is organized as follows. Our modified entropy for representing visual characteristics is described in Section II. The watermark embedding and watermark extraction algorithms are presented in Section III. Experimental results are shown in Section IV. Finally, the conclusion is given in Section V.

## II. HUMAN VISUAL CHARACTERISTICS

Human visual characteristics can be presented by entropy and edge entropy [2], which can be used to select the region of embedding blocks. Entropy and edge entropy values indicate less distortion areas that are suitable for watermark embedding as these areas do not have significant distortion effect in the host image. Entropy and edge entropy are used to select the significant embedding watermark. The modified entropy of an  $N$ -state is defined by:

$$E_p = -\sum_{i=1}^N \frac{p_i \exp(1-p_i) + p_i \log_2(p_i)}{2} \quad (1)$$

where  $p_i$  denotes the occurrence probability of  $i$ -th pixel with  $0 \leq p_i \leq 1$  and  $1-p_i$  represents the uncertainty or ignorance of the pixel value.

## III. PROPOSED SCHEME

Referring to the psychovisual threshold, the gaps between the psychovisual threshold and minimum quantization values of JPEG compression can be utilized to embed the watermarks.

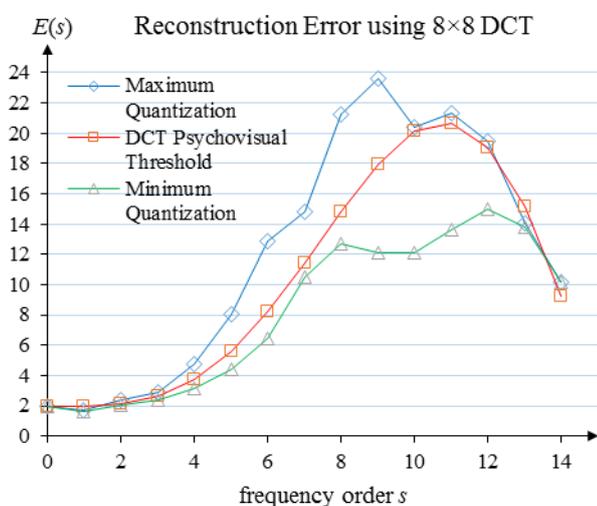


Figure 1. DCT psychovisual threshold

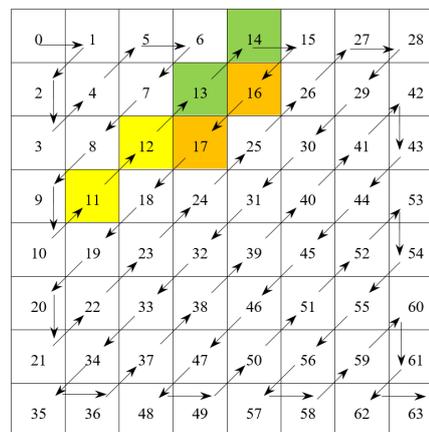


Figure 2. Selection of DCT coefficients using zig-zag order based on psychovisual threshold.

These gaps can be noticed between the curve marked with circle and the curve marked with triangle as shown in Fig. 1 which plots reconstruction errors for DCT psychovisual threshold, maximum quantization and minimum quantization values, where  $E$  denotes reconstruction errors. A detail description can be found in [9]-[20]. The DCT coefficients are arranged in a zig-zag order as shown in Fig. 2. Referring to the gap between the psychovisual threshold (the curve marked with circle) and the minimum quantization value (the curve marked with triangle) in frequency orders 4 to 5, some coefficients are selected for potential watermark embedding. These selected coefficients are ordered into a vector of coefficients as shown in Fig. 3. We assume that those locations provide less distortion and more robustness under image compression attacks.

This experiment uses two thresholds:  $\alpha$  and  $\beta$ , where  $\alpha$  denotes a threshold for the first coefficient and  $\beta$  represents a



Figure 3. Selected DCT coefficients

threshold for the second coefficient. The thresholds  $\alpha$  and  $\beta$  are set as negative or positive values based on the certain condition in Algorithm 1.

### Algorithm 1: Setup of threshold values

**Input:**  $T, A$   
**Output:**  $\alpha, \beta$

1. **for**  $x=0$  to 2
2.     **if** ( $A(2x) < 0$ ) **then**
3.          $\alpha = -T$ ;
4.     **else**
5.          $\alpha = T$ ;
6.     **end (if)**
7.     **if** ( $A(2x+1) < 0$ ) **then**
8.          $\beta = -T$ ;
9.     **else**
10.          $\beta = T$ ;
11.     **end (if)**
12. **end (for)**

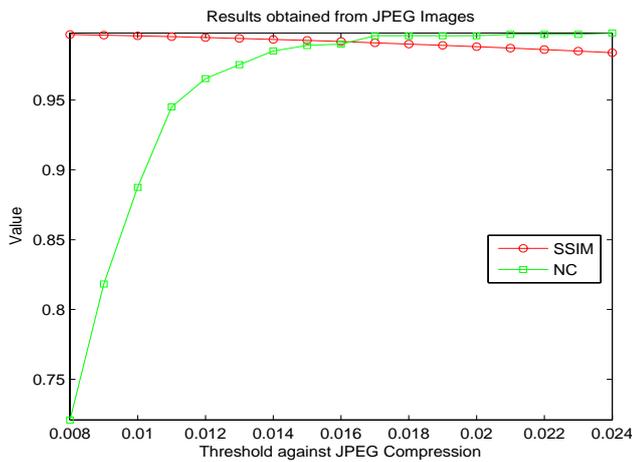


Figure 4. Trade-off between SSIM and NC values for Lai scheme [3]

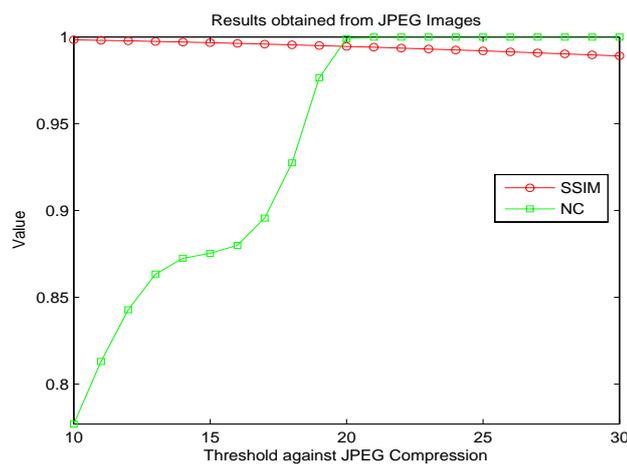


Figure 5. Trade-off between SSIM and NC values for our scheme

In the algorithm,  $T$  represents a threshold value of the trade-off between the imperceptibility and robustness of watermark image under JPEG compression.  $T$  is measured based on the relationship between SSIM values and NC values. Based on the experimental results, we find an optimal threshold  $T$  for Lai scheme [3] and our proposed scheme as shown in Figs. 4 and 5. We find an optimal threshold for  $U$  orthogonal matrix from the hybrid DCT-SVD as 0.016; and the middle DCT coefficient as about 20. The embedding and extraction algorithms are described in Algorithms 2 and 3.

#### A. Watermark Embedding Algorithms

Watermark embedding process is described in Algorithm 2. In the algorithm, watermarks are embedded in the frequency coefficients of each block of an image using the technique given in step 8. Note that for  $x=0,1$ , and  $2$ ,  $A(2x)$  represents  $A(0)$ ,  $A(2)$ ,  $A(4)$  and  $A(2x+1)$  denotes  $A(1)$ ,  $A(3)$ ,  $A(5)$  as shown in Fig. 3.  $\alpha$  and  $\beta$  present variant thresholds for watermark embedding. If  $A(y)<0$ , for  $y=0,1, \dots, 5$ , the threshold value is negative, otherwise the threshold is positive as given in Algorithm 1.

**Algorithm 2:** Embedding watermark

---

**Input:** Host image; watermark; threshold ( $\alpha$  and  $\beta$ )

**Output:** Watermarked image containing a logo

---

**Pre-processing:**

- Step 1: The cover image of size  $M \times N$  is divided into  $8 \times 8$  non-overlapping blocks.
- Step 2: Modified entropy for each non-overlapping block is computed by Equation (1).
- Step 3: Blocks that have lowest modified entropy values are selected and their  $x$  and  $y$  coordinates are saved.
- Step 4: Selected blocks are converted into frequencies using DCT
- Step 5: DCT coefficients are converted into a vector by using zig-zag order as shown in Fig. 2.
- Step 6: Certain coefficients based on the psychovisual threshold as shown in Fig. 3 are selected.
- Step 7: The binary watermark is scrambled using Arnold chaotic map.

**Watermark embedding:**

- Step 8: Each bit of binary watermark is embedded according to the rules as follows:

```

for x=0 to 2
  if U<length(Watermark) then
    if Watermark(U)=1 then
      if (|A(2x)|<|A(2x+1)|) then
        C=A(2x);
        A(2x)=A(2x+1)+β;
        A(2x+1)=C;
      else
        A(2x)=A(2x)+α;
        A(2x+1)=A(2x+1);
      end (if)
    else
      if (|A(2x)|<|A(2x+1)|) then
        A(2x)=A(2x)+β;
        A(2x+1)=A(2x+1);
      else
        C=A(2x);
        A(2x)=A(2x+1);
        A(2x+1)=C+α;
      end (if)
    end (if)
  end (for)
  
```

**Post-processing:**

- Step 9: The modified values of the vector are set into the two-dimensional matrix in Fig. 2.
  - Step 10: The inverse DCT on each selected block is performed.
  - Step 11: The modified selected blocks are merged to reconstruct the watermarked image.
- 

#### B. Watermark Extraction Algorithms

Watermark extraction process is described in Algorithm 3.

**Algorithm 3:** Watermark extraction

**Input:** Watermarked image;  $x$  and  $y$  coordinate location of selected block; Threshold ( $\alpha$  and  $\beta$ )

**Output:** Watermark recovery

---

**Pre-processing:**

- Step 1: The  $x$  and  $y$  coordinates are used to determine the embedded blocks. Then, each selected block is transformed using DCT.
- Step 2: Each block-based DCT is converted into a vector using zig-zag order.

**Watermark extraction:**

- Step 3: Certain coefficients as shown in Fig. 3 are selected. Each watermark bit is recovered according to the following rule as follows:  
**if**  $A_k < A_{k+1}$  for  $k=0,2,4$  **then**  
     watermark bit =1,  
**else**  
     watermark bit =0.  
**end (if)**

**Post-processing after embedding:**

- Step 4: The scrambled binary watermark is inversed by Arnold chaotic map to obtain the original watermark.

IV. EXPERIMENTAL RESULTS

Our scheme has been evaluated for four grayscale images of size  $512 \times 512$  pixels as shown in Fig. 6. A binary logo image with  $32 \times 32$  pixels is scrambled by Arnold chaotic map to provide extra security as given in Fig. 7. The proposed scheme is tested under different types of attacks such as average filter, wiener filter, median filter, Gaussian low pass, Gaussian noise, speckle noise, pepper and salt noise, sharpening, Poisson noise, adjust, histogram equalization attack, cropping, scaling, JPEG compression and combination attacks.



Figure 6. (a) Lena, (b) Cameraman, (c) Airplane, (d) Pepper images

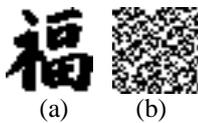


Figure 7. (a) Binary watermark, (b) Scrambled watermark

The performance of our scheme is evaluated by imperceptibility and robustness. Imperceptibility is measured by Structural Similarity (SSIM) index defined by:

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (2)$$

where  $\alpha > 0, \beta > 0, \gamma > 0$ , are parameters which can be adjusted to signify their relative importance. Robustness of the proposed scheme has been evaluated after applying different types of attack using Normalized Cross-Correlation (NC) and Bit Error Rate (BER) which are defined by:

$$NC = \frac{\sum_{i=1}^M \sum_{j=1}^N W(i, j) \cdot W^*(i, j)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N W(i, j)^2 \sum_{i=1}^M \sum_{j=1}^N W^*(i, j)^2}} \quad (3)$$

$$BER = \frac{\sum_{i=1}^M \sum_{j=1}^N W(i, j) \oplus W^*(i, j)}{M \times N} \quad (4)$$

where  $W^*(i, j)$  is the extracted watermark and the  $W(i, j)$  is the original watermark.  $M$  and  $N$  denote the row and column size. A trade-off between imperceptibility (SSIM) and robustness (NC) is used to determine the threshold value. The experimental results of our scheme are given in Table I.

TABLE I. WATERMARKED IMAGE QUALITY

Image	Lai Scheme [3]		Our Scheme	
	ARE	SSIM	ARE	SSIM
Lena	0.3751	0.9919	<b>0.3126</b>	<b>0.9946</b>
Cameraman	<b>0.2729</b>	0.9935	0.3038	<b>0.9936</b>
Airplane	0.6450	0.9834	<b>0.3035</b>	<b>0.9935</b>
Pepper	0.3696	0.9923	<b>0.3173</b>	<b>0.9940</b>
Average	0.4156	0.9902	<b>0.3093</b>	<b>0.9939</b>

Referring to Table I, our scheme has been verified by ARE and SSIM values that demonstrate higher imperceptibility and less distortion in the watermarked images. The details of NC values of the proposed scheme under different types of attack are shown in Fig. 8. All the visual perception of extracted

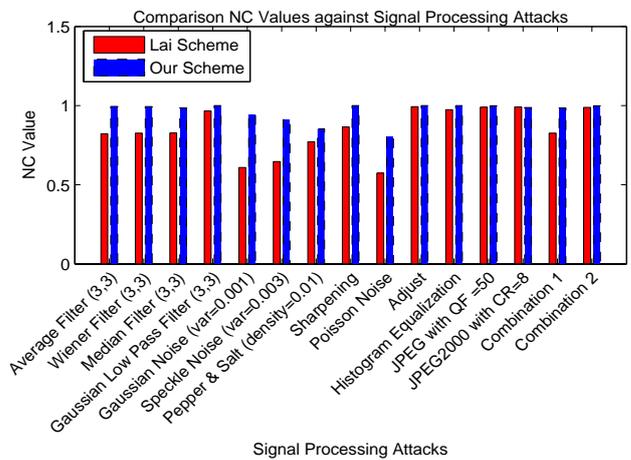


Figure 8. Comparison NC values between Lai scheme and proposed scheme against image processing attacks

watermark under different types of image-processing and geometrical attacks are presented in Fig. 9. Comparison of the proposed scheme and existing techniques is presented in Table II. Our technique produces higher NC values than Lai scheme. This proves that the diagonal frequency order based on the psychovisual threshold can better defend against different types of attack. The proposed scheme proves the supremacy over the existing techniques in terms of adding noise, median filter, low-pass filter, and histogram equalization.

TABLE II. COMPARISON NC VALUES OF THE PROPOSED SCHEME AND EXISTING SCHEME UNDER DIFFERENT TYPES OF ATTACK FOR WATERMARKED LENA IMAGE

Attack	Our scheme	Lai [3]	Roy [7]	Das [8]
Gaussian low-pass Filter (3,3)	<b>1</b>	0.9672	1	0.9118
Image sharpening	<b>1</b>	0.8653	1	0.9327
Centred cropping 25% (128x128 by white)	<b>1</b>	0.9941	0.9969	0.9954
JPEG (QF=50)	0.9990	0.9902	<b>1</b>	0.9810
JPEG (QF=70)	<b>1</b>	1	1	0.9918
Combinational attacks pepper & salt (density=0.003) and Median Filter (3,3)	<b>0.9563</b>	0.8227	0.8746	0.8641
Salt & pepper noise, density = 0.01	0.8314	0.7745	<b>0.9466</b>	0.8122
Gaussian noise variance=0.001	<b>0.9393</b>	0.6348	0.8998	0.8816
Median filtering with filter size 3x3	<b>0.9854</b>	0.8272	0.9687	0.9118
Histogram equalization	<b>1</b>	0.9747	0.9313	0.9253

## V. CONCLUSION

This paper presents a new embedding technique for image watermarking based on the psychovisual threshold. The watermarks are embedded by examining the psychovisual threshold on the selected middle frequency-coefficient pairs. A watermark image is not directly embedded into a host image, certain coefficients in the middle frequencies are modified with a threshold. Then the watermark image is scrambled before it is embedded into the host image. The proposed scheme utilizes modified entropy to determine the embedding regions. Furthermore, our scheme is evaluated under different types of attacks such as image noise, low-pass filter, sharpening, median filter, JPEG and JPEG2000, geometrical attacks like image cropping and scaling. The proposed scheme is compared to the existing methods in terms of imperceptibility and robustness. The experimental results verified that our scheme performs better than existing techniques in terms of SSIM and NC values. The proposed scheme also proves the superiority under combinational attacks.

## ACKNOWLEDGMENT

This work was supported by Fundamental Research Grant Scheme (FRGS) No. RDU160102 from Ministry of Higher Education, Malaysia.

## REFERENCES

- [1] F. Ernawan, M. N. Kabir, M. Fadli, Z. Mustafa, "Block-based Tchebichef image watermarking scheme using psychovisual threshold," *International Conference on Science and Technology-Computer (ICST)*, pp. 6-10, 27-28 Oct. 2016.
- [2] F. Ernawan, M. Ramalingam, A. S. Sadiq, Z. Mustafa, "An improved imperceptibility and robustness of 4x4 DCT-SVD image watermarking using modified entropy," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 9, no. 2-7, pp. 111-116, 2017.
- [3] C.C. Lai, "An improved SVD-based watermarking scheme using human visual characteristics," *Optics Communications*, vol. 284, no. 4, pp. 938-944, 2011.
- [4] X. Wu, W. Sun, "Robust copyright protection scheme for digital images using overlapping DCT and SVD," *Applied Soft Computing*, vol. 13, no. 2, pp. 1170-1182, 2013.
- [5] N.A. Abu, F. Ernawan, N. Suryana, S. Sahib, "Image watermarking using psychovisual threshold over the edge," *Information and Communication Technology*, vol. 7804, pp. 519-527, 2013.
- [6] F. Ernawan, "Robust image watermarking based on psychovisual threshold," *Journal of ICT Research and Applications*, vol. 10, no. 3, pp. 228-242, 2016.
- [7] S.Roy, A.K. Pal, "A blind DCT based color watermarking algorithm for embedding multiple watermarks," *AEU International Journal of Electronics and Communications*, vol. 72, pp. 149-161, 2017.
- [8] C. Das, S. Panigrahi, V.K. Sharma, Mahapatra K.K., "A novel blind robust image watermarking in DCT domain using inter-block coefficient correlation," *AEU International Journal of Electronics and Communications*, vol. 68, no. 3, pp. 244-253, 2014.
- [9] F. Ernawan, M. N. Kabir, J. M. Zain, "Bit allocation strategy based on psychovisual threshold in image compression," *Multimedia Tools and Applications*, pp. 1-24, 2017.
- [10] F. Ernawan, N. Kabir, K.Z. Zamli, "An efficient image compression technique using tchebichef bit allocation," *Optik - International Journal for Light and Electron Optics*, vol. 148, pp. 106-119, 2017.
- [11] F. Ernawan, Z. Mustafa, L.B. Aji, "An Efficient Image Compression Using Bit Allocation based on psychovisual Threshold," *Information (Japan)*, vol. 9(9B), pp. 4177-4182, 2016.
- [12] F. Ernawan, S.H. Nugraini, "The optimal quantization matrices for jpeg image compression from psychovisual threshold," *Journal of Theoretical and Applied Information Technology*, vol. 70, no. 3, pp. 566-572, 2014.
- [13] F. Ernawan, N.A. Abu, N. Suryana, "An adaptive JPEG image compression using psychovisual model," *Advanced Science Letters*, vol. 20, no. 1, pp. 26-31, 2014.
- [14] N.A. Abu, F. Ernawan, "A novel psychovisual threshold on large DCT for image compression," *The Scientific World Journal*, vol. 2015, no. 2015, pp. 001-011, 2015.
- [15] N.A. Abu, F. Ernawan, N. Suryana, "A generic psychovisual error threshold for the quantization table generation on JPEG image compression," *9th International Colloquium on Signal Processing and its Applications*, pp. 39-43, 2013.
- [16] F. Ernawan, S.H. Nugraini, "The optimal quantization matrices for JPEG image compression from psychovisual threshold," *Journal of Theoretical and Applied Information Technology*, vol. 70, no. 3, pp. 566-572, 2014.
- [17] F. Ernawan, N.A. Abu, N. Suryana, "An optimal tchebichef moment quantization using psychovisual threshold for image compression," *Advanced Science Letters*, vol. 20, no. 1, pp. 70-74, 2014.
- [18] F. Ernawan, N.A. Abu, N. Suryana, "Adaptive tchebichef moment transform image compression using psychovisual model," *Journal of Computer Science*, vol. 9, no. 6, pp. 716-725, 2013.
- [19] N.A. Abu, F. Ernawan, "Psychovisual threshold on large tchebichef moment for image compression," *Applied Mathematical Sciences*, vol. 8, no. 140, pp. 6951-6961, 2014.
- [20] F. Ernawan, N.A. Abu, N. Suryana, "TMT quantization table generation based on psychovisual threshold for image compression," *International Conference of Information and Communication Technology*, pp. 202-207, 2013.



Figure 9. Result under different type of attacks and recovered watermark images from (a) Average Filter (2.2) (b) Wiener Filter (2.2) (c) Median Filter (2.2) (d) Gaussian Low Pass Filter (2.2) (e) Gaussian Noise (var=0.001) (f) Speckle Noise (var=0.003) (g) Pepper and Salt Noise (density=0.01) (h) Sharpening (i) Poisson Noise (j) Adjust (k) Histogram Equalization Attack (l) Centred Cropping 50% (256x256 by black) (m) Centred Cropping 50% (256x256 by white) (n) Cropping rows off 25% (128 rows by black) (o) Cropping columns off 25% (128 columns by black) (p) Scaling 0.8 (q) JPEG with QF=40 (r) JPEG with QF=50 (s) Combination Pepper & Salt (density=0.003) and Median Filter (3.3) (t) Combination JPEG with QF=50 and Centre Cropping 25%

# Data Protection by Design in Systems Development

## From legal requirements to technical solutions

Fredrik Blix, Salah Addin Elshekeil, Saran Laoyookhong

Department of Computer and Systems Sciences

Stockholm University

Stockholm, Sweden

blix@dsv.su.se

**Abstract**—Data protection by design is a principle to systems development meaning that the protection of personal data is built into the systems design from the start. For many jurisdictions, this principle is becoming a legal requirement. Using a research approach based on design science, a framework is constructed helping systems developers achieve privacy by design in a systematic manner. The framework articulate how the business requirements can be captured, assessed, and implemented in the systems development. Examples of how the data protection principles can be concretely implemented is also presented.

**Keywords:** Privacy by design, data protection, systems development, information security, general data protection directive, GDPR

### I. INTRODUCTION

The European General Data Protection Directive GDPR is the new European-wide regulation aimed at all businesses and other legal entities that are processing personal data (information that is relating to an identified or identifiable living person). In practice, GDPR means that businesses will have to do a lot of work to secure the processing of such data. The difference between general cybersecurity efforts and these efforts are that GDPRs focus is not on cybersecurity risks in general – it is concerned with the protection of basic human rights (such as the right not to be discriminated against because of ethnicity, sex, etc.). One of the requirements of GDPR is that all systems – both old legacy systems and new systems that are developed now, if they handle personal data – need to be designed with privacy (that is, data protection) in mind. This principle is called *data protection by design*.

The problem is that most businesses have not been required to consider privacy by design when they developed their systems, since this has not been a legal requirement before. As more and more jurisdictions require this, the problem is growing. Even though there are no systematic surveys, one can estimate that probably less than 2 systems out of 10 would be considered designed with privacy in mind if analyzed. The problem is owned by businesses owning (or renting) systems processing personal data.

The practical relevance of this study is take one step towards the helping businesses solve how they can achieve data protection by design in a cost-effective manner. The theoretical relevance is filling a void in studies on data protection by

explicitly tackling privacy by design – not in principle only, but how it relates directly to systems development.

This paper presents the study and its results. In section 2, presents a systematic literature review in summary. In section 3 the overall research approach and - method is covered. Section 4, presents the resulting framework and finally section 5 contains concluding remarks.

### II. LITERATURE REVIEW

The authors aimed to answer the following questions through the literature review:

- What are the existing privacy by design principles?
- How can privacy by design principles be implemented in IT Systems?
- Are there any existing case studies concerned with the implementation of privacy by design?
- Are there any established standards for implementing privacy by design?

The systematic literature review followed was based on Rondolph's guidelines [1] on how to conduct a thorough literature review. The sources of the data collection were academic and popular articles identified through online digital libraries for research publications as well as through web search engines. The data was then evaluated, analyzed to find patterns, and the answers to the questions posed.

#### A. Definition of privacy by design

Privacy has been a controversial topic for long time; there were many attempts to define what privacy is according to the context and the environment [2]. Warren and brandies defined privacy term and the “right to be left alone” back in 1890 [3]. After that, there were number of definitions of privacy. On the legal side, privacy was perceived and defined differently from country to country, especially in the EU where privacy was lawfully granted for EU Citizens in the European commission of human rights.

Privacy by design was introduced in the 90's by Ann Cavoukian who defined the concept as “Refers to the philosophy and approach of embedding privacy into design specification of various technologies” [4]. The seven principles were introduced

and became the foundation of Privacy by Design concept for many research. Furthermore, P. Schaar [5] introduced six different objectives to consider when designing processing systems, and argued that these principles should be bonded for technology designers, and producers. Also, the 11 privacy principles were introduced according to the standard ISO/IEC 29100 [6].

*B. Privacy by design principles*

There is no consensus about privacy by design principles, authors and regulations introduced different principles. Ann Cayoukian [4] who defined privacy by design introduced seven principles, these principles are generic and not easily can be translated into IT requirements. While Schaar [5] has 6 principles, which are missing important principle such as accountability. Standards such as ISO 29100 introduced 11 principles, that are detailed and overlapped with GDPR and other privacy principles. Purpose limitation for example is newly introduced in GDPR, while data quality principle was introduced by Schaar [5], but ISO 29100 called it accuracy and quality. Accountability principle was introduced in ISO29100 and GDPR. Consent is also another principle was introduced by ISO29100, but no one else introduced this principle. In GDPR it is considered part of the lawfulness, fairness and transparency. GDPR [7] the new EU data protection regulations has introduced seven data protection principles table 1, which will be the focus on this research. These principles have to be implemented in the systems that process personal data by design and by default according to GRPR article 25. All establishments who are processing EU Citizens personal data inside and outside EU will be subject for compliance. The notation (DPP 1-7) is used to refer to these principles in this paper.

Table I. Data Protection Principles

#	Notation	Data Protection Principle
1.	DPP1	Lawfulness, fairness and transparency
2.	DPP2	Purpose Limitation
3.	DPP3	Data Minimization
4.	DPP4	Storage Limitation
5.	DPP5	Integrity and Confidentiality
6.	DPP6	Accuracy
7.	DPP7	Accountability

*C. Approaches to privacy by design implementation*

Privacy implementation approaches can be divided into two, management and process approach and engineering and technology approach. In Fig. 1, previous researches grouped based on patterns and approaches.

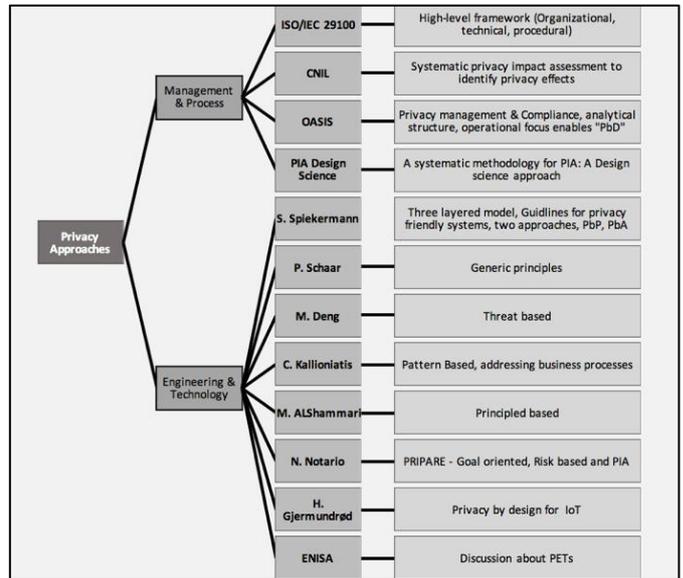


Figure 1. Privacy by design implementation approaches

*D. Approaches to privacy impact assessment (PIA) and related approaches*

Wright defined that Privacy Impact Assessment (PIA) is “a form of risk assessment, an integral part of risk management” [8], while Wadhwa [9] viewed the PIA as a privacy management tool. These definitions can be implied that the PIA can manage the risk concerning the privacy and include a similar process to the risk management. Moreover, CNIL defined a privacy risk given the processing of personal data that has impacts on the privacy of data subjects and also stated that another aspect of the PIA is to determine suitable measures that complied with the legal requirements [10].

The development of PIA went back since 1996 when the US Internal Revenue Service issued its IRS PIA, endorsed by the Federal Chief Information Officers Council [10]. Throughout the years, public sectors concerning information and privacy commissioner from various countries like Hong Kong, Canada, New Zealand and Austria has released their version of PIA to suit each need. Then, in December 2007, the Information Commissioner’s Office (ICO)’s PIA Handbook was released and is considered the first PIA Handbook in Europe. After that, in 2008, the International Organization for Standardization (ISO) published the PIA 22307:2008 standard. In the following year, the European Commission (EC) issued a recommendation on the implementation of privacy and data protection principles for RFID application which was endorsed by Article 29 [11] and later became EC’s PIA framework for RFID in February 2011. Also, there are similar tools such as Data Protection Impact Assessment (DPIA) which was introduced by the European Commission as an evaluation and decision-making tool.

*E. Challenges of privacy by design*

Privacy by design is a new topic, with divergent views on both the principles and the approaches. There is no comprehensive framework that covers all aspects of privacy and risk-based approach. The lack of framework might be attributed

to the controversial topic of privacy and no consensus on the privacy principles. As a result, there are different views on how privacy can be realised in organizations in general and in IT systems. At the same time, few case studies represent privacy implementation in IT systems. In conclusion, no established comprehensive risk-based standards or framework covers legal, technical and organisational requirements.

### III. METHOD

Since there is no comprehensive framework for translating GDPR into system requirements, the focus is on develop a new framework and evaluate it through application. This research work follows a design science methodology to build the artefact (here, the framework) and use a case study to evaluate and improve the artefact. Both interviews and literature review were used in parallel to explicate problem and define the requirements of the artefact. There was a total of five interviews conducted with both management teams from an IT security company and members from start-up company who is developing a system that processes the sensitive personal data based on artificial intelligence.

In the first interview, together with the management team, the researchers concluded that the lack of knowledge to interpret the GDPR into clear technical requirements was the most pressing problem. The CEO of the start-up company attended the second interview, resulting in that all respondents agreed on the need to development a versatile and flexible framework with the goal to guide businesses to identify the needed organisational and technical information- and cybersecurity measures, resulting in privacy by design, for an IT system. The researchers used creative methods to collect and select ideas using perspectives as *'how the framework should be built'* and *'how the developed framework will be used in the research work'*. The researchers developed the framework by identifying GDPR requirements, data protection risks and cybersecurity measures (security controls), map them with data protection principles, and finally by identifying existing approaches and security requirements as well as identifying other important factors.

The rest of the interviews were conducted while applying the developed framework through the case study, to form a basis for the evaluation. The researchers analyzed the system of the start-up company and their business context to select an initial set of appropriate security measures, then assessed the system to define the data protection requirements. Upon applying the framework to the case study, the researchers learned about the necessity to improve the framework and synthesised the new findings to create an improved and refined version of the suggested framework (as presented in conclusions).

### IV. RESULTS

#### A. Construction of a framework

The data protection principles (DPP1-7 in Table 1), identified through the literature review, forms the basis for the results since these principles represent what is to be achieved by any privacy efforts. In addition, these principles are referred to from the European GDPR regulation, and thus forms part of legal requirements in many jurisdictions. The literature review also revealed other approaches to privacy, privacy impact assessment and privacy risk identification that proved useful.

Even though the set of potential privacy measures (organizational and technical) that can be applied is almost unlimited, a limited number of such measures – the most commonly used – were listed and subsequently mapped back to the seven data protection principles. The mapping answers the question, “What privacy measures can help satisfy which the data protection principle *DPPX*?”.

Which privacy measures are *suitable* in each situation depends on many other internal (systems-related) and external (outside of the system) factors. GDPR mentions seven factors which were supplemented with four factors from other sources.

In total, this forms the basis for the suggested framework presented here (see figure 2). This framework is the improved and refined version after the evaluation (see *method* chapter for evaluation details).

#### B. The APSIDAL framework

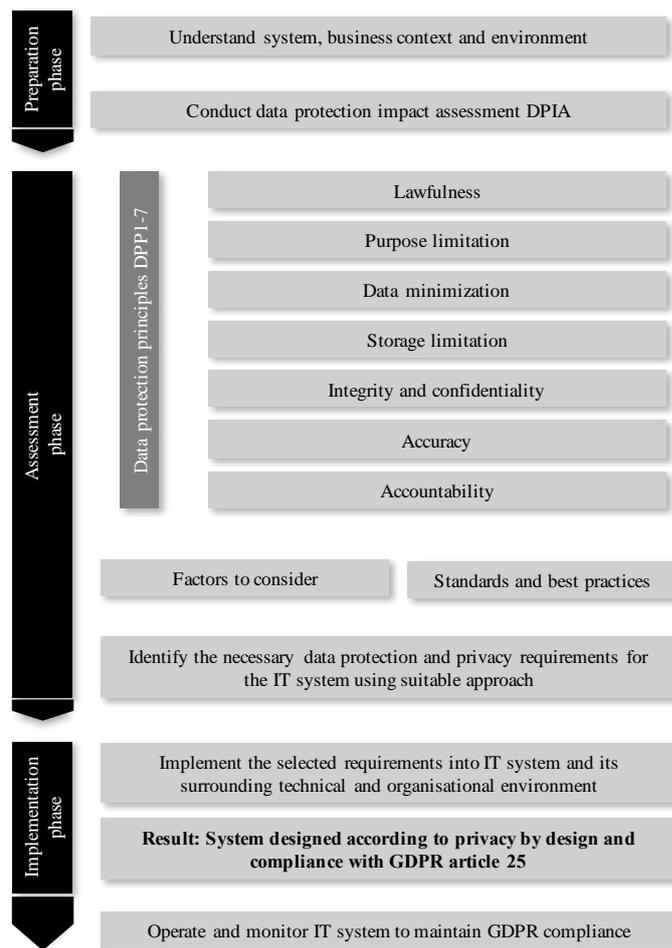


Figure 2: Framework for Privacy by Design, APSIDAL

### V. CONCLUSIONS

By applying the suggested APSIDAL framework, businesses and other organizations can take an important step towards complying with the principle of privacy by design as required in an increasing number of jurisdictions. The framework shows step-by-step activities that needs to be carried out, tied to the

three main identified phases (preparation, assessment and implementation phase).

For those needing more specific guidance, appendix 1 shows the identified organizational and technical data protection and security measures found potentially valuable, tied to each of the seven data protection principles (see Appendix 1).

REFERENCES

- [1] J. Randolph, "A Guide to Writing the Dissertation Literature Review," *Practical Assessment, Research & Evaluation*, vol. 14, no. 13, pp. 1-13, 2009.
- [2] C. T. Di Iorio and F. Carinci, *Privacy and Health Care Information Systems : Where Is the Balance?*, Berlin: Springer, 2013, pp. 77-105.
- [3] S. D. Warren and L. D. Brandeis, "The Right to Privacy," *Harvard Law Review.*, vol. IV, no. 5, p. 193–220, 15 Dec 1890.
- [4] A. Cavoukian, "Privacy by Design - Information and Privacy Commissioner," 2007. [Online]. Available: <https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf>. [Accessed 22 9 2017].
- [5] P. Schaar, "Privacy by Design," *Identity in the Information Society*, vol. 3, no. 2, p. 267–274, August 2010.
- [6] International Standard ISO/IEC 29100, "Privacy framework Technologies 29100," ISO/IEC, 2011.
- [7] Official Journal of the European Union, "General Data Protection Regulation - GDPR," *Official Journal of the European Union*, vol. 2016/679, p. 119/1, 27 4 2016.
- [8] D. Wright, R. Gellert, S. Gutwirth and M. Friedewald, "Minimizing technology risks with PIAs, precaution, and participation," *IEEE Technology and Society Magazine*, vol. 30, no. 4, pp. 47-54, 2011.
- [9] K. Wadhwa and R. Rodrigues, "Evaluating privacy impact assessments," *Innovation: The European Journal of Social Science Research*, vol. 26, pp. 161-180, 2013.
- [10] Commission Nationale de l'Informatique et des Libertés, "PRIVACY IMPACT ASSESSMENT (PIA)," 6 2015. [Online]. Available: <https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-1-Methodology.pdf>. [Accessed 22 9 2017].
- [11] European Commission, "Privacy and Data Protection Impact Assessment Framework for RFID Applications," no. January, pp. 1-24, 2011.

I. APPENDIX

### DPP1: Lawfulness, Fairness and transparency

**GDPR Provision**

*"processed lawfully, fairly and in a transparent manner in relation to the data subject"*  
GDPR Art.5.1(a)

**Objective**

Processing of personal data no matter how minor the processing is has to stand on a firm legal basis. The data controllers have to provide clear views on how the processing works and the consequences on the data subject before collecting and processing the data.

**Organizational Measures** 

**Strategy:** embed privacy into the organization strategy to demonstrate strong commitment to privacy

**Policies, processes, and procedures:** to assess, manage the lawfulness and the adherence to the legal requirements at all stages of the data processing. This includes the roles and responsibilities of different individuals involved in the data processing

**Legal Measures:** clear guidelines and measure of selecting, enforcing and documenting the right legal bases of the processing

**Technical Measures** 

**Embedded Transparency Measures:** embedding necessary forms, dialogues, notifications into information systems. Such as asking for consent before collecting location data of a mobile app user

**Embedded Legal Measures:** For example, an embedded database that capture consents from users of the information system, and map it to the user data. This can help the users to exercise their rights and the organization to manage their obligations

**Non-Repudiation Services:** Implementing the non-repudiation service from the data subject. For example, the digital signatures shall be implemented when the collected data is sensitive.

A. DPP1 Lawfulness, fairness and transparency

### DPP2: Purpose Limitation

**GDPR Provision**

*"collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes"* GDPR Art. 5.1(b)

**Objective**

Purpose limitation requires the data controllers to define the reasons for data collection and processing. Both data controllers need to ensure that the data is processed according to the original purpose. The data should not be processed when the purpose has changed without firm legal ground. Also, data traceability throughout the data lifecycle is needed to know when the data processing is not according to the original purpose anymore.

**Organizational Measures** 

**Strategy:** Privacy strategy aligned with organization strategy, and in some cases changing to the business model to uphold the purpose limitation

**Policies, processes, and procedures:** to assure the personal data is used to the purpose it was collected for. The processing must be according to the purpose during the data processing lifecycle. Any deviation must be captured, and dealt with.

**Technical Measures** 

**Data Inventory:** functionality to trace back the purpose of the data collection, this can be achieved by adding meta-data or tagging to the personal data, this can be linked to the data minimization and storage limitation purposes

**Reporting:** automatic notification and reporting functionality embedded into the information systems

B. DPP2 Purpose limitation

### DPP3: Data Minimization

**GDPR Provision**  
*"adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed"* GDPR Art. 5.1(c)

**Objective**  
 Data minimization involves reducing the amount of the collected data to what is necessary that the interaction with the data subject is satisfied without.

**Organizational Measures** 

**Strategy:** embed privacy into the organization strategy to demonstrate strong commitment to privacy

**Policies, processes, and procedures:** to assess, manage the lawfulness and the adherence to the legal requirements at all stages of the data processing. This includes the roles and responsibilities of different individuals involved in the data processing

**Identity & Access Management:** clear guidelines and measure of selecting, enforcing and documenting the right legal bases of the processing

**Technical Measures** 

**Centralized Storage:** limiting the distribution of storing personal data in distributed environment, and consolidate the personal data in central storage can contribute to data minimization

**Data Pseudonymization:** Encrypting some of the data fields that can link the data to persons such as names, birthdays, address. Keys should be handled carefully to prevent recovering the encrypted fields.

**Strip Unused meta data:** files, documents, videos can contain personal data about the author, the geographical location of the user and IP addresses that may not be necessary for the data processing. Cleaning these meta-data might be necessary

**Intermediary Proxies:** Data Collection and data transmission through intermediary proxies to filter and anonymize personal data such as IP addresses, and cookies

C. DPP3 Data Minimization

### DPP5: Storage Limitation

**GDPR Provision**  
*"kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organizational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject"* GDPR Art. 5.1(e)

**Objective**  
 Storage limitation principle focuses on keeping the identifiable data for only the period that the data serve its purpose. The data controllers have full responsibility to maintain track of the data and remove it when it is no longer being processed for its original purpose.

**Organizational Measures** 

**Awareness:** awareness can be vital to limit the storage of personal data. Employees tend to keep as much as date if they needed it in the future.

**Data Lifespan:** clear policies, and guidelines on the lifespan of every data sets. This needs to be defined with the business functions within the organization to avoid deleting necessary data that can be important for to conduct business or other obligations.

**Technical Measures** 

**Tractability:** functionality in the information system that can report and visualize the route the actual data data to the backup, and other distributed copies. This can enable organization to remove unnecessary data

**Self-wiping:** For example, an embedded database that capture consents from users of the information system, and map it to the user data. This can help the users to exercise their rights and the organization to manage their obligations

**Reporting:** Implementing the non-repudiation service from the data subject. For example, the digital signatures shall be implemented when the collected data is sensitive.

E. DPP5 Storage Limitation

### DPP4: Accuracy

**GDPR Provision**  
*"adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed"* GDPR Art. 5.1(c)

**Objective**  
 To take necessary steps to ensure the accuracy of data obtained and to verify the data source. Furthermore, any challenges to the accuracy of information shall be considered and keep up to date when necessary. In the information technology system, the accuracy of a digital record can also be measured by the ability of anyone to understand what the record says correctly.

**Organizational Measures** 

**Data Completeness Awareness:** raise the awareness among employees and anyone responsible for data collection and processing about the importance of collecting complete and accurate personal data, therefore, preventing wrong decisions

**Data Management:** roles, responsibilities, and procedures to insure data quality across the organization. Controls on data collection, and processes to allow data subjects keeps their data accurate and always up to date

**Data Normalization:** Guidelines and procedures to on the accepted data formats such as date, address, name, and language formats.

**Technical Measures** 

**Input Validation:** Validate all inputs to information systems, and allow only valid data and formats to the system. For example allow only UK date format.

**Data Dispute Handling:** functionalities to notify data subjects about any changes to personal data in the information systems, and measures to allow data subjects object or report any issues about their personal data processed in the systems

**Data Cleansing:** functionalities to analyze the correctness, completeness and consistency of personal data in information systems, and delete inaccurate or incomplete data.

D. DPP4 Accuracy

### DPP6: Confidentiality and Integrity

**GDPR Provision**  
*"processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures"* GDPR Art. 5.1(f)

**Objective**  
 Integrity and confidentiality are part of the foundation of information security. Protecting the privacy of the data subject by maintaining its integrity is to maintain the accuracy and consistency of stored data. Also, the confidentiality of the data is maintained by protecting the information from disclosure to unauthorised access. The measures for this principle shall be implemented and operated throughout the data lifecycle.

**Organizational Measures** 

**Identity & Access Management:** embed privacy into the organization strategy to demonstrate strong commitment to privacy

**Encryption & Key Management:** to assess, manage the lawfulness and the adherence to the legal requirements at all stages of the data processing. This includes the roles and responsibilities of different individuals involved in the data processing

**Physical security and data processing locations:** clear guidelines and measure of selecting, enforcing and documenting the right legal bases of the processing

**Technical Measures** 

**End to end encryption:** encrypting personal data at rest, in transit, and in use when possible. Depending on the sensitivity of personal data. Other techniques such as anonymization and pseudonymization can achieve the same objective

**Data Validation:** preserving the integrity of the data in transit and at rest by using hashes and input validation when capturing and storing personal data

**Authentication:** adequate level of authentication depending on the level of risk based on (something you are something you have or something you know) principle. Combination of two can be necessary for sensitive personal data processing

**Authorization:** access rights management, by using models such as role based or any other models to insure data always kept secret, and prevent the risk of disclosure of personal data to unauthorized individuals.

F. DPP6 Confidentiality and Integrity

## DPP7: Accountability

### GDPR Provision

“kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organizational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject”  
GDPR Art. 5.1(e)

### Objective

Accountability is a new concept introduced in GDPR. The data controllers must be accountable and be able to demonstrate compliance with the provisions of the regulations. The demonstration can be achieved in several ways, from not processing un-legal personal data, to implement the privacy principles into IT systems.

### Organizational Measures



**Strategy:** By considering privacy in to the business and operational strategies, organizations can demonstrate strong commitment to GDPR

**Policies, standards, awareness:** adopting best practices, enforcing policies, and implementing processes can demonstrate strong commitment to GDPR. Considering people, processes, and technologies may also be necessary

**Certified Products:** one way to demonstrate compliance is to adopt products and services that are build and certified with privacy in mind. EuroPriSe project is one example

### Technical Measures



**Authentication & Authorization:** embedding necessary forms, dialogues, notifications into information systems. Such as asking for consent before collecting location data of a mobile app user

**Tamper proof audit trails:** Information systems should provide audit logs that can be used for monitoring compliance, and any future investigation or reporting. Tamper proof audits trails should be implemented strictly in IT Systems

**Monitoring:** the goal with privacy should be always to prevent data leakage. Although monitoring can be seen as a re-active approach, but it can help the organization to prevent future incidents

**Data Loss Prevention:** preventing data from leakage on the way out from information systems or organizations considered to be one of the important functions for data protections.

## G. DPP7 Accountability

## **Session 2: Cyber Security**

Title: A Comparison Between API Call Sequences and Opcode Sequences as Reflectors of Malware Behavior

(Authors: Saja Alqurashi, Omar Batarfi)

Title: Android Botnet Detection: An Integrated Source Code Mining Approach

(Authors: Basil Alothman, Prapa Rattadilok)

Title: An Analysis of Home User Security Awareness and Education

(Authors: Fayez Alotaibi, Nathan Clarke, Steven Furnell)

Title: Monitoring Darknet Activities by Using Network Telescope

(Authors: Shaikha AlShehyari, Chan Yeob Yeun, Ernesto Damiani)

Title: Enhancing cyber security awareness with mobile games

(Authors: Faisal Alotaibi, Steven Furnell, Ingo Stenge, Maria Papadaki)

# A Comparison Between API Call Sequences and Opcode Sequences as Reflectors of Malware Behavior

Saja Alqurashi  
King Abdulaziz University  
Jeddah, Saudi Arabia  
ssmalqurshi@kau.edu.sa

Omar Batarfi  
King Abdulaziz University  
Jeddah, Saudi Arabi  
obatarfi@kau.edu.sa

**Abstract**— The volume of malware detected annually is increasing exponentially, and malware programs are written in such a way that they can often escape detection tools. Some are can even modify themselves and alter their appearance for each infection. Thus, for malware detection, it is important to analyze malware behavior, and application programming interface (API) call sequences and operational code (opcode) sequences usefully reflect the behavior of malware. Moreover, a hidden Markov model (HMM) is a robust learning model for malware detection. In this work, we therefore compared API call sequences and opcode sequences using the HMM learning model. The results showed that learning in API call sequences is more accurate than that of opcode sequences. We conclude that API call sequences are therefore better for malware detection.

**Keywords**—malware, HMM, API call sequences, opcode sequences

## I. INTRODUCTION

Malware is software developed for malicious intent [1]. It can steal sensitive data from a computer or infect files one after another, spreading the infection throughout the computer. Malware needs to be caught and removed from the infected computer promptly to avoid the leakage of sensitive data or any other malicious activity in the computer. Malware programs can be classified into viruses, worms, Trojans, spywares, adware and Rootkits [2].

Malicious codes are currently written in such a way that they can escape detection tools. Some malicious codes have the ability to modify themselves and alter their appearance for each infection. This is called the obfuscation technique [3].

To successfully detect malicious code, it is important to analyze the code's behavior and its effects on the system promptly using a process called malicious code analysis. Complete removal of the malicious code from the infected machine requires not only the deletion of malicious software but also the removal of associated processes, services, and registry entries. To accomplish this, it is necessary to understand the behavior of malicious code. Several detection techniques that analyze malicious code behavior have therefore been proposed [4].

There is a substantial research field now focusing on studying the working behavior of malware to determine the most effective analytical techniques. Recent studies have shown that

application programming interface (API) call sequences and operational code (opcode) sequences reflect malware behavior.

### A. API/system calls

An API (Application Programming Interface) refers to a set of instructions used to program software applications [5]. The core of an operating system usually contains a comprehensive API (called OS API or system calls) used by the programmers to request system resources as well as carry out a range of core functionalities. The system calls of a malware program reflect its high-level functionality and can be used to understand its overall behavior [5]. System call analysis is therefore an effective method of malware analysis [5].

### B. Operational Code (Opcode) sequences

An operational code (Opcode) “is the portion of a machine language instruction that specifies what operation to be performed by the CPU” [5]. A machine language instruction refers to an instruction in binary or hexadecimal code that can be executed directly on a CPU. Examples of such operations include addition, logical comparison, instruction control, and so on. Some machine language instructions only contain Opcode, while others contain one or more operands.

The complete set of machine language instructions for a software program comprises a sequence of Opcodes. The frequency and sequence of the Opcodes can be effectively used to analyze the behavior of a malware program [5].

### C. Problem statement

Like any other program, malware needs to perform actions to accomplish its objectives. These actions enable us to distinguish normal from malicious behavior. In this work, malicious behavior between API and opcode sets will be studied and compared using an HMM learning model. The main contribution of this work will be to compare malware analyses using dynamically extracted API sequences with a statically extracted opcode.

## II. RELATED WORK

### A. API call Sequences

Iwamoto et al. [6] proposed an automatic method of malware classification. Their method used a large sample of 4684 unpacked malware programs. After disassembling the sample malware, the authors analyzed their control flow to develop graphs of their function calls. Following this process, API sequence graphs were developed by extracting specific pairs of API calls. The similarity between the malware programs was then calculated by comparing API sequences using Dias Coefficient [6]. A hierarchical cluster analysis was then used to visually identify malware with similar features. During their experiment, the authors observed that the automatic “hierarchical cluster analysis” was conducted quickly and resulted in a significant number of clusters of the variants. The similarity between samples within the cluster was high -- 0.8 or greater. This work indicates that API call sequences can be effectively utilized as features of malware behavior in malware detection techniques.

Uppal et al. [7] presented a novel malware identification approach based on API call sequences. In their work, the authors used an odds ratio as a feature selection method. They then applied a support vector machine (SVM) algorithm to classify the suspect program as either benign or malware. This approach captures API calls by tracking the execution of the suspect program. Following API sequence extraction, the authors selected distinct API sequences in two steps: first, they generated the call grams; second, they calculated the odds ratio of each gram and generated the feature vector. Various machine-learning algorithms, such as Naïve Bayes, SVM, Decision Tree, and Random Forest, were then applied to construct the proposed model for classification. The SVM was the most accurate of all the algorithms. This work also concludes that API call sequences can be effectively used in behavior-based malware detection.

### A. Opcode sequences

Moskovitch et al. [9] proposed the use of n-grams of opcode to detect malware. The opcodes were generated by disassembling the executable files. Using opcode n-gram features, they trained several classifiers including decision trees, SVM, Naïve Bayes, KNN, and similarity based classifiers. They implemented the proposed method on 30,000 files and the results yielded an accuracy greater than 99%. They found that the best algorithms were decision tree (J48), KNN and SVM. Overall, J48 performed the most effectively. The authors concluded that opcodes can be successfully used to detect malware.

Santos et al. [10] proposed a novel algorithm to detect variants of known malware. They evaluated the similarity measure for each malware and its set variant using frequency opcode sequences. Additionally, they also evaluated the similarity of the suspect software to the entire dataset of benign files.

The authors in [11] aimed to overcome the limitations of static malware analysis by proposing a two-phase framework. The first phase analyzes the dynamic behavior of the malware and the second phase classifies malware behavior based on the outcome of the dynamic behavior analysis. This analysis was

performed in two steps: runtime analysis and resource monitoring. The malware behavior of each sample was then classified using artificial intelligence (AI) techniques. The proposed framework produced favorable results in terms of malware detection.

### B. The Hidden Markov Model as a malware detection technique

Priyadarshi et al. [12] developed a “code emulator” to run the malware in an emulated environment. The code emulator comprises five major components: CPU emulation, memory, hardware, operating system, and the emulation controller and analyzer. The code emulator can remove the instructions that were inserted using code obfuscation techniques. Once the metamorphic virus is morphed by the code emulator, HMM can successfully distinguish between viruses and normal files.

The authors in [13] also used an HMM-based classification method to detect malware belonging to a metamorphic family. In the first step, they used two statistical parameters “normalized term frequency” and “term frequency-inverse document frequency” to eliminate unnecessary files. They then trained HMM with n-grams of opcode sequences extracted from different viruses’ codes. For each group of viruses, a separate HMM was trained. They calculated the probabilities of observing n-gram sequences of a given code and compared them with the trained sequences.

## III. RESEARCH METHODOLOGY

A Hidden Markov model (HMM) is a machine learning technique that acts as a state machine. A Hidden Markov model has states, and known probabilities of the state transitions are termed a Markov model [14]. A Markov model has states that are visible to the observer. In contrast, a hidden Markov model (HMM) has states that are not directly observable [15].

Each state is associated with a probability distribution for viewing a set of observation symbols. The transition between the states has fixed probabilities.

In general, HMM is used for statistical pattern analysis although it is also used in speech recognition [15] and malicious code detection [15,16].

To train an HMM, we used the observation sequences to represent a set of data [14]. We then matched an observation sequence against a trained HMM to determine the probability of seeing such a sequence. If the probability is high, the observation sequence is similar to the training sequences.

As mentioned in [17], the notations used in the hidden Markov models are as follows (see Fig. 1):

- T = length of the observation sequence
- N = number of states in the model
- M = number of distinct observation symbols
- Q = distinct states of the Markov Model
- V = set of possible observations
- A = state transition probability matrix

B = observation probability matrix  
 $\pi$  = initial state distribution  
 O = observation sequence

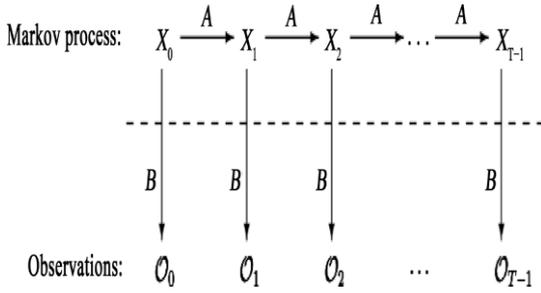


Figure 1. Hidden Markov Model Notations [17]

A hidden Markov model is defined by the matrices A, B and  $\pi$ . An HMM is thus denoted as  $\lambda = (A, B, \pi)$

The following three problems can be solved efficiently using HMM algorithms[17]:

Problem 1: Given a model  $\lambda = (A, B, \pi)$  and an observation sequence O, we need to find  $P(O|\lambda)$ , an observation sequence that can be scored to see how well it fits a given model

Problem 2: Given a model  $\lambda = (A, B, \pi)$  and an observation sequence O, we can determine an optimal state sequence for the Markov model. This is the most likely hidden state sequence that can be uncovered.

Problem 3: Given O, N and M, we can find a model  $\lambda$  that maximizes the probability of O. This involves training a model to optimally fit an observation sequence.

These three problems can be efficiently solved using the following three algorithms:

- The Forward algorithm.
- The Backward algorithm.
- The Baum-Welch re-estimation algorithm.

The forward algorithm is used to calculate the probability of being in a state  $q_i$  at time t given an observation sequence O [17]. The forward algorithm, or  $\alpha$  pass,

determines  $P(O|\lambda)$ . The algorithm can be stated as follows.

For  $t = 0, 1, \dots, T-1$  and  $i = 0, 1, \dots, N-1$ .

$$\alpha_t(i) = P(O_0, O_1, \dots, O_t, x_t = q_i | \lambda)$$

The probability of the partial observation sequence up until time t is  $\alpha_t(i)$ . Using the forward algorithm,  $P(O|\lambda)$  can be computed as follows:

Let  $\alpha_0(i) = \pi_i b_i(O_0)$ , for  $i = 0, 1, \dots, N-1$

For  $t = 1, 2, \dots, T-1$  and  $i = 0, 1, \dots, N-1$  compute

$$\alpha_t(i) = \sum_{j=1}^{n-1} \alpha_{t-1}(j) a_{ij} b_i(O_t)$$

$$\text{Then, } P(O|\lambda) = \sum_{j=0}^{n-1} \alpha_{t-1}(j)$$

The backward algorithm helps determine the most likely optimal state sequence. This algorithm can be stated as follows [17]:

For  $t = 0, 1, \dots, T-1$  and  $i = 0, 1, \dots, N-1$  define

$$\beta_t(i) = P(O_{t+1}, O_{t+2}, \dots, O_{T-1}, x_t = q_i | \lambda)$$

$\beta_t(i)$  can then be calculated using the following steps:

Let  $\beta_{T-1}(i) = 1$ , for  $i = 0, 1, \dots, N-1$

For  $t = T-2, T-3, \dots, 0$  and  $i = 0, 1, \dots, N-1$ , compute:

$$\beta_t(i) = \sum_{j=1}^{n-1} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$$

For  $t = 0, 1, \dots, T-2$  and  $i = 0, 1, \dots, N-1$ , define:

$$\gamma_t(i) = P(x_t = q_i | O, X)$$

The relevant probability up until time t is given by:

$$\gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{P(O|X)}$$

The most likely state at any time t is the state at which  $\gamma_t(i)$  is at a maximum.

The Baum-Welch algorithm helps to iteratively re-estimate the parameters A, B, and  $\pi$  [21]. It provides an efficient way to best fit the observations. The number of states N and number of unique observation symbols M are constant. However, other parameters such as A, B and  $\pi$  change in line with the row stochastic condition [17]. This process of re-estimating the model can be explained as follows [17]:

Initialize  $\lambda = (A, B, \pi)$  with an appropriate guess or random values. For example:

$$\pi = 1/N, A_{ij} = 1/N, B_{ij} = 1/M.$$

Compute  $\alpha_t(i)$ ,  $\beta_t(i)$ ,  $\gamma_t(i)$  and  $\gamma_t(i, j)$  where  $\gamma_t(i, j)$  is a digamma. The digammas can be defined as:

$$\gamma_t(i) = \alpha_t a_{ij} b_j \left( \frac{O_{t+1} \beta_{t+1}(j)}{P(O|\lambda)} \right)$$

$\gamma_t(i)$  and  $\gamma_t(i, j)$  are related by:

$$\gamma_t(i) = (x + a)^n = \sum_{j=0}^{n-1} \gamma_t(i, j)$$

Re-estimate model parameters as follows:

For  $i = 0, 1, \dots, N-1$ , let:

$$\pi_i = \gamma_0(i)$$

For  $i = 0, 1, \dots, N-1$  and  $j = 0, 1, \dots, N-1$ , compute:

$$a_{ij} = \frac{\sum_{t=0}^{T-2} \gamma_t(i, j)}{\sum_{t=0}^{T-2} \gamma_t(i)}$$

For  $j = 0, 1, \dots, N-1$  and  $k = 0, 1, \dots, M-1$ , compute:

$$b_j(k) = \sum_{\substack{t \in \{0,1,\dots,T-2\} \\ O_t=k}} \gamma_t(j) / \sum_{t=0}^{T-2} \gamma_t(i)$$

If  $P(O|\lambda)$  increases, proceed to step 3.

A. Research method

The datasets are initially prepared for the HMM process. Once the HMM accesses the dataset records, it starts to build learning objects for each record in the dataset. Each learning object will be related to a class that possesses an average probability from the contained objects (each object contains information about its learned data). The probability for each object derives from the learning chain equations in the HMM, which tries to connect objects to one another. Learning classes are connected to each other at the top level, while at lower levels a learning object from a class may be connected to learning objects in other classes.

In our research, we trained an HMM using the observation sequences (API call sequences and Opcodes sequences) to represent a set of data. We matched an observation sequence against a trained HMM to determine the probability of observing such a sequence (score). If the probability (score) was high, the observation sequence was similar to the training sequences; if it was not high, the sequences were different. The HMM learning process is illustrated in the following chart:

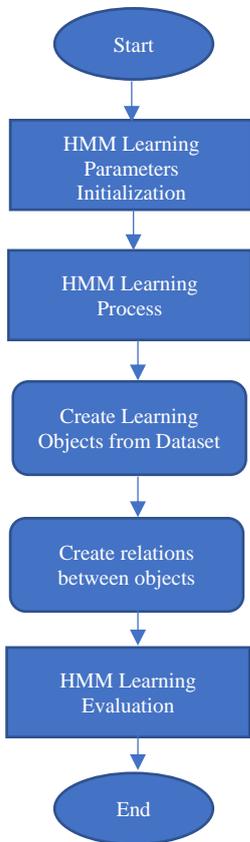


Figure 2. Flowchart for HMM learning process

As Fig. 2 shows, the datasets are initially dealt with by preparing them for the HMM process. Once the HMM accesses the dataset records, it will start to build learning objects for each record in the dataset. Each object will be related to a chain that possesses within itself an average probability from the contained objects (each object contains information about its learned data). The likelihood probability (score) for each object derives from the learning chain equations in the HMM, which tries to connect objects to one another. Meanwhile, learning chains are connected to each other on the top level, while on the level below a learning object from one chain might be connected to its counterparts in other chains.

Upon completion of the HMM process, the code extracts a learning model based on its accuracy.

B. Dataset extraction and preparation

The structure of the adopted dataset aims to reflect malware behaviour in two ways; the first is based on process Opcode sequences [15], while the second is based on API call sequences [18]. The authors of [18] dynamically analysed approximately 23,000 instances of malware and extracted the API call sequences, as well as obtaining the Opcodes sequences from [15]. Conversely, the authors of [15] statistically analysed 8,000 malware programs to extract Opcode sequences (more information about each dataset can be found in the references for this research) [15][18]. In our project, we chose common malware in both datasets (6,900 malware items) to equalize the size of datasets between the experiment based on API call sequences and the experiment based on Opcode sequences, and then used these in our experiment. Each dataset contained several records so that we could train our model accordingly.

C. Experiment Environment

In our project, we used MATLAB R2013a for programming. MATLAB has the ability to operationalise complex mathematical expressions. Researchers have often utilised it for prototyping new machine learning algorithms. Additionally, it also reduces the effort and time required for research and development [19].

MATLAB also has many toolboxes for machine learning models and algorithms, such as the genetic algorithm and K-means classifier. It is also useful for representing and working with matrices [19].

In our experiment, we used 6,900 malware programs to build an HMM model; 70% were used to learn the model and 30% to test it.

IV. RESULTS AND DISCUSSION

To present an example for the HMM learning log value called score, we show the learning behavior for the first six learning objects in each dataset. These are presented in Fig. 8 and Fig. 9. To simplify the learning process for clarity, each line displays the learning behavior from time 0 until the top level of learning is reached. The increasing level of the line means that this object is still learning and constructing a probability value for objects or called score; however, once this line becomes

smooth, it means that the learning process has stopped and there is no learning score above this line.

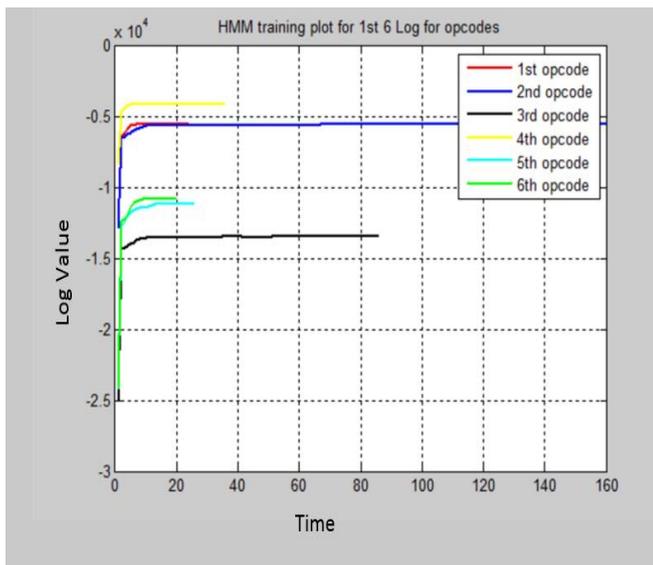


Figure 3. First six learning objects’ log score behaviour in HMM for the opcode dataset

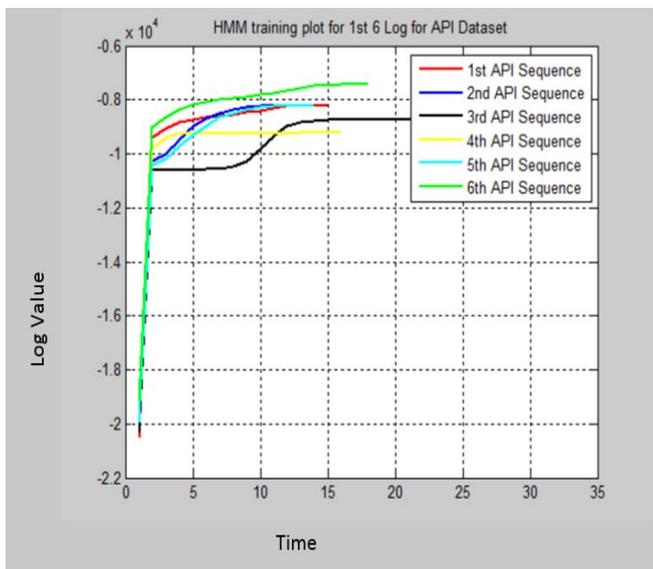


Figure 4. First six learning objects’ log score behaviour in HMM for the API dataset

In Fig. 5, the accuracy of the learning process for the API dataset and the Opcode sequences dataset is compared. In HMM, the accuracy of the learning process is represented by a log value which, according to HMM documentation, represents the value of observation probability (the same as the accuracy of the chains collaboration). After a training model is built using the Baum Welch algorithm, each object in the learning model has a probability value. The probability for each set of objects (scores) stems from the learning chains equations in HMM.

The learning process involves trying to connect objects to one another. To compare both datasets, we used the average value of probability values or scores for all sequences in each dataset. Each one of these datasets was trained within the HMM to provide a learning model. Hence, this learning model will be used to evaluate whether a software program is a form of malware. Optimal accuracy in the results was based on the average accuracy value of the learning objects during the learning process. Thus, the training accuracy in the first test for the Opcode dataset was 1.4556; for the API dataset, the accuracy was 1.5365, higher than that of the Opcode dataset.

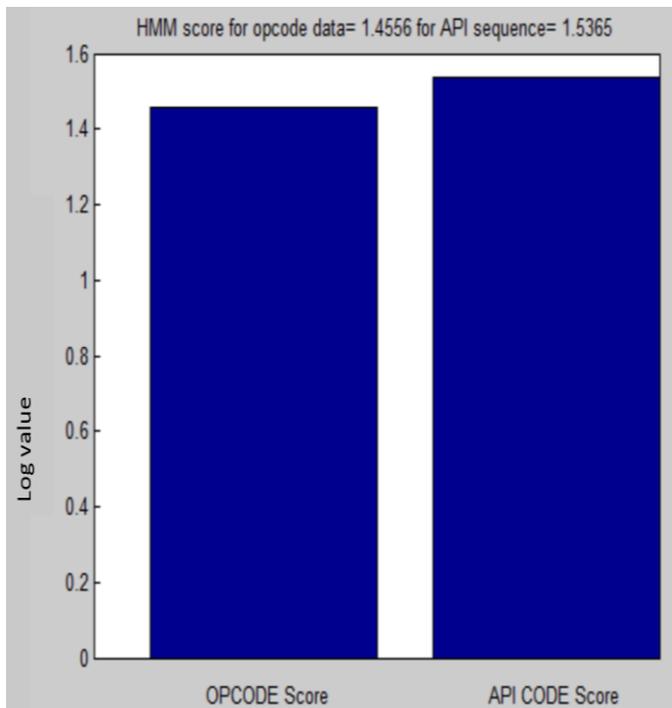


Figure 5. Accuracy of HMM training for API and Opcode data

V. EVALUATION

The evaluation process compares the learning accuracy of the HMM for both the API and opcode datasets. Ten tests were carried out to determine the average accuracy. The accuracy values are presented in the following table.

TABLE 1: HMM LEARNING ACCURACY FOR API AND OPCODE DATASETS

#test	API	Opcode
1	1.3	1.1
2	1.4	1.49
3	1.1	1.4
4	1.3	1.2

#test	API	Opcode
5	1.4	1.2
6	1.2	1.24
7	0.9	1.05
8	1.48	0.97
9	1.3	0.8
10	1.4	1.2

As shown in Table 1, in the first phase of evaluation the API score was 1.278 out of 1.5 (learning rate; we simplified the absolute error to obtain a value from 0 to 1.5 for each learning iteration); the opcode score, on the other hand, was 1.165 from 1.5. Thus, API yielded the best results in the HMM learning process due to the composite structure of the opcode data records.

Smoothing of the records in the API dataset, rather than the opcodes dataset, occurred because the structure of API call sequences is readable compared to the Opcode sequences dataset records. This led to the superior result in the API dataset. Therefore, the API call sequence reflects the behavior of malware better than Opcode, and contains better sequence representation.

## VI. CONCLUSION

Malware has become a major threat to computers and information systems. In response, malware developers use obfuscation techniques to generate malware with the ability to escape anti-malware systems. Recent research has focused on how to detect malware based on its behavior. In this work, we studied API call sequences and opcode sequences, known to reflect malware behavior, and compared their ability to detect malware based on a hidden Markov learning model. The results from this work show that, based on the HMM, API call sequences are more accurate for detecting malware. In future work, we intend to classify malware based on hidden Markov learning using machine-learning algorithms.

## REFERENCES

- [1] M. Alazab, S. Venkataraman, and P. Watters, "Towards understanding malware behaviour by the extraction of API calls," in Proc. - 2nd Cybercrime Trust. Computer Work. CTC 2010, November 2009, pp. 52–59, 2010.
- [2] G. McGraw and G. Morrisett, "Attacking malicious code: A report to the Infosec Research Council," IEEE Software, vol. 17, no. 5, pp. 33–41, 2000.
- [3] C. Torrano-gimenez, A. Perez-Villegas, G. Alvarez, C. Torrano-Giménez, A. Perez-Villegas, G. Álvarez Marañón, and others, "An Anomaly-Based Approach for Intrusion Detection in Web Traffic," Digital. Journal Information. Assurance and Security 5(4) 446-454, pp. 1–9, 2010.
- [4] U. Bayer, A. Moser, C. Kruegel, and E. Kirda, "Dynamic analysis of malicious code," Journal Computer Virology, vol. 2, no. 1, pp. 67–77, 2006.

- [5] Z. Bazrafshan, H. Hashemi, S. M. H. Fard, and A. Hamzeh, "A survey on heuristic malware detection techniques," IKT 2013 - 2013 5th Conference on Information and Knowledge Technology, pp. 113–120, May 2013.
- [6] K. Iwamoto and K. Wasaki, "Malware classification based on extracted API sequences using static analysis," in Proceedings of the Asian Internet Engineering Conference on - AINTEC '12, 2012, no. June, pp. 31–38.
- [7] D. Uppal, R. Sinha, V. Mehra, and V. Jain, "Malware Detection and Classification Based on Extraction of API Sequences," in Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on IEEE), 2014, pp. 2337–2342.
- [8] M. Alazab, "Profiling and classifying the behavior of malicious codes," Journal of Systems and Software, vol. 100, pp. 91–102, 2015.
- [9] R. Moskovitch, C. Feher, N. Tzachar, E. Berger, M. Gitelman, S. Dolev, and Y. Elovici, "Unknown malcode detection using OPCODE representation," in Intelligence and Security Informatics, 2008, pp. 204–215.
- [10] I. Santos, F. Brezo, J. Nieves, Y. K. Peña, B. Sanz, C. Laorden, and P. G. Bringas, "Idea: Opcode-sequence-based malware detection," in International Symposium on Engineering Secure Software and Systems, 2010, vol. 5965 LNCS, pp. 35–43.
- [11] A. Jantan, "An approach for malware behavior identification and classification," in 2011 3rd International Conference on Computer Research and Development, 2011, vol. 1, pp. 191–194.
- [12] S. Priyadarshi, "Metamorphic Detection via Emulation Metamorphic Detection via Emulation," MSc thesis, San Jose State University, 2011.
- [13] S. P. Thunga and R. K. Neeliseti, "Identifying Metamorphic Virus Using n-grams And Hidden Markov Model," in Advances in Computing, Communications and Informatics (ICACCI), 2016, pp. 2016–2022.
- [14] C. Annachhatre, "Hidden Markov Models for Malware Classification," MSc thesis, San Jose State University, 2013.
- [15] C. Annachhatre, T. H. Austin, and M. Stamp, "Hidden Markov models for malware classification," Journal Computer Virology and Hacking Techniques, vol. 11, no. 2, pp. 59–73, May 2015.
- [16] A. Kalbhor, T. H. Austin, E. Filiol, S. Josse, and M. Stamp, "Dueling hidden Markov models for virus analysis," Journal Computer Virology and Hacking Techniques, vol. 11, no. 2 pp.103-118, November 2014.
- [17] M. Stamp, "A revealing introduction to hidden Markov models," Department of Computer Science San Jose State University, 2004.
- [18] Y. Ki, E. Kim, and H. K. Kim, "A Novel Approach to Detect Malware Based on API Call Sequence Analysis," International Journal of Distributed Sensor Networks, vol.11 Issue 6, 2015.
- [19] MATLAB, "Machine Learning with MATLAB." [Online]. Available: <https://www.mathworks.com/products/matlab.html>. Accessed: 12-Oct-2016].

# Android Botnet Detection

## An Integrated Source Code Mining Approach

Basil Alothman

Faculty of Technology  
De Montfort University  
Leicester LE1 9BH, Great Britain  
p14029266@my365.dmu.ac.uk

Prapa Rattadilok

Faculty of Technology  
De Montfort University  
Leicester LE1 9BH, Great Britain

**Abstract**—Android is one of the most popular smartphone operating systems. This makes it one of the default targets for malicious cyber-attacks. Android’s Play Store is not very restrictive which makes installing malicious apps easy. Botnets are amongst the most dangerous hacking approaches that are used nowadays on the internet. It is common for botnet developers to target smartphone users to install their malicious tools and target a larger number of devices. This is often done to gain access to sensitive data such as credit card details, or to cause damage to individual hosts or organisation resources by executing denial of service attacks. In this paper, we propose an approach to identify botnet Android mobile apps by means of source code mining. We analyse the source code via reverse engineering and data mining techniques for several examples of malicious and non-malicious apps. We use two approaches to build datasets. In the first, we perform text mining on the source code and construct several datasets and in the second we build one dataset by extracting source code metrics using an open-source tool. After building the datasets, we run several classification algorithms and assess their performance. Initial results show a high level of accuracy.

**Keywords**—component; Android Botnet Detection; Document Analysis; Source Code Mining; Android App; Reverse Engineering; Malware Detection

### I. INTRODUCTION

Android is a very popular operating system that is now growing amongst smartphone users. This operating system is open source and user-friendly. Also, it is relatively easy to write Java applications that run smoothly on it. With the huge number of existing Java apps, it becomes difficult to know whether a new app is safe to use. To overcome this issue, amongst other challenges, it would be useful to automate the process of checking how safe a new app is (i.e. to use the smartphone itself to predict whether or not an app is safe).

In this paper, we attempt to solve this problem by automatically reverse engineering Android apps, obtaining their Java source code and using the source code to make predictions. We achieve this by using data mining techniques to analyse the Java source code of these apps and try to predict whether a given app is botnet or not. Observe that these apps come as Android Application Package (APK) files.

We have gathered a collection of botnet apps (obtained from the ISCX data-set [3], [4]) and safe, or normal, apps. We have used Dex2jar[10] to reverse engineer these android apps and

convert them into Java source code. We had a collection of 21 apps (9 botnets and 12 safe) and used them to create datasets suitable for data mining. In our datasets, each app is an instance (i.e. example). We extracted several attributes (i.e. features) for each app from its Java source code, and our class variable is either botnet (positive) or not (negative). After building several datasets, we used them to run several classifiers and evaluate their performance.

To the best of our knowledge, we have not seen any work to identify Android botnet apps by directly mining their source code. Therefore, our contributions can be summarised as follows: Our approach uses data mining techniques to analyse the Java source code in two ways. In the first method, we treat the Java source code as if it was normal text by using Natural Language Processing (NLP) methods[17]. And in the second approach, we extract several statistical measures from the source code and use these metrics as attributes. Our approach can be considered static as we do not require the execution of the Android app itself. Our idea is that as soon as an Android app is downloaded, it is reverse engineered and its Java source code is obtained and used to predict whether this app is safe or malicious. It is noteworthy that the advantage here is being “pro-active”. In other words, we attempt to identify danger before it occurs.

Another point we would like to state is that our work can be considered behaviour-based instead of signature-based. We believe that using signature-based methods have the general disadvantage of relying on other people to report whether a certain app is malicious (signature-based methods work by comparing signatures, or hashes, of files or file contents on a system to a list of known malicious files).

The remainder of this paper is organised as follows: Section 2 summarises related work and existing approaches. Section 3 provides a brief overview of botnets and Section 4 explains in detail how we constructed our datasets. The 5th Section has a short description of the algorithms we have used. Sections 6 and 7 have our experimental results and discussion and conclusions respectively. The paper finishes with our planned future work in Section 8.

### II. RELATED WORK

Many approaches for botnet detection have been reported in the literature. Zhao et al. [23] introduced an approach that is

based on the analysis of network traffic. They extract a set of features (i.e. attributes) from traffic chunks and then use machine learning algorithms to identify whether the traffic is malicious or not. Other works include BotMiner [20] and BotHunter [21].

As our work is focused on botnet detection on the Android operating system, we will summarise the key existing approaches in the remainder of this section.

An interesting approach is DENDROID [6]. In this work, malware Android apps were grouped into families by analysing their source code. A similarity measure was used to taxonomise apps and create a phylogenetic-tree like structure.

One approach to detect malicious Android apps was the work of Sheen et al. [2]. In this work, features such as the API calls and permission requests that an APK file makes are used in separate datasets and an ensemble of classifiers (collaborative decision fusion) was used to perform predictions. Another approach that is related to permissions is that of Wei et al. [7]. In their work, they use techniques from the text mining domain to analyse the relationship between permission requests that an Android app makes and its textual description. Singh et al. [5] stated that the use of permission mechanism is not effective in general.

Several approaches use static or dynamic analysis techniques. In static analysis, attempts are made to detect malicious activities without the need to execute the Android apps. The main idea is to model how the Android apps work by constructing and analyse some graphical models. An example of this type is the recent work of Junaid et al. [1]. They presented an approach to detect malicious behaviour in Android apps using models of their life cycles. Reverse engineering was used in this approach to construct a life cycle model for each Android app. After that, possible event sequences are derived from these models and used in attack detection. They developed a system called Dexteroid to identify SMS (Short Message Service) when they are sent to costly numbers as well as whether sensitive data is being leaked. Other examples include the work carried by Gordon et al. [22] who built a tool called DroidSafe, Feng et al. [27], Arzt et al. [24] who built a tool called FlowDroid and Yang and Yang [25] who developed a tool called LeakMiner. On the other hand, approaches that employ dynamic analysis try to execute the Android apps to perform specific tasks and use the resulting data to detect malicious attacks. Some recent examples are the works of Bai et al. [26], Zhang et al. [27] and Yan and Yin [28].

Singh et al. [5] proposed a method that is based on the analysis of manifest files to identify malicious behaviour. Additionally, Chen et al. [29] proposed an approach to build data flow models from the reverse engineered source code of Android apps. Their method tries to build data flow models by detecting where data enters an application and how this data moves through it. In other words, they build trees of classes, methods and variables and use these trees to identify malicious code.

### III. BOTNET OVERVIEW

Botnets are groups of networked computers which carry out certain malicious activities such as stealing confidential information such as login credentials or credit card details, send

spam emails or perform DDoS attacks [13]. These computers are compromised and controlled remotely by a botmaster [11].

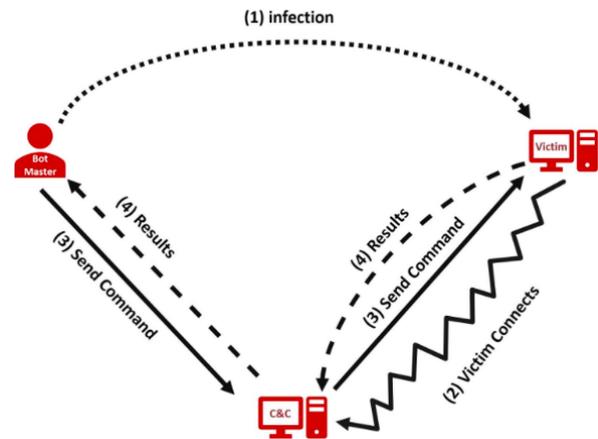


Figure I: Anatomy of a Bot Attack

Usually, there is a botmaster and botnet clients (i.e. victims). The botmaster is the attacker, or ontroller, of the botnet [11] (most likely the developer of the botnet malware). A bot client, or host, is the victim host which is exploited and remotely controlled by the botmaster. The botmaster is often known as the server Command & Control (C&C) as it sends commands and instructions to the botnet clients to carry out malicious tasks.

Botnets follow a systematic way to launch attacks. As we show in Fig. I, the first step is for the Botmaster to infect a victim with a Bot. As we have mentioned previously, there are several ways to infect a victim. Observe that the number of infected victims can be very large and, therefore, the attacker can have an army of Bots under his, or her, control. After the victims are infected, they connect to the C&C server and wait for instructions. This connection can be established using one of the known protocols such as HTTP or IRC. Then, the C&C server sends its commands to the victims which in turn execute the commands and report back the results to the C&C.

### IV. DATASET FORMATION AND FEATURE EXTRACTION

To obtain the Java source code of the Android apps, we reverse engineered them. As the APK files are compressed files, we renamed them to .zip and then unzipped them which resulted in .dex files [8]. After that, we used the Dex2jar [10] tool to convert the “dex” files into Java “jar” files. As now we have Java jar files, we used the Java decompiler “JD-GUI” [9] to regenerate the Java source code of the APK apps. The entire process is illustrated in Fig. II. We made sure that the normal (or not botnet) apps were network applications. We provide a list of the apps we have used in Table I.

TABLE I. A LIST OF THE APPS WE HAVE USED

Botnet Apps	Not Botnet Apps
Anserverbot	AndroIRC
Bmaster	SimpleIRC
DroidDream	Kik
Geinimi	LOVOO
Nickspy	Line
PJapps	WhatsApp
Pletor	Hi5

Zitmo	SKOUT
Rootsmart	Viber
	Messenger
	WeChat
	SnapChat

To be able to predict whether a given app is BotNet or Not, we need to build predictive models. For this purpose, we used the WEKA [31] open source machine learning platform (version 3.6.13). We prepared our datasets as follows:

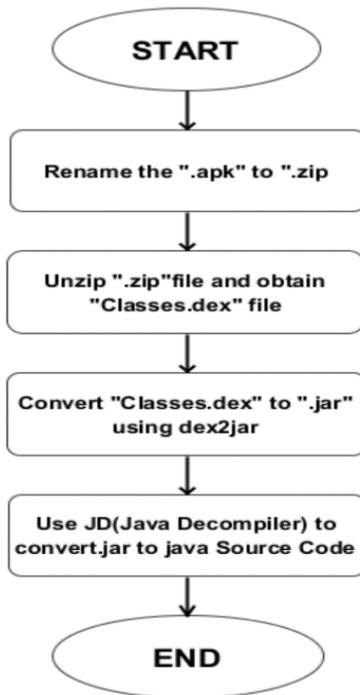


Figure II. APK File Reverse Engineering

A. Text-Mining Approach

Text mining [5] aims to process textual information, which is normally unstructured, and use the resulting structured data to build predictive models and, or, to understand the original textual information better. The structured data is usually obtained by deriving numerical summaries about the documents based on the words they contain.

To be able to use the Java source code we obtained to perform text mining, we concatenated all the Java code of each app into one file (which means we now have one Java source code file per app) and then we created a dataset that has three columns: The app’s name, the apps Java source code and the app’s class (botnet or not).

After that, we applied WEKA’s TextToWordVector filter with Term Frequency and Inverse Document Frequency (TF-IDF) on all the Java code. TF-IDF [17] is a widely used transformation in NLP where terms (or words) in a document are given importance scores based on the frequency of their appearance across documents. We use this idea so that any information, such as Java class, method and variable names, or words used in comments, are assigned scores. A word is important and is assigned a high score if it appears multiple times

in a document (i.e. Java source code of an app). However, it is assigned a low score (meaning it is less important) if it appears in several documents (Java code of several apps).

We used WEKA’s default parameters for this filter except for the “number of words to keep”. This parameter is 1000 by default, and we changed it to 3000 and 5000. This is because some of the apps had a large number of Java classes and lines. An example dataset resulting after this process is shown in table II. Observe that the features are Weights (w’s) of words that result after applying the TF-IDF filter. And the dots “...” mean so forth.

TABLE II. DATASET RESULTING AFTER APPLYING TEXTTOWORDVECTOR AND THEN TF-IDF FILTERS

App Name	W1	W2	...	Class (Botnet or not)
App 1	0.069	0.034	...	Yes
App 2	0	0.018	...	No
...	...	...	...	...
App n	0.009	0	...	No

B. Source Code Metrics Approach

In this approach, we aimed to use software metrics as characteristics (or features) of the Java source code we obtained. For this purpose, we used the tool CodeAnalyzer [30] and we were able to obtain several quantitative measures. These include statistics such as the total number of files, the total number of code lines and the code to comment ratio. An example dataset resulting after this process is shown in table III. Observe that SCM stands for Source Code Metric.

TABLE III. DATASET RESULTING AFTER APPLYING TEXTTOWORDVECTOR WITH TF-IDF FILTER

App Name	SCM1	SCM2	SCM m	Class (Botnet or not)
App 1	33921	0.11	...	Yes
App 2	128998	0.21	...	No
...	...	...	...	...
App n	45635	0.33	...	No

C. Feature Selection

We applied WEKA’s StringToWordVector with TF-IDF filter with a various number of words to keep, the resulting datasets had much more features than examples. For example, the number of features in dataset W3000 is 4332 (see table IV) and the number of examples we have is 21. This means that, in this dataset, the number of features is more than 200 times the number of examples. Having a large number of features makes it practically impossible to interpret models and can cause overfitting [19]. Therefore, reducing the number of features can help avoid overfitting and build models which are easier to interpret and with better predictive performance. Having a smaller number of features can also reduce the computational time considerably [18].

As feature selection tries to identify the most informative features and removes the uninformative, irrelevant, noisy or unreliable features, we applied WEKA’s SubSetEval feature selection algorithm on each of these datasets.

The selected features included words such as “lock”, “state” and “concurrent” and the removed features included words like “audio”, “recycle” and “widget”. Table IV provides a description of all our datasets.

TABLE IV. A SUMMARY OF THE DATASETS WE HAVE CREATED

Dataset Name	No of Features	Dataset Description
Metrics	15	Resulted after extracting code metrics using CodeAnalyzer
W1000	1332	Resulted after applying WEKA’s StringToWordVector with number of words to keep = 1000
W1000FS	24	Resulted after applying WEKA’s SubSetEval feature selection algorithm to dataset W1000
W3000	4332	Resulted after applying WEKA’s StringToWordVector with number of words to keep = 3000
W3000FS	85	Resulted after applying WEKA’s SubSetEval feature selection algorithm to dataset W3000
W5000	7697	Resulted after applying WEKA’s StringToWordVector with number of words to keep = 5000
W5000FS	21	Resulted after applying WEKA’s SubSetEval feature selection algorithm to dataset W5000

### V. ALGORITHMS USED IN THIS STUDY

In this study, we have used the following algorithms in WEKA:

#### A. NaiveBayes

The Naive Bayes classifier [14] is based on Bayes theorem with independence assumptions between input variables (predictors). Suppose  $x$  was the input variables and  $c$  was the class, Bayes theorem introduces a method of calculating the posterior probability,  $P(c|x)$ , from  $P(c)$ ,  $P(x)$ , and  $P(x|c)$ . This classifier assumes that the effect of the value of an input variable ( $x$ ) on a given class ( $c$ ) is independent of the values of other input variables. This assumption is known as class conditional independence.

#### B. KNN

K-Nearest Neighbours (KNN) [15] is an algorithm that stores all available examples (i.e. instances) and classifies new examples based on a similarity measure (e.g. distance function). In more detail, a majority vote of neighbours is used to classify new examples. This is achieved by the choosing the most common class amongst the K-Nearest neighbours. In our study we used five neighbours ( $K=5$ ).

#### C. J48

J48 is a variation of the well-known C4.5 algorithm which is decision tree [12]. The way it works is by building classification or regression models in the form of a tree structure. This is done by breaking down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes. A decision node has two or more branches, and a Leaf node represents a classification or decision.

#### D. RandomForest

This algorithm works by building many decision trees at training time and using them to vote for the class of a new

example [16]. To construct each tree, the training data is obtained by randomly sampling both the examples and input variables (with replacement).

#### E. SMO

This is WEKA’s Support Vector Machine (SVM) [13]. This algorithm carries out classification by finding the hyperplane that finds the maximum possible margin between two classes.

### VI. EXPERIMENTAL RESULTS

We have introduced the algorithms we used in Section V. We have run each algorithm on each of our datasets. We have used 10 fold cross-validation to calculate the average classification accuracy. It is worth mentioning here that we ran these algorithms with their default parameters unless mentioned otherwise. Our results are summarised in Table V.

TABLE V. A SUMMARY OF OUR RESULTS

Dataset	NB	KNN(5)	J48	RF	SMO
Metrics	<b>85.71%</b>	80.95%	80.95%	<b>85.71%</b>	80.95%
W1000	71.43 %	76.19 %	80.95 %	76.19 %	<b>90.47 %</b>
W1000FS	80.95%	<b>95.24 %</b>	80.95 %	90.47 %	90.48 %
W3000	85.71 %	<b>95.24 %</b>	66.67 %	80.95 %	<b>95.24 %</b>
W3000FS	90.48 %	<b>95.24 %</b>	80.95 %	<b>95.24 %</b>	95.23 %
W5000	76.19 %	<b>95.23 %</b>	85.71 %	85.71 %	90.48 %
W5000FS	90.47 %	<b>95.24 %</b>	95.23 %	<b>95.24 %</b>	<b>95.24 %</b>

### VII. DISCUSSION AND CONCLUSIONS

For each dataset, we display the best performing algorithm in bold. As it can be seen, some algorithms perform equally for some datasets. The feature values are positive real numbers in all of these datasets. For the metrics dataset, the number of features is less than the number of examples. As for the other datasets, some of them had much more features than examples as can be calculated using the *No of Features* column in table IV.

We have talked about the disadvantage of having much more features than examples in Section 4 and our experimental results support what we discussed there. It is clear from the table that the performance improves significantly after applying feature selection. This case is true regardless of the number of features before applying feature selection (recall we used several values for the number of words to keep parameter when we applied TextToWord filter).

By analysing the experimental results further, we can make an interesting observation. That is, the performance of the Decision Tree (J48) algorithm is always the worst regardless of the dataset.

Another observation is that the k-Nearest Neighbour algorithm (with five neighbours in our case) seems to be the best performing algorithm in general.

Some might suggest that code obfuscation might be used to evade detection. We would like to clarify that this is exactly what our method is about. Even if a known Botnet app uses such techniques, we just need to add it to our training data and use it in building predictive models.

## VIII. FUTURE WORK

A method for the automatic detection of botnet apps was introduced in this paper. This method is based on the analysis of the Java source code of such apps. It is worth mentioning here that our current dataset is small and it more data is needed to draw more sound conclusions. It is costly to obtain data and we currently only have 21 examples (9 known botnets and 12 normal apps). However, we plan to extend our dataset soon and extend the approach in the near future. Additionally, we intend to use more data mining algorithms and techniques. For example, we want to use ensemble techniques such as bagging and stacking. Another idea we would like to explore is to merge the two types of datasets we have created (i.e. the metrics and text mining data). This is illustrated in Table VI (where w's are weights resulting after TF-IDF and M's are metrics as we explained in previous sections). After that, we will investigate whether we gain any performance improvement.

TABLE VI. AN EXAMPLE DATASET RESULTING AFTER MERGING TEXT MINING AND METRICS DATASETS

App Name	W1	W2	...	M1	M2	...	Class
App 1	1	0	...	33921	0.11	...	Yes
App 2	0	0	...	128998	0.21	...	No
...	...	...	...	...	...	...	...
App n	0	1	...	45635	0.33	...	No

## REFERENCES

- [1] JUNAID, Mohsin, DONGGANG LIU, y DAVID KUNG, . "Dextroid: Detecting malicious behaviours in Android apps using reverse-engineered life cycle models". *Computers & Security* . 2016, vol 59, p. 92 - 117.
- [2] SHEEN, Shina, R. ANITHA, y V. NATARAJAN, . "Android based malware detection using a multifeature collaborative decision fusion approach ". *Neurocomputing* . 2015, vol 151, Part 2, p. 905 - 912.
- [3] Andi Fitriah A.Kadir, Natalia Stakhanova, Ali A. Ghorbani, "Android Botnet: What URLs are telling us" , 9th International Conference on Network and System Security (NSS), November 3-5, 2015, New York City, USA
- [4] Gonzalez, Hugo, Natalia Stakhanova, and A. Ghorbani. "Droidkin: Lightweight detection of android apps similarity." *Proceedings of the 10th SECURECOMM (2014)*.
- [5] WITTE, R., Q. LI, , Y. ZHANG, y J. RILLING, . "Text mining and software engineering: an integrated source code and document analysis approach". *IET Software*. 2008, vol 2, núm. 1, p. 3-16. SINGH, Pooja, PANKAJ TIWARI, y SANTOSH SINGH, . "Analysis of Malicious Behavior of Android Apps ". *Procedia Computer Science* . 2016, vol 79, p. 215 - 220.
- [6] SUAREZ-TANGIL, Guillermo, E. TAPIADOR, Juan, PEDRO PERIS-LOPEZ, y JORGE BLASCO, . "Dendroid: A text mining approach to analyzing and classifying code structures in Android malware families ". *Expert Systems with Applications* . 2014, vol 41, núm. 4, Part 1, p. 1104 - 1117.
- [7] WEI, M., X. GONG, y W. WANG, . Claim What You Need: A Text-Mining Approach on Android Permission Request Authorization. 2015. p. 1-6.
- [8] ZHANG, Xiaolu, FRANK BREITINGER, y IBRAHIM BAGGILI, . "Rapid Android Parser for Investigating DEX files (RAPID) ". *Digital Investigation* . 2016, vol 17, p. 28 - 39.
- [9] GitHub - java-decompiler/jd-gui: A standalone Java Decompiler GUI. [ONLINE] Available at: <https://github.com/java-decompiler/jd-gui>. [Accessed 13 Nov 2017].
- [10] SourceForge. 2016. dex2jar download | SourceForge.net. [ONLINE] Available at: <https://sourceforge.net/projects/dex2jar/>. [Accessed 22 Oct 2017].
- [11] Mehedy Masud, 2011. *Data Mining Tools for Malware Detection*. 1 Edition. Auerbach Publications.
- [12] Lior Rokach, 2014. *Data Mining With Decision Trees : Theory and Applications (2nd Edition) (Series in Machine Perception and Artificial Intelligence) (Series in Machine Perception and Artificial Intelligence)*. 2 Edition. World Scientific Publishing Company.
- [13] Ingo Steinwart, 2008. *Support Vector Machines (Information Science and Statistics)*. 2008 Edition. Springer.
- [14] WIKIPEDIA, S.. *Classification Algorithms: Artificial Neural Network, Naive Bayes Classifier, Support Vector MacHine, Boosting, Linear Classifier, Case-Based Reasonin*. University-Press Org, 2013.
- [15] Larose, Daniel T. , 2014. *Discovering Knowledge in Data: An Introduction to Data Mining*. 2nd ed.: WILEY.
- [16] Sumeet Dua, 2011. *Data Mining and Machine Learning in Cybersecurity*. 1 Edition. Auerbach Publications.
- [17] Sholom M. Weiss, 2009. *Text Mining: Predictive Methods for Analyzing Unstructured Information*. Softcover reprint of hardcover 1st ed. 2005 Edition. Springer.
- [18] Huan Liu, 1998. *Feature Selection for Knowledge Discovery and Data Mining (The Springer International Series in Engineering and Computer Science)*. 1998 Edition. Springer.
- [19] Huan Liu, Hiroshi Motoda, 2007. *Computational Methods of Feature Selection (Chapman & Hall/CRC Data Mining and Knowledge Discovery Series)*. Edition. Chapman and Hall/CRC.
- [20] Guofei Gu , Roberto Perdisci , Junjie Zhang , Wenke Lee, BotMiner: clustering analysis of network traffic for protocol- and structure-independent botnet detection, *Proceedings of the 17th conference on Security symposium*, p.139-154, July 28-August 01, 2008, San Jose, CA
- [21] Guofei Gu , Phillip Porras , Vinod Yegneswaran , Martin Fong , Wenke Lee, BotHunter: detecting malware infection through IDS-driven dialog correlation, *Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium*, p.1-16, August 06-10, 2007, Boston, MA
- [22] MI Gordon, D Kim, JH Perkins, L Gilham, N Nguyen, MC Rinard, 2015, *Information Flow Analysis of Android Applications in DroidSafe*. NDSS
- [23] Liping Feng, Hongbin Wang, Qi Han, Qingshan Zhao and Lipeng Song. *Modeling Peer-to-Peer Botnet on Scale-Free Network*. Abstract and Applied Analysis.
- [24] Steven Arzt, Siegfried Rasthofer, Christian Fritz, Eric Bodden, Alexandre Bartel, Jacques Klein, Yves Le Traon, Damien Oceau and Patrick McDaniel. 2014. *FlowDroid: Precise Context, Flow, Field, Object-sensitive and Lifecycle-aware Taint Analysis for Android Apps*.
- [25] Zhemin Yang and Min Yang. 2012. *LeakMiner: Detect Information Leakage on Android with Static Taint Analysis*. WCSE'12 Proceedings of the 2012 Third World Congress on Software Engineering
- [26] Guangdong Bai, Yongzheng Wu, Jun Sun, Jianliang Wu, Yang Liu, Qing Zhang, Jin Song Dong. 2014. *DroidPF: A Framework for Automatic Verification of Android Applications*
- [27] Zhemin Yang, Min Yang, Yuan Zhang, Guofei Gu, Peng Ning, and X. Sean Wang. 2013. *AppIntent: analyzing sensitive data transmission in android for privacy leakage detection*. *Proceedings CCS '13 Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*.
- [28] Lok Kwong Yan and Heng Yin. 2012. *DroidScope: Seamlessly Reconstructing the OS and Dalvik Semantic Views for Dynamic Android Malware Analysis*. 21st USENIX Security Symposium (USENIX Security 12).
- [29] Chia-Mei Chen, Je-Ming Lin and Gu-Hsin Lai. 2014. *Detecting Mobile Application Malicious Behaviors Based on Data Flow of Source Code*. *International Conference on Trustworthy Systems and their Applications (TSA)*, 2014.
- [30] CodeAnalyzer. <http://www.codeanalyzer.teel.ws> (Accessed on 29 JUL 2017).
- [31] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); *The WEKA Data Mining Software: An Update*; SIGKDD Explorations, Volume 11, Issu

# An Analysis of Home User Security Awareness & Education

Fayez, Alotaibi<sup>1</sup>, Nathan, Clarke<sup>1,2</sup> and Steven Furnell<sup>1,2,3</sup>

<sup>1</sup>Centre for Security, Communications and Network Research (CSCAN)  
Plymouth University, Plymouth, United Kingdom

<sup>2</sup>Security Research Institute, Edith Cowan University, Western Australia

<sup>3</sup>Center for Research in Information and Cyber Security,  
Nelson Mandela Metropolitan University, Port Elizabeth, South Africa  
Fayez.alotaibi@plymouth.ac.uk

*Abstract* - The human factor is a major consideration in securing systems. People use an increasingly wide range of digital devices such as smartphones, tablets, laptops and smart TVs, with each device having a different operating system, security configurations and threats. This presents users with an unenviable and potentially insurmountable task of securing them. This paper presents an analysis of efforts being made in providing suitable awareness and education with home users. The analysis shows that whilst significant efforts are made, a focus upon a “one-fits-all” solution that does not take into account individual users – their needs, prior knowledge, learning styles and security priorities – results in unnecessary information overload and a need to spend an excessive amount of time reading web-based content that may have little relevance to them specifically. This review indicates that there is a need for an approach that can provide the users with bespoke awareness information. It is recommended that a holistic information security management system for home users can be proposed and designed which can provide users with bespoke awareness information based on the technologies, applications and services that users use in a manner that is acceptable and timely.

## I. INTRODUCTION

The number of the internet users in the world has risen from 1 billion in 2005 to 3.17 billion in 2015, and is likely to further increase to 5 billion by 2020 [1], [2]. Many people around the world use different digital technologies to access different online services. For example, the percentage of adults who use a computer for online services has increased from 67% in 2009 to 81% in 2014 in the UK. Furthermore, the number of smartphone users has risen sharply from 28% to 61% between 2009-2014 [3].

Alongside this significant increase in the number of Internet users and different platforms, different devices and a wide range of online applications and services, a significant increase in cyber-related threats have also been experienced [4]. A recent Internet Security Threat Report published by Symantec [5], shows that 54 zero-day vulnerabilities were discovered in 2015, compared to only 23 and 24 vulnerabilities in 2013 and 2014 respectively. In addition, the number of the identities exposed in online breaches was 429 million in 2015, up 23% on the previous year. Ransomware malicious attacks also rose from 269,000 attacks in 2014 to 362,000 attacks in 2015. The number of the web attacks blocked per day increased more than the double in 2015 with 1.1 million attacks. Users not only have an increasing array of

technologies and services to protect but against a backdrop of an increasing threat landscape that requires them to take this seriously.

Information security awareness and education therefore has a key role in ensuring users are informed and educated on how to remain secure. Several studies have tried to assess the online safety and the information security behavior of home users. A study among UK adults highlighted that 55% were aware and used a firewall at home, 24% were aware but did not use firewall and 20% were not aware of firewall [3].

The National Cyber Security Alliance and McAfee conducted an online safety survey among Internet home users in the U.S. The findings of the study showed that 25% of the home users never changed their passwords without being forced by the service provider and only 15% changed them in the last year. In addition, 53% of users use an unsecured wireless network to connect to the internet and only 32% used a wireless network protected by passwords [6]. Another study was conducted to evaluate the security procedures for the smartphone users in the U.S. [7], the results indicated that only 34% used PIN for their smartphones, more than the half of the respondents (55%) did not configure the PIN code to protect their devices. Less than one quarter of the respondents (21.6%) installed security software such as antivirus, compared with 41% of participants installing gaming applications.

The aforementioned studies analysed the situation of the information security awareness in USA and the UK which are considered as developed countries. The lack of cyber awareness is highly likely to increase among the developing nations. A study was conducted to evaluate information security awareness for home users in India, 64% of home users did not use anti-virus software, 71% were not sure about the security settings in the browsers, around 80% had no clue about malware, spyware and phishing and personal firewalls [8].

The above findings from different countries around the world indicate that significant gaps in understanding and knowledge exist even with the most basic of concepts. These leaves users open to a variety of attacks that would compromise their systems and information. This paper examines the efforts being made, both in real-world implementations and within research towards awareness and education of users with cyber security. Section 2 provides an analysis of online awareness portals, with section 3 presenting a critique of the research literature. Section 4 presents a comparative analysis alongside technology growth and

identifies a number of requirements that future awareness systems must consider.

## II. INFO-SEC AWARENESS WEB PORTALS

Web-based portals are one of the most commonly used tools which are available for educating home users on how to use technologies in a secure manner. Most notable amongst them include:

- Get Safe Online [9].
- Microsoft YouthSpark [11]
- Norton Family Resources [13]
- Internet Matters [15].
- Stay Safe online [10].
- Google Safety Centre [12]
- UK Safer Internet Center [14]
- Childnet International [16]

Table I presents an analysis of the cyber security awareness portals which are available for home users based on some criteria such as context awareness, age groups and delivery methods.

TABLE I. A REVIEW FOR THE CYBER SECURITY PORTALS FOR HOME USERS

		Get Safe Online	Stay Safe online	Microsoft YouthSpark	Google Safety Centre	Norton	UK Safer Internet	Internet Matters	Childnet
<b>Context awareness</b>	Generic Information	✓	✓	✓	✓	✓	✓	✓	✓
	PC	✓		✓					
	Laptop	✓						✓	
	Tablet	✓					✓	✓	
	Smartphones	✓	✓	✓			✓	✓	
	Game Console			✓			✓	✓	✓
	Smart TV	✓					✓	✓	
<b>Age groups</b>		✓					✓	✓	✓
<b>Content delivery method</b>	Quiz	✓	✓				✓	✓	✓
	Games		✓				✓		✓
	Infographics		✓						
	Videos	✓	✓	✓	✓	✓	✓	✓	✓
	Presentations			✓			✓		✓
	Checklists			✓			✓	✓	✓
	Leaflets	✓						✓	
	Posters		✓	✓					
<b>Separate website</b>		✓	✓				✓	✓	✓

Most of these websites try to deliver the awareness content via many methods such as quiz, posters and videos. In addition, some of the portals classify and organise the topics

and the content into several groups based on the age of the user such as children, teens and parents. However, there are some points which might make these portals less useful and productive:

- The majority of these websites only provide generic information without providing a customized awareness based on what the users currently need.
- The portals do not cover most of the technologies, devices and online threats related to the home environment.
- Most of the websites provide a long written tips and advice which might frustrate the users to interact and read the awareness contents.
- Some of the portals have some usability issues such as poor navigation, bad structure and inappropriate colours. Microsoft YouthSpark , Norton Family Resources and Google Safety are operating and linked under the main website of their companies which could make it difficult for people to find these portals.

## III. AN ANALYSIS OF THE CURRENT STATE OF THE ART

With a view to examining the current state of the art, this section seeks to explore the academic literature within the information security awareness domain. A total of 21 papers have been identified and categorised into four domains:

- General security awareness: (3 papers)
- Security guidelines and controls (4 papers)
- Web threat awareness (7 papers)
- Gamification (7 papers)

The papers are summarised in Table II.

### A. General Security Awareness

A theoretical E-Awareness Model (E-AM) was suggested by Kritzinger and Von Solms [17] which contains an awareness and enforcement component. The awareness topics are classified based on the level of the home user knowledge which is divided into three levels: novice, intermediate and advanced and the users can be evaluated and tested in each level. In the enforcement stage: they suggest that the portal can be hosted in with regulating services such as information services providers (ISP) to ensure that all the users cannot access the Internet without accessing E-Awareness portal. This type of restricted enforcement might annoy and disturb the users' activity which might lead them to try to bypass the portal. In addition, this suggestion which allow the ISPs to manage the security awareness could raise some concerns regarding some aspects such as additional financial cost, legal and privacy issue, technical issues and dealing with multiple devices connecting via different ISP.

Labuschagne & Eloff [18] claimed that many internet users use shared resources and computers in some African rural areas and they are not aware of online risks and threats. Therefore, they decided to use the point of shared devices as a positive point and proposed a Shared Public Security Awareness (SPSA) system based on a virtualized environment. The system tried to enforce the user to answer the security awareness questionnaire or to access the awareness content before using the internet. The users will be

TABLE II. STUDIES PROPOSING INFORMATION SECURITY AWARENESS TOOLS FOR HOME USERS

Authors	Awareness method	Domain	Timeliness	Enforcement	Motivation	Evaluation
Tolnai & Von Solms (2009) [19]	Awareness Portal	General Security awareness	No	No	No	No
Caceres & Teshigawara (2010) [20]	Security Guideline tool	Security guideline & controls	No	No	No	Yes
Kritzinger and Von Solms (2010) [17]	Theoretical model	General security awareness	Yes	Yes	No	No
Sharifi et al. (2011) [22]	Web browser extension	Web threats awareness	Yes	Yes	No	No
Maurer et al. (2011) [23]	Web browser extension	Web threats awareness	Yes	Yes	No	Yes
Arachchilage & Cole (2011) [24]	Awareness game	Web threats awareness	No	No	Yes	No
Labuschagne et al. (2011) [25]	Awareness game	General Security awareness	No	No	Yes	No
Labuschagne & Eloff (2012) [18]	Security Awareness system	General Security awareness	No	Yes	No	No
Magaya & Clarke (2012) [26]	risk analysis tool	Security guideline & controls	Yes	No	No	Yes
Jahankhani et al. (2012) [27]	Theoretical awareness tool	Web threats awareness	Yes	Yes	No	No
Fruth et al. (2013) [28]	Awareness game	General Security awareness	No	No	Yes	Yes
Juhari & Zin (2013) [29]	Awareness game	General Security awareness	No	No	Yes	Yes
Smith et al. (2013) [30]	Awareness portal	Web threats awareness	No	No	No	Yes
Serrhini & Moussa (2013) [31]	Safe browser tool	Security guideline & controls	Yes	Yes	No	Yes
Potgieter et al. (2013) [32]	Theoretical browser extension	Web threats awareness	Yes	Yes	No	No
Cetto et al. (2014) [33]	Awareness game	Web threats awareness	No	No	Yes	No
Volkamer et al. (2015) [34]	Web browser extension	Web threats awareness	Yes	Yes	No	Yes
Rani & Goel (2015) [35]	Expert tool	Security guideline & controls	No	No	No	No
Hale et al. (2015) [36]	Awareness game	Web threats awareness	Yes	No	Yes	Yes
Giannakas et al. (2015) [38]	Awareness game	Web threats awareness	No	No	Yes	Yes
Karavaras et al. (2016) [39]	Theoretical application	Web threats awareness	Yes	No	No	No

prevented from accessing only the internet if their level of security knowledge is not satisfactory. In addition, the proposed awareness contents are not provided based on their needs which might be generic and not useful for them. In addition, this approach does not have a centralized management as it is proposed to work only with shared computers on virtual platforms.

Tolnai & Von Solms [19] proposed an Information Security Awareness Portal (ISAP) to be used as an educational source to learn about online threats. The proposed portal consists of several limited categories such as the internet, online transactions, countermeasures.

#### B. Security Guideline and Controls

Caceres & Teshigawara [20] proposed an approach to design a security guideline tool for home users based on international standards to help them to understand the online

threats and allow them to stay protected. The tool succeeded in improving the cyber security knowledge of the users about some online threats. Nevertheless, the approach requires some information which needs to be provided by the users such as the symptoms of the threats or the current risk that they currently face. Therefore, it might not be useful if the user is not aware of the current threat. In addition, the tool is designed based on particular security standards which means the threat which are not mentioned in the standards will not be included in the guideline.

A web-based risk analysis tool (WEBRA) was proposed for home users by Magaya & Clarke [26]. The tool utilised the ISO 27002 and NIST SP 800 – 30 standards in order to specify assets and the implemented controls. The user online behaviour will be evaluated and awareness by answering 18 questions in different topics such as passwords, backup and encryption. In addition, the risk of the missing controls will

analyzed and prioritized from high to low. As a final stage, the tool will provide the users with a recommendation page which contains recommendations and links to websites in order to help them to implement the missing security controls for each device. The tool is very easy to use and does not require any prior practical skills. However, the process of detecting the controls currently implemented is done in a manual way from the user side which could be difficult for some novice users. Moreover, the tool does not have the ability to check the effectiveness of each implemented control. For instance, if the user selects that the password is configured, the tool will exclude it from the missing control list without identifying the password strength which might be weak.

Serrhini & Moussa [31] claimed that many home users do not have an appropriate knowledge in which security features should be enabled or disabled in the browsers. Therefore, they developed a tool for home users called Automatic Safe Browser Launcher which can allow a user to surf the Internet securely. The tool has the ability to scan and detect all the installed web browsers. Once the user selects and clicks the preferred web browser, most of the required security setting features which are identified by the authors, such as updating web browser, network settings, encryption settings and JavaScript settings, will be applied (enable or disable) in the selected browser before it is launched. All the preference reconfiguration for the browsers is applied in the Browser Preferences File or the Windows Key Registry. The user also can view all the misconfigured settings for each browser in the configuration page which has a description of each associated security risk and an activate button to enable each feature automatically. However, the tool provides does not provide the security configuration based on the current needs of the users, only one fixed reconfiguration list for all the users, which might not suit some users and restrict their online activities. This kind of enforcement might result in switching the tool off and using the normal browsers to avoid the restriction.

Some Internet users do not have appropriate knowledge to assist them to identify the issues related to their systems. Rani & Goel [35] designed an Expert System for Cyber Security Attack Awareness (CSAAES) assist the internet users to identify and solve the issues that their computers experience such as viruses, social engineering, SQL injection and data modification. The system has two options to be selected by the users: attack identifier and information about a specific attack type that the user would like to get more information about it. The attack identifier provides a checkbox list which contains 25 symptoms which might the user face some of them. The tool requires some symptoms which might be difficult for the home users to provide them due to their poor knowledge. In addition, the tool might not be accurate in identifying the threat because many threats have the same symptoms. Moreover, the suggested countermeasures are provided without a guideline how to implement each countermeasure.

### C. *Web threats Awareness*

A number of studies has focused on mitigating the risk from browsing the Internet. Sharifi et al. [22] proposed a browser extension which can help to make users aware of the potential threats when they browse websites. The browser extension allows the users post and receive comments about the websites which are being visited. In addition, the tool can

calculate the scam percentage by collecting data about the current websites from several sources. The tool tries to make the user aware of the possible threats at the right time and this depends on the quality of the mutual cooperation between the users as it is a collaborative/community based approach. The tool does not try to provide the users with awareness topics or materials once the users are infected by online scams. Moreover, it has not been evaluated yet by the end users for assessing the functionality and the usability of the tool.

Another Firefox plugin was designed by Maurer et al. [23] to raise the cyber security awareness about phishing attacks, draw the users' attention and make them aware when they deal online with confidential data such as credit card numbers, passwords and transaction authentication number (TAN) at the appropriate time. The tool has been evaluated in different case studies and the results showed the tool was acceptable by the participants and they were able to identify the phishing websites easily.

Another portal was designed by Smith et al. [30] which aims at providing awareness of social engineering threats and risk including materials and quizzes. The quizzes are divided into the levels in order to motivate and encourage the users to answer them. These two portals try to provide cyber awareness content about specific threats rather than providing general awareness.

A theoretical model was proposed by Potgieter et al. [32] which aims to promote information security awareness based on behavioral activities when a web browser is used. The model called Targeted Awareness Browser Extension plans to provide the users with particular awareness content when a possible threat might be experienced. For example, if a user is browsing a banking account, an awareness topic about phishing attacks will be shown to the user or awareness content about the risk of the malicious programs and attacks will be displayed if a user is browsing a website which has a malicious application or code. The main idea of the tool is generally good. However, the motivation option is not mentioned in the tool which could affect the functionality of the tool and might lead to uninstall the extension from the browser.

Volkamer et al. [34] developed a tool called PassSec which works as an add-on in Firefox browsers to provide security awareness about using passwords in unsafe websites. PassSec offers two main functions. The first task is to highlight all the password fields in different colors: green if the website is using HTTPS or red if it is using HTTP. The second task is to provide the users with an awareness dialogue when a password is typed in an unsafe website which is using http. The dialogue has a warning headline: "it is insecure to enter a password". In addition, it contains a warning message about the possible result when typing the password insecurely: "Your password could fall into the hands of unauthorized persons and could be used to access your personal data". Moreover, PassSec can provide a secure mode which redirects the users to a secure connection (https) if it is available in the website. , the tool was successful in providing the users with an awareness notification in the right time but it is very limited and only deals with password security in PCs and laptops. Furthermore, the tool has a lack of encouragement and behavior analysis which could enhance the functionality further.

Another theoretical framework, which is called Soc-Aware, was proposed by Karavaras et al. [39] which can provide awareness about the malicious links threats which might be experienced by Facebook users. Soc-Aware filters and check URLs posted in the Facebook account. . Once a post is detected as a malicious URL, the system will notify the user about how many times they experienced malicious actions and provide them with a Facebook page which includes security awareness materials and guidelines in order to mitigate the threats experienced. The tool is only applicable for Facebook users and only covers one threat which is malicious URLs. In addition, the application requires an access permission in order to work which might be considered as a privacy threat.

#### D. Gamification

Cetto et al. [33] designed a game called Friend Inspector to enhance the privacy awareness among users of social network websites. The user are asked to select the most personal photo to him from two photos are brought from his Facebook profile. Next, the selected picture are presented with 20 profiles (user's friends and strangers), the user has to select the correct profiles who can view his shared pictures. At the end of the game, the user receives the overall score and recommendations to enhance their privacy settings. The game succeeded in making the user aware of the possible vulnerabilities which might be caused by the current implemented settings. However, the game requires access to a user's Facebook account which might be risky even though the authors claim that personal data is secured.

Arachchilage & Cole [24] designed a mobile game for home users to educate them how to avoid phishing attacks supported with a reference guide. Another educational mobile game called CyberAware, was developed by Giannakas et al. [38] which aims to allow children to learn about cyber security principles and online issues such as malware and spam while they are playing the game. Only children from 9 to 11 years are the targeted group in this game. In addition, the game did not try to provide awareness materials based on the weaknesses of the users after completing all levels of the game.

Another game called CyberPhishing was proposed by Hale et al. [36] in order to provide online phishing awareness. The game simulates the phishing attacks in only three aspects: email, web browsing and social media. The users' behavior and actions are analysed which can help to identify the user's weakness and the required training in order to mitigate the possible risk. The tool does not have any awareness materials or topics which can be provided to the users based on the simulation results. Therefore, the tool should tailor the users to the required awareness material based on the result of the analysis behaviour.

#### IV. DISCUSSION

A number of studies have tried to provide the home users with cyber security awareness which are tailored to their needs in different aspects. Providing an appropriate content based on the level of the cyber knowledge for the users was suggested by Kritzinger and Von Solms [17]. This approach might not be accurate due to the difference in the knowledge between the users. Another attempt was proposed by Magaya & Clarke [26] to provide a bespoke recommendation and guideline based on the result of the risk assessment for the

home users but the security controls are identified manually by the users which might be difficult for the novice users. Therefore, it would be beneficial if the automation option was exploited in the tool.

Other attempts have been done by researchers to provide a particular awareness when the users are browsing the internet. Sharifi et al. [22], Maurer et al. [23] and Potgieter et al. [32] proposed a browser extension to make the user aware of the phishing websites and the possible threats while surfing online, whereas Volkamer et al. [34] designed a tool to show an awareness notification when the users are browsing insecure websites with password fields. While Karavaras et al. [39] and Cetto et al. [33] introduced approaches which can provide a tailored awareness for the Facebook users.

Some studies such as Kritzinger and Von Solms [17] Labuschagne & Eloff [18] Serrhini & Moussa [31] have tried to restrict the home users' online activity and force them to apply security settings or to read awareness materials. This type of enforcement could create a level of undesirability which will result in the solution being switched off or uninstalled. In addition, it has been suggested in some studies that the cyber security awareness can be managed by the ISPs.

However, this is not a workable solution because of many issues such technical, privacy, financial and legal issues. The functionality of this approach can become more complicated if there are multiple technologies which are working on multiple ISPs. The motivation for this suggestion is that the authors want to give this responsibility to someone who is better able to manage it. Therefore, it would be a good idea if and individual can take the lead. For example, a member of each family, who is interested in technology, can manage the home network and digital devices. In addition, the enforcement option might not be the best approach within the home environment due to the lack of auditing, policies and penalties. In addition, the enforcement can cause a resistance from the family members which can lead to the system is not being used.

The vast majority of the educational games are dedicated in a single area, limited scope and they do not successfully adapt to the multi-threat, multi technology and services. They have not tried to provide tailor-made awareness content based on the present needs of the users. In addition, they are designed to offer cyber awareness for children without providing valuable awareness for the rest of the family members.

As most of the tools are optional to be used and their main stakeholder is the home users, it is important to encourage and motivate the users to get engaged with tool in order to promote the cyber knowledge and awareness. The majority of the tools have not introduced any kind of motivations such as scores or digital certificates. For example, it can be easily introduced digital badges or a cyber hero of the week in the family unit which could help to create a motivational environment between the family members.

In addition, the studies conducted by Rao & Pati [8] and Howe et al. [40] revealed that there is a clear need to design an friendly usable approach which can manage security controls, security configurations and installed software in a wide variety of platforms and devices in order to promote

cyber security awareness and provide protection for internet users.

From the prior discussion, it is clear that there is a need for bespoke individualized personalized approach that takes into account knowledge and awareness of the technologies, applications and services that users use and provides bespoke information directly based upon the current security posture. In order to measure and understand how the home users are doing something, well or badly? , it needs to be defined against something by using security policy in order to deliver customized awareness contents. Despite of the fact that many approaches and tools have been proposed for the home users to promote cyber security, they are providing general, static and limited awareness content. Therefore, there is a need to provide the users with some kinds of polices which can deliver a customized awareness content.

#### V. CONCLUSION AND FUTURE WORK

This paper reviewed the existing security awareness tools for home users in terms of the timeliness,, the mechanism and the effectiveness. The above discussion indicates that there is a lack in providing the home users with a bespoke awareness when they really need it. Further research will also be undertaken in proposing and designing a holistic information security management system for home users which can provide tailored security awareness by applying some groups of policies.

#### REFERENCES

- [1] Statista, "Number of internet users worldwide 2005-2016 | Statista," 2016. [Online]. Available: <https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/>. [Accessed: 22-Dec-2016].
- [2] Futuretimeline, "Future 2020 | Internet use reaches 5 billion worldwide," 2016. [Online]. Available: <http://www.futuretimeline.net/21stcentury/2020.htm#internet-2020>. [Accessed: 22-Dec-2016].
- [3] Ofcom, "Adults' media use and attitudes," 2015. [Online]. Available: [http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/media-lit-10years/2015\\_Adults\\_media\\_use\\_and\\_attitudes\\_report.pdf](http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/media-lit-10years/2015_Adults_media_use_and_attitudes_report.pdf). [Accessed: 25-Jun-2017].
- [4] H. Taylo, "Biggest cybersecurity threats in 2016," 2015. [Online]. Available: <http://www.cnbc.com/2015/12/28/biggest-cybersecurity-threats-in-2016.html>. [Accessed: 05-Mar-2017].
- [5] Symantec, "2016 Internet Security Threat Report," 2016. [Online]. Available: <https://www.symantec.com/content/dam/symantec/docs/reports/istr-21-2016-en.pdf>. [Accessed: 25-Jun-2017].
- [6] NCSA and McAfee, "2011 NCSA / McAfee Internet Home Users Survey," 2011. [Online]. Available: [https://staysafeonline.org/download/datasets/2068/NCSA\\_McAfee\\_Online\\_User\\_Study\\_Final\\_11\\_15\\_11.pdf](https://staysafeonline.org/download/datasets/2068/NCSA_McAfee_Online_User_Study_Final_11_15_11.pdf). [Accessed: 22-Jun-2017].
- [7] NCSA and PayPal, "2013 NATIONAL ONLINE SAFETY STUDY," 2013. [Online]. Available: [https://staysafeonline.org/download/datasets/7358/2013\\_NCSA\\_Online\\_Safety\\_Study.pdf](https://staysafeonline.org/download/datasets/7358/2013_NCSA_Online_Safety_Study.pdf). [Accessed: 22-Jun-2017].
- [8] U. H. Rao and B. P. Pati, "Study of internet security threats among home users," *2012 Fourth Int. Conf. Comput. Asp. Soc. Networks*, pp. 217–221, 2012.
- [9] GetSafeOnline.org, "Get Safe Online | Free online security advice," 2016. [Online]. Available: <https://www.getsafeonline.org/>. [Accessed: 22-Dec-2016].
- [10] StaySafeOnline.org, "National Cyber Security Alliance | StaySafeOnline.org," 2016. [Online]. Available: <https://staysafeonline.org/>. [Accessed: 22-Dec-2016].
- [11] Microsoft, "Online Safety - YouthSpark Hub," 2016. [Online]. Available: <https://www.microsoft.com/about/philanthropies/youthspark/youthsparkhub/programs/onlinesafety/>. [Accessed: 22-Dec-2016].
- [12] Google, "Google Safety Centre," 2016. [Online]. Available: <https://www.google.co.uk/intl/en/safetycenter/>. [Accessed: 22-Dec-2016].
- [13] Norton, "Norton Security Center, A Complete Online Security Resource," 2016. [Online]. Available: <https://us.norton.com/security-center/>. [Accessed: 22-Dec-2016].
- [14] Saferinternet.org.uk, "UK Safer Internet Centre," 2016. [Online]. Available: <http://www.saferinternet.org.uk/>. [Accessed: 22-Dec-2016].
- [15] Internetmatters.org, "Information, Advice and Support to Keep Children Safe Online," 2016. [Online]. Available: <https://www.internetmatters.org/>. [Accessed: 22-Dec-2016].
- [16] Childnet, "Childnet International," 2016. [Online]. Available: <http://www.childnet.com/>. [Accessed: 22-Dec-2016].
- [17] E. Kritzinger, S. Von Solms, and S. H. Von Solms, "Cyber security for home users: A new way of protection through awareness enforcement," *Comput. Secur.*, vol. 29, pp. 840–847, 2010.
- [18] W. A. Labuschagne and M. Eloff, "Towards an automated security awareness system in a virtualized environment," in *11th European Conference on Information Warfare and Security 2012, ECIW 2012*, 2012, pp. 163–171.
- [19] A. Tolnai and S. Von Solms, "Solving security issues using information security awareness portal," *2009 Int. Conf. Internet Technol. Secur. Trans.*, pp. 1–5, 2009.
- [20] G. R. Caceres and Y. Teshigawara, "Security guideline tool for home users based on international standards," *Inf. Manag. Comput. Secur.*, vol. 18, no. 2, pp. 101–123, 2010.
- [21] E. Kritzinger and S. H. Von Solms, "Cyber security for home users: A new way of protection through awareness enforcement," *Comput. Secur.*, vol. 29, no. 8, pp. 840–847, 2010.

- [22] M. Sharifi, E. Fink, and J. G. Carbonell, "SmartNotes: Application of crowdsourcing to the detection of web threats," *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, pp. 1346–1350, 2011.
- [23] M. Maurer, A. De Luca, and S. Kempe, "Using Data Type Based Security Alert Dialogs to Raise Online Security Awareness," *SOUPS '11 Proc. Seventh Symp. Usable Priv. Secur.*, p. Paper 2, 2011.
- [24] N. A. G. Arachchilage and M. Cole, "Design a mobile game for home computer users to prevent from phishing attacks," *Int. Conf. Inf. Soc. (i-Society 2011)*, pp. 485–489, 2011.
- [25] W. A. Labuschagne, I. Burke, N. Veerasamy, and M. M. Eloff, "Design of cyber security awareness game utilizing a social media framework," *2011 Inf. Secur. South Africa - Proc. ISSA 2011 Conf.*, 2011.
- [26] R. T. Magaya and N. L. Clarke, "Web-based risk analysis for home users," *Proc. 10th Aust. Inf. Secur. Manag. Conf. AISM 2012*, pp. 19–27, 2012.
- [27] H. Jahankhani, T. Jayaraveendran, and W. Kapuku-Bwabw, "Improved awareness on fake websites and detecting techniques," *Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng.*, vol. 99 LNCS, pp. 271–279, 2012.
- [28] J. Fruth, C. Schulze, M. Rohde, and J. Dittmann, "E-learning of IT security threats: A game prototype for children," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8099 LNCS, pp. 162–172, 2013.
- [29] S. F. Juhari and N. A. M. Zin, "No Educating Children about Internet Safety through Digital Game Based Learning," *Int. J. Interact. Digit. Media*, vol. 1, no. 1, pp. 65–70, 2013.
- [30] A. Smith, M. Papadaki, and S. M. Furnell, "Improving awareness of social engineering attacks," *IFIP Adv. Inf. Commun. Technol.*, vol. 406, pp. 249–256, 2013.
- [31] M. Serrhini and A. A. Moussa, "Home users security and the web browser inbuilt settings, framework to setup it automatically," *J. Comput. Sci.*, vol. 9, no. 2, pp. 159–168, 2013.
- [32] M. Potgieter, C. Marais, and M. Gerber, "Fostering Content Relevant Information Security Awareness through Browser Extensions," *Inf. Assur. Secur. Educ. Train. 8th IFIP WG 11.8 World Conf. Inf. Secur. Educ.*, pp. 58–67, 2013.
- [33] A. Cetto, M. Netter, and G. Pernul, "Friend Inspector: A Serious Game to Enhance Privacy Awareness in Social Networks," *arXiv Prepr. arXiv ...*, 2014.
- [34] M. Volkamer, K. Renaud, G. Canova, B. Reinheimer, and K. Braun, "Design and Field Evaluation of PassSec: Raising and Sustaining Web Surfer Risk Awareness," in *Trust and Trustworthy Computing: 8th International Conference, TRUST 2015, Heraklion, Greece, August 24–26, 2015, Proceedings*, M. Conti, M. Schunter, and I. Askoxylakis, Eds. Cham: Springer International Publishing, 2015, pp. 104–122.
- [35] C. Rani and S. Goel, "CSAAES: An expert system for cyber security attack awareness," *Int. Conf. Comput. Commun. Autom. ICCCA 2015*, pp. 242–245, 2015.
- [36] M. L. Hale, R. F. Gamble, and P. Gamble, "CyberPhishing: A game-based platform for phishing awareness testing," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2015–March, pp. 5260–5269, 2015.
- [37] F. Giannakas, G. Kambourakis, and S. Gritzalis, "CyberAware: A mobile game-based app for cybersecurity education and awareness," *2015 Int. Conf. Interact. Mob. Commun. Technol. Learn.*, no. November, pp. 54–58, 2015.
- [38] F. Giannakas, G. Kambourakis, and S. Gritzalis, "CyberAware: A mobile game-based app for cybersecurity education and awareness," *2015 Int. Conf. Interact. Mob. Commun. Technol. Learn.*, no. November, pp. 54–58, 2015.
- [39] E. Karavaras, E. Magkos, and A. Tsohou, "Low User Awareness Against Social Malware: An Empirical Study and Design of a Security Awareness Application," *13th Eur. Mediterr. Middle East. Conf. Inf. Syst. (EMCIS 2016)*, pp. 1–10, 2016.
- [40] A. E. Howe, I. Ray, M. Roberts, M. Urbanska, and Z. Byrne, "The Psychology of Security for the Home Computer User," *2012 IEEE Symp. Secur. Priv.*, pp. 209–223, 2012.

# Monitoring Darknet Activities by Using Network Telescope

Shaikha AlShehyari, Chan Yeob Yeun, Ernesto Damiani  
Khalifa University of Science and Technology, Information Security Research Center,  
PO Box 127788, Abu Dhabi, UAE

Email: {shaikha.alshehyari, chan.yeun, ernesto.damiani}@kustar.ac.ae

**Abstract**— Today, many hosts are connected to the Internet worldwide. Those hosts are intentionally or even accidentally targeted in a daily basis by malicious activities. Thus, it is essential to monitor Internet traffic in order to acquire the knowledge of the threats that is targeting computers and network systems. Security experts initiated numerous approaches for monitoring traffic that includes malicious activities, and network telescope was one of them. Network telescopes are valuable in the term of monitoring and gathering data associated with Internet attack activities. Analysis of traffic captured by network telescopes has been shown as an effective measure in characterizing evil traffic caused by worm propagation and distributed denial of service attacks. This paper will discuss monitoring traffic coming from Darknet using network telescope. Additionally, we will identify the security issues and threats introduced if no monitored device is placed within a network infrastructure. The objective of this paper is to introduce the effect of such a technology on entities, provide successful solution, and encourage future research in this area of interest.

**Keywords**— Darknet, Network Telescope, DDoS, Worm

## I. INTRODUCTION

With the advance innovation in technology, devices are getting cheaper and easily available to public. As a result, more and more devices are connected to the ubiquitous networks and virtualized frameworks in [1-4]. This expands the attack vector for these substances and guarantees an expanding populace of hosts for future assaults. Darknet traffic is usually associated with Cyber-attacks that occurs within the Cyber pool [5]. It is a network where it has a set of unused IP addresses with no server allocated. It can be accessed through specific applications, configurations, and often used via non-standard communications protocols and ports. Early 2000s, network traffic was directed to a Darknet, and by analyzing it, one could identify remote network behaviour or even malicious Internet activities. Malicious activities in the Darknet include network scanning, address spoofing, and Internet worms [6] used to disrupt networks [7].

The dark web has shown exemplary results in terms of privacy. This platform has a good record of online anonymity to a great extent. The prevalence of systems like The Onion Router (Tor) [8] can serve as examples of the level of privacy that Darknet offers. The negative side of these kinds of platforms is the ability to grant cover to persons in regressive regimes that require protection for the purpose of web crawling. Illegal markets, trolls among other evil acts are some

of the dark sides of the platform. Massive criminal marketplaces reside in Darknet. Anything could be bought from drugs to assassinations. The most well-known criminal marketplace was Silk Road. Recently, a major police operation took down the Silk Road which was famous of selling illegal items including hacked PayPal accounts, drugs, fake passports, and other IDs. Darknet users uses bitcoins to conduct transactions and this is because they want to remain entirely anonymous within Darknet [9]. One of the effective ways to watch Internet activity is to employ passive monitoring using sensors or traps such as Darknet [10].

Monitoring packets destined to unused Internet addresses has become an increasingly important measurement technique for detecting and investigating malicious Internet activity [11]. Proactive Cyber-security tools provide basic protection as today's Cyber-criminals utilize legitimate traffic to perform attacks and remain concealed quite often until it is too late [12]. It is known that no automated solution (IDS or IPS) can prevent malicious activities from getting inside the network. Therefore, it is easy for the attacker to target an organization using Darknet traffic to infect its infrastructure by malware. It is well-known that Darknet traffic is often logged, however it is not frequently that it is carefully considered. Even worse, entities do not filter the accurate Darknet traffic, thus, they always fall into obtaining false positive alarms. Internet Service Providers and entities are interested in traffic that is anomalous relative to regular traffic received. Entities that is highly targeted by threat actors should deploy their own network telescope within their IT infrastructure. Network telescopes can be integrated into an organisation's existing security framework easily [13]. It is important to build up an artificial intelligence using data mining to aid in the identification of network incidents in a relatively short period of time. Data mining is used to discover patterns in large data sets. WEKA [14] is a famous data mining tool for analyzing and building the predictive model for the unlabelled data classification and prediction. It consists of a collection of machine learning algorithms such as Clustering, Feature selection/Attribute subset selection, Classification and Association Rule mining [14]. Network telescope (NT) is called BlackHoles [15], Darknets, or Internet Sinks [16]. Network Telescope is an Internet system that allows user to monitor diverse and large-scale of events taking place on the World Wide Web (Internet) [17]. The idea is to perceive traffic targeting the unused address-space of the network

within the Darknet. Since all traffic associated with those addresses is suspicious, one can gain information about possible network attacks. Network Telescope acts like trap-based monitoring systems that aim to deploy online sensors to trick and trap adversaries to collect malicious activities [10]. There is two type of network telescope. One can monitor the traffic passively, while the other could perform active monitoring achieved by advanced telescopes. Telescopes are able to provide empirical data based on malicious traffic, unbiasedly targeted towards unused address space. Honeypots are very similar in nature to network telescopes [14], they do however serve a more specific purpose in that they are used to emulate a vulnerable service or host in order to attract malicious traffic [18].

The motivation of this paper is to propose a framework that will be a baseline for establishing a network telescope that could be deployed in different entities within UAE. The telescope named as “U-Telescope” will help in capturing Darknet traffic. Limited number of related work have been done before in this area, however, additional work has to be performed to develop the limitations and the shortcomings of the framework related to pervious works. In addition, there is a need to enhance the data visualization to the data collected by the telescope. Through correct visualisation, large sets of data can be summarised in a compact and easily understandable format [19]. Moreover, data mining to the whole data captured will be taken to measure the traffic in different perspectives such as location, source IP and destination IP.

The remainder of the paper is structured as follows. The next section discusses the related work. The proposed framework for deploying network telescope in IT environment is described in Section III. Finally, conclusions and future works are discussed in Section IV.

## II. RELATED WORK

Capturing traffic and data received from Darknet is associated with several challenges: improper implementation of network framework, network misconfiguration and the lack of accurate data related to the Darknet traffic. Below, relevant background information about network telescope will be provided.

### A. Topological Models and Effectiveness of Network Telescopes [20]

Various attacks could be lunched through unused IP addresses network. The quick broadcast of malware such as worms, remain a significant threat. The low latency and high bandwidth of such networks facilitates extremely rapid attack patterns and worm propagation, leaving very little time for active countermeasures [20]. Therefore, it is hard to obtain as much early warning information as possible to contribute in configuring appropriate defensive mechanisms. Network Telescopes NT, a new method for monitoring and measuring the network proposed by Moore et al. from the Cooperative Association for Internet Data Analysis (CAIDA). Network

telescopes detect suspicious behaviour by detecting unusual phenomena that exist in Darknet.

F. Gagadis And S. Wolthusen [20] argues that at present, network telescopes are used regularly for academic motivations and data collection on distributed topologies for Internet examinations. As described above, traffic that are illegitimate will be observed by the telescope. This traffic could be the consequence of worms scanning, misconfiguration or backscatter traffic from spoofed addresses. Monitoring large space of IP addresses will contribution in detecting massive number of security incidents that will help in collecting more data for analysis. Network telescopes have the ability to survive an incident or attack by itself. It does not trigger false alarms [20]. There are two general topological models of network telescopes:

- Passive telescope: observes the packets arriving, keeps logs and later discards them without further interactions with the attacker [20].
- Active telescope: observes the incoming packets and respond to back to them to establish communication channels until the incident is known. It emulates services, analyzes attacks and keeps tracks of the attacker.

Network telescopes are operative and useful tools for witnessing large-scale events and requiring careful observations. In deploying network telescopes, researchers must study the various constraints of a network to eliminate false alarms and false positive results.

### B. Practical Darknet Measurement [11]

These days, Internet has become a place for constant attacks targeting users and entities’ infrastructure. There are many different methods to identify these attacks. One of these methods is observing the unused network addresses. The reason for that is that many attacks propagate randomly, so threats can be detected by only monitoring unused spaces between live addresses. Sensors that observe unused address space are called telescopes. They capture significant information about an assorted range of attacks such as Internet worms, denial of services attacks, and botnets. In this paper, the author depicted and analyzed the important measurement issues associated with prevailing Darknet, estimating the position and service configuration of Darknet, and analyzing the information collected by Darknet. The main aim of this is to give a general overview of Darknet measurement and provide researchers with the necessary information to deploy and analyze the information from Darknet monitoring systems.

To illuminate the analysis, the information was utilized from the worldwide deployed Internet Motion Sensor (IMS); distributed Darknet monitoring system. The IMS comprises of 60 Darknet blocks at 18 associations including broadband suppliers, major service providers, huge enterprises, and academic networks in 3 continents. Over 17 million addresses are being monitored.

The paper mentioned three general strategies for transmitting packets to a Darknet monitoring system. The most straightforward approach is to configure the observing box to send ARP responses for each unused address to the router. He also discussed the Darknet placement within a network. If the Darknet monitor is set behind a firewall or any other infrastructure preservation or a filtering device, it will mostly not monitor externally sources attacks. Ideally, a Darknet deployment that contains of monitors deployed both inside and outside network perimeters ought to have the greatest potential visibility. This analysis has endeavoured to present that building and running a Darknet monitor is a simple and productive technique of earning significant extra visibility into network dangers and the condition of local network and Internet entirely.

C. Passive IP Traceback: Capturing the Origin of Anonymous Traffic through Network Telescopes [21]

IP traceback is used to discover the origin of anonymous traffic. Internet Service Providers (ISPs) do not provide support when it comes to Internet-scale IP traceback systems. Thus, such systems could not be deployed. The author of this article presents an Internet-scale Passive IP Traceback (PIT) mechanism that does not require ISP deployment. Fig. 1 below shows the structure of IP traceback approach.

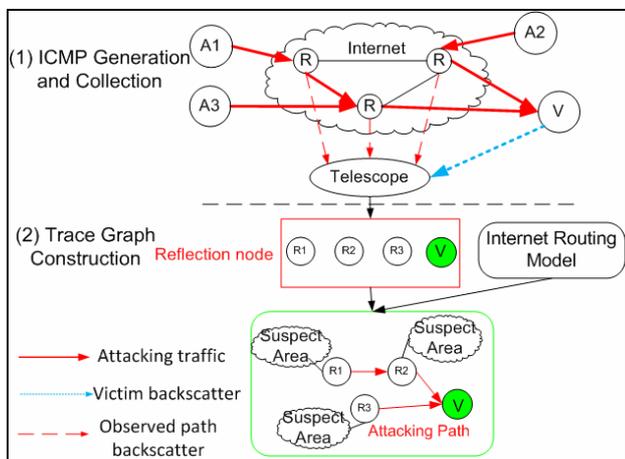


Fig. 1: Passive IP Traceback

The PIT system examines packets collected from network telescopes, and deduces the locations of spoofed traffic. The idea behind PIT system is that spoofing flows may discover ICMP error messages at routers on its way to the victim. Those messages are sent to the spoofed nodes [21]. Assuming that attackers has a routine of using randomly forged addresses, network telescopes will retrieve some ICMP messages. The addresses of routers carrying the message can be joint with an Internet route model to re-build the attack path and discover the locations of the spoofers. As an example, Fig. 2 shows the geolocations of reflection routers identified during a February 2008 attack against a victim in Taiwan. IP addresses were mapped to locations using <http://www.ipaddresslocation.org/>.

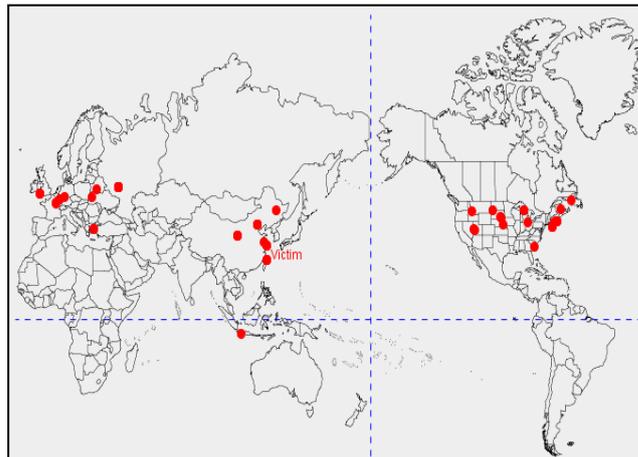


Fig. 2: Reflection locations of an actual attack in Feb. 2008

The paper presented a good approach named as Passive IP traceback mechanism (PIT) and proven its advantages by not requiring new deployment at any router or ISP. Initial results show it is practical though not perfect.

The related work provided above are quite good, but still have limitation when it comes to framework design, data gathering and data visualization.

III. DISCUSSIONS AND ANALYSIS

Currently, cyber security experts face big challenges when it comes to cyber criminals. They provide their entities with the best tools and equipment to ensure that they are in safe from attacks. However, criminals always find a way to get inside the targeted organization.

The proposed related work provided in section II was limited each in its area of studies. The table below shows the differences between each approach taken based on framework design, data gathering and data visualization.

TABLE I. DIFFERENCES BETWEEN THE THREE RELATED WORK

Related work in section II	Telescope framework	Data gathering	Analysis	Data visualization (console)
Topological Models and Effectiveness of Network Telescopes	X	X	X	X
Practical Darknet Measurement	√	√	√	X
Passive IP Traceback	√	√	X	X

From above table, the second related work succeeds to provide better framework design, data gathering and analysis of the data gathered.

In order to evade the threats associated with darknet, a fair understanding of the network framework should be adopted. This will help in placing the network telescope in the right

place to capture all traffic needed. Network experts should collect all the relevant data to be visualized in a console for alert in later stages. Thus, there is a need to place such a “U-Telescope” device to passively monitor their network assets. The project is done in five phases as mentioned below:

1. **Design the “U-Telescope” framework:** it is a single machine that acts as a monitoring device. It is allocated in front of the firewall in order to capture all traffic retrieved from outside world. Fig. 3 shows the “U-telescope” framework.

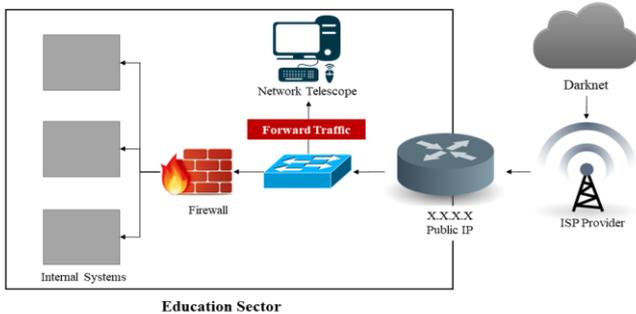


Fig. 3: “U-Telescope” Framework

2. **Data gathering:** the host will have the Wireshark tool that can capture all incoming traffic and filter unwanted/used IP addresses. A dump file from the tool is then generated to view the data.
3. **Data Analysis:** data will be presented in statistics view and graphs. The data will be sorted/filtered by source IP address, destination IP address, location and protocol.
4. **Data simulation:** The simulation will be used on the collected data and programmatically will create a partial data file where it is updated in a specific period of time. This data will be displayed in a console that will trigger the change of data to update its data.
5. **Data mining:** all data gathered will be combined and shown in graphs and charts. This will aid in finding patterns for the current data captured.

Table II below summarizes the provided solution:

TABLE II. PROPOSED SOLUTION

Related work in section II	Telescope framework	Data gathering	Analysis	Data visualization (console)
U-Telescope	√	√	√	√

U-Telescope solution will assure data gathered, analyzed and displayed in a precise manner. The framework was designed in a way to capture all traffic coming from the Darknet. The trap machine (network telescope) gets a copy of all incoming traffic before it goes through firewall. Thus, it will not be filtered.

On 27 of September, 2017, U-Telescope was placed in the Information Technology department within one of the education sector in the UAE as shown Fig. 3. The Darknet traffic was captured using Wireshark tool. Wireshark is a network packet analyzer that capture network packets and tries to display that packet data as detailed as possible [22]. Any traffic that hits the destination public IP X.X.X.X/24, will be triggered and logged within the tool. Fig. 4 below displays the data:

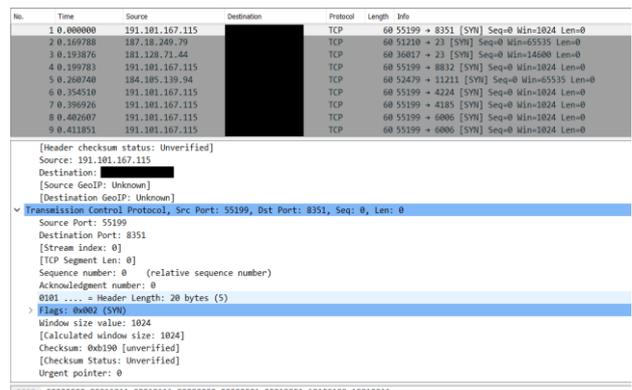


Fig. 4: Darknet traffic captured by Wireshark. The destination IP address was hidden due to the right of privacy

In order to get the data shown in Fig. 4, we had to input a filter to sort out the IP addresses that are in used. While monitoring/capturing the data, a file will be written to the desk with all the packets details. Once the file reach 30MB size, a new file will be created to complete the capture.

After gathering the data, traffic analysis should be established to eyewitness any malicious activities in place. Data will be presented in form of statistics view and in a more human readable format. Fig. 5 below shows the IP addresses coming from the Darknet.

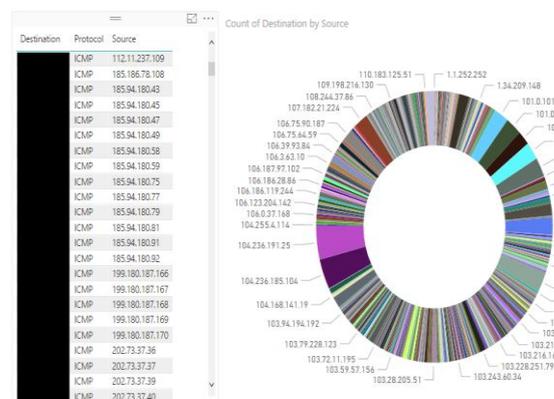


Fig. 5: Statistical view of Darknet traffic

Large volume of data was received to the target IP address. We were able to visualize the data to facilitate the delivery of information. In order to achieve that, we used Microsoft Power BI Desktop tool. The data was shown in different views like bar, pie and scatter charts. We observed diverse protocols such as ICMP, TCP and SNMP. Huge number of Darknet IP addresses were logged. To confirm the malicious intent of those IP addresses, we queried for backlisted IP addresses and compared them to the list found within the data gathering phase. We used the website <http://www.ipvoid.com/ip-blacklist-check/> to query the IP address. Fig. 6 reveals an example of a blacklisted IP address found within the Darknet traffic.

Analysis Date	2017-10-08 18:00:58
Elapsed Time	6 seconds
Blacklist Status	<b>BLACKLISTED 4/97</b>
IP Address	<b>182.84.156.40</b> <a href="#">Find Sites</a>   <a href="#">IP Whois</a>
Reverse DNS	Unknown
ASN	<a href="#">AS4134</a>
ASN Owner	No.31,Jin-rong Street
ISP	China Telecom Jiangxi
Continent	Asia
Country Code	 (CN) China
Latitude / Longitude	<a href="#">28.55 / 115.9333 Google Map</a>
City	Nanchang
Region	Jiangxi

Fig. 6: Blacklist IP address found within Darknet traffic captured

As shown above in Fig. 6, the blacklisted IP address is coming from China as it was used previously to conduct malicious activities.

#### IV. CONCLUSION

In conclusion, network telescopes have proven to be a beneficial approach in the capture of nefarious network traffic. Network security is of growing worry and by employing network telescopes, it allows defending against malicious entities. It is always known that the best solution is to be capable of addressing both network security holes and external entities.

In future, U-Telescope could be allocated in different geolocation within entities all over UAE. This is known as distributed network telescope. This approach will help us to observe internet activities and cyber-attacks via passive monitoring. In addition, it will provide realistic measures and analysis of darknet information because all of these data will be compared to the original darknet data. This project will

attract and assist the cybersecurity firms and organization who are interested in monitoring darknet data. It aims to support the organization in detecting and preventing if possible cybercrime. The project could be improved in later stages to include an alert system. For example, the system will trigger an alarm once it receives massive traffic from a single source, which in theory means that attackers are trying to launch DDOS attack against the victim.

#### REFERENCES

- [1] C.Y. Yeun, E.K. Lua, J. Crowcroft, "Security for emerging ubiquitous networks," In proceeding of the 62nd IEEE Vehicular Technology Conference, Vol. 2, pp. 1242-1248, September 2005.
- [2] D.M. Konidala, C.Y. Yeun, K. Kim, "A secure and privacy enhanced protocol for location-based services in ubiquitous society," In proceeding of IEEE Global Telecommunications Conference, GLOBECOM'04, Vol. 4, pp. 2164-2168, 2004
- [3] J. Baek, Q.H. Vu, A. Jones, S., Al Mulla, C.Y. Yeun, "Smart-frame: A flexible, scalable, and secure information management framework for smart grids," In proceeding of International Conference for Internet Technology And Secured Transactions, pp. 668-673, December 2012.
- [4] F.A. Bazargan, C.Y. Yeun, J. Zemerly, "Understanding the security challenges of virtualized environments," In proceeding of International Conference for Internet Technology And Secured Transactions, pp. 67-72, December 2011
- [5] T. Ban, L. Zhu, J. Shimamura, S. Pang, D. Inoue, and K. Nakao, "Behavior Analysis of Long-term Cyber Attacks in the Darknet", National Institute of Information and Communications Technology 4-2-1 Nukui-Kitamachi, Tokyo, 184-8795, Japan.
- [6] M. Bailey, E. Cooke, D. Watson, F. Jahanian, and J. Nazario, "The Blaster Worm: Then and Now.", IEEE Security & Privacy, 3(4):26-31, 2005.
- [7] B. Irwin , "A Baseline Study of Potentially Malicious Activity Across Five Network Telescopes", International Conference on Cyber Conflict, 2013.
- [8] J. Buxton, T. Bingham, "The Rise and Challenge of Dark Net Drug Markets", Global Drug Policy Observatory, January 2015.
- [9] E. Jardine, "The Dark Web Dilemma:Tor, Anonymity and Online Policing", Centre for International Governance Innovation, SEPTEMBER 2015.
- [10] C. Fachkha, M. Debbabi, "Darknet as a source of cyber intelligence: Survey, taxonomy, and characterization," IEEE Communications Surveys & Tutorials. Vol. 18, No. 2, pp. 1197-227, January 2016.
- [11] Bailey, M. et al. "Practical Darknet Measurement." 2006 40th Annual Conference on Information Sciences and Systems, 1496-1501, 2006.
- [12] P. Chatziadam, I. Askoxylakis, A. Fragkiadakis, "A Network Telescope for Early Warning Intrusion Detection", 10.1007/978-3-319-07620-1\_2, 2014.
- [13] B. Irwin, "A framework for the application of network telescope sensors in a global IP network", Rhodes University, 2011.
- [14] D. Singh, E. Leavline, "Data Mining In Network Security - Techniques & Tools: A Research Perspective", Journal Of Theoretical And Applied Information Technology, 20th November 2013.
- [15] E. Cooke, M. Bailey, Z. Mao, D. McPherson ,D. Watson, F. Jahanian , "Toward Understanding Distributed Blackhole Placement" , Association For Computing Machinery,2004.
- [16] K.Priya, Selvi.U, "Passive IP Traceback: Tracing spoofer IP and Blocking IP spoofer from Data access using Greedy Algorithm", International Journal of Emerging Trends in Science and Technology, Pages 4119-4131, 2016.
- [17] D. Moore, G. Voelker and S. Savage, "Quantitative Network Security Analysis",CAIDA/SDSC and CSE Department,University of California, San Diego, Dec 4, 2002.
- [18] S. Hunter, "Literature Review: Network Telescope Dashboard and Telescope Data Aggregation", 20 June 2010.

- [19] S. Hunter, "A Network Telescope Information Visualisation Framework", Rhodes University, November 2010.
- [20] F. GAGADIS and S. WOLTHUSEN, "Topological Models and Effectiveness of Network Telescopes", the Information Security Group at Royal Holloway, University of London in conjunction with TechTarget, 2008.
- [21] G. Yao, J. Bi, Z. Zhou, "Passive IP Traceback: Capturing the Origin of Anonymous Traffic through Network Telescopes". Computer Communication Review - CCR. 413-414. 10.1145/1851275.1851237, 2010.
- [22] U. Lamping, R. Sharpe, and E. Warnicke, "Wireshark User's Guide", 2004-2013.

# Enhancing Cyber Security Awareness with Mobile Games

F.Alotaibi<sup>1</sup>, S.Furnell<sup>1,2,3</sup>, I.Stengel<sup>1,4</sup>, M.Papadaki<sup>1</sup>

<sup>1</sup> Centre for Security, Communications and Network Research, University of Plymouth, Plymouth, UK

<sup>2</sup> Security Research Institute, Edith Cowan University, Perth, Western Australia

<sup>3</sup> Centre for Research in Information and Cyber Security, Nelson Mandela University, Port Elizabeth, South Africa

<sup>4</sup> Hochschule Karlsruhe, University of Applied Sciences Karlsruhe, Germany

{Faisal.Alotaibi, Steven.Furnell, Maria.Papadaki}@plymouth.ac.uk, Ingo.Stengel@hs-karlsruhe.de

**Abstract-** The ever-increasing threats on cybersecurity has consequently increased the need for enhanced awareness about cybersecurity and its various threats among public. This paper presents the design aspects of the two mobile gaming applications including Malware Guardian game, and Password Protector game. Further, different mobile games concept developed during the course of the study is also presented. The malware Guardian was aimed at educating the users about different security threats, security issues, the risks associated with it and the tools to be used for preventing these attacks. The Password Protector is aimed at educating the users about the need for creating strong and complex passwords, remembering and changing them frequently. Both the games are in the final phase of design, and will be developed once the design specifications are outlined. The major design aspects of the two gaming applications are presented in this paper, which are carefully crafted after a peer review of functions and technologies to be associated with.

**Keywords –** CyberSecurity, Gaming Technologies, CyberSecurity Awareness, Mobile Applications, Gamifying

## I. INTRODUCTION

The ever increasing threats on the security of information technology systems and in the cyber world have raised the importance of the cybersecurity awareness for both individuals and organizations. However, being aware about cybersecurity is not given enough importance and is often neglected, especially in workplaces. For any organization, it is important that the employees are aware and up-to-date with cybersecurity threats as it can potentially reduce incidents of security breaches. However, according to UK Cyber Security Breaches Survey, it was observed that only 30% of the organizations provide awareness regarding cybersecurity to their employees [1]. The research study presented aims at addressing the issue of creating cybersecurity awareness by employing gameplay methods over traditional approaches to make awareness process more engaging. As a part of study, 12 game concepts were shortlisted, and a brief overview of these games is presented, along with the design concepts of two selected games for implementation.

## II. BACKGROUND STUDY

It is important that the users are trained to be cybersecurity aware. An effective approach to achieve this is employing serious games that are designed to achieve a purpose along with entertaining the players. Serious games are considered to be effective in achieving impactful training and behavioral change among users. The approach of using games for training is referred to as games based learning approach. These games are popularly used in school education. More recently, they are also adopted for healthcare, advertising, behavioral change, and cybersecurity training [2].

A key feature of games based learning method is that it provides an interactive approach to train and educate users regarding the chosen topic. The games can impart or enhance the target skills of the players via a fun, engaging manner. The design of the games is flexible and can be adapted to meet the requirements of virtually every topic of interest for training [3]. A good game takes the player into a virtual world and simulate real world scenarios to connect the players to situations that might arise in the real world and its consequences. The player witnesses the consequences and would also be given tools in the games that they can engage to avert the consequences in the games. The player is rewarded for right actions and penalized for the wrong ones. This process encourages the player to use the correct tools and thereby empowers them with right knowledge that can extended to the real world. The games based learning method is not only engaging but also cost effective, scalable to cater to large number of users, and customizable [4]. In the literature, there are fewer studies that have applied games based learning approach for cybersecurity awareness. The study presented is aimed at addressing this research gap.

The games based learning method has several advantages. The obvious advantage is that games based learning method provides an interactive approach to train or educate the users about a specific program. It enables the players to acquire skills and to enable thought processes in a fun and interactive way. The adaptability and flexibility of games approaches enables to

design the game to suit almost every training subject possible [8]. A well-designed game enables the user to enter a virtual environment that is similar to the real environment and would enable the player to draw connection between the learning in the virtual environment to the real world. The games based approach would motivate the player to move towards the goal with required actions and also see the consequences without facing penalties in the real life. Further, when compared to the traditional method of training, games based methods are more engaging, relatively cost effective, easily transferrable to a large number of trainees, customized to each player, etc. [9].

Game based learning is a process which uses exercises (competitive/scoring) to motivate users in learning according to specific learning objectives [10]. Usually the games involve an interesting and interactive narrative specifically designed to meet the learning objectives. Scoring is one of the major aspects of the game which is essential for developing the interest among the users. Another important aspect is GBL environment, which must be effective and encourage the users to learn and adapt [11]. Many research studies were conducted and still being conducted to analyze the impact of gaming in creating awareness about various activities in various fields. It was suggested that though GBL do not result in better performance in some instances, they don't generally result in worse performance, and may have additional benefit of a more positive attitude toward the subject of GBL. The performance in GBL can be attributed to various aspects including its environment, visual appeal, interactivity etc. [12]. It was found that using immersive environments in the game, where all the said aspects were effectively designed would result in significant improvements in the learning aspect of the users [13].

### III. IDEAS FOR CYBER SECURITY AWARENESS GAMES

A few research studies have shown that games can be used to impact security awareness [5,6,7]. This study approaches the games based learning in a manner similar to brain-training apps. The games in this study are designed to engage the players in short activities frequently, eventually leading to enhanced cybersecurity awareness. To develop the games for training, 12 potential game concepts were shortlisted at the conceptual stage. These games were based on the security aspects that the end-users would likely encounter, and the concepts which they need to be familiar with. An overview of these games is presented in this section. The process involved outlining the security learning/awareness objective, key elements of the gameplay, and a storyboard level design of the game

#### A. Vulnerability Patching

The main aim of this puzzle game is to educate the players about fixing vulnerabilities. In this game, different vulnerabilities are presented on the screen as cracks, which will spread over a period of time. Player needs to stop it from spreading by taking appropriate steps, which involves in applying appropriate patch to fix it. Various cracks keep on

appearing on the screen and the player needs to act quickly because the more time he takes to fix one patch new ones will start to appear even more rapidly. As the time progresses the difficulty increase in terms of more complicated vulnerabilities will start to appear on the screen, which takes more time to solve. The goal for the player here is to fix the crack in shortest period of time by following the correct procedure to apply the patch. Throughout the game the player will acquire new techniques (patches), which will help him to fix the vulnerability more quickly.

#### B. Leak Data Game

This is an action and tactic based game outlined with an objective of training the users about the importance of data privacy and security and intends to train them about being aware of the type of data being collected from their devices. The aim of this game is to make the users aware about the importance of data privacy and the type of information being collected from their devices from the Internet. The idea behind the game is to represent the data transfer between the user device and the Internet in the form of network packets. The type of data being transferred in the network packet will be illustrated to the user. For instance, the data being carried from the network packet maybe something like user location, banking details, user's personal information, pictures, etc. The user needs to decide which popping network packets needs to be allowed to the Internet.

#### C. Backup Cloud

This is a puzzle game outlined with an objective of creating awareness among players about the importance of taking backup of their files. The focus of this game is to train the players about the importance of data backup. In the game, the user will be provided various scenarios in which they are required to store the given files according to the file type into their appropriate folders in the cloud. As the game proceeds the difficulty level of the game will be increased. In the next level, the number of file types will be increased to sixteen types but the number of folders in the cloud will be same, i.e. eight folders in the cloud.

#### D. Phishing Email

This game aims at enabling the users to identify phishing emails from legitimate emails and take appropriate action such as deleting the phishing email. This is an interactive action type game where the user is posed with a list of emails on the screen with different types of email contents. The email list scrolls on the screen with different content types. The user identifies if it is a phishing email and drags it to the left to bin it and drags it to the right if it is a safe email.

#### E. Cybersecurity Helpdesk

This is a puzzle game, where the player would be enacting the role of a cybersecurity support worker. On the game interface, is a computer network with different types of computing devices connected to the cyber space. The

computing devices might be facing a cybersecurity issue and is indicated on the screen by the presence of a “happy” or “sad” emoticon on the computing device. The user needs to identify the “sad” device and investigate the cybersecurity issue the device is facing. In the game interface, first there will be a network with computers and the player identifies the computer with an issue and selects it. After this step, the issue faced by the device will be shown on the screen. The user needs to select the right solution for the issue. As the game progress, the number of devices and issues would be increasing.

#### F. Anti-virus

This is an action game with complex graphics in which malware attacks are posed to the player that needs to be stopped by the users using anti-virus update as ammunition. This is an action game where the player will encounter malware attacks. The setup will be of the user travelling on the screen in different directions where he will encounter malware during the journey that occurs in the cyber space. The user will be posed with malwares and viruses which the user needs to destroy to reach their destination. The user will be provided with ammunition which they have to use to kill the malware or the viruses. The game will be a multi-level where the difficulty in each level increases with respect to time and the number of attacks. The game will be fun because the attackers would be dynamic and be able to avoid the shooting of the player. Hence, the user needs to be careful while choosing the target and destroying them. The game maybe made complicated by having non-threatening objects such as files or required software which the user should not shoot. If the user shoots such file he will lose points as well as part of their ammunition.

#### G. Network Tunnel

This is an action game where the user, with his devices, travels through a simulated computer network connection and encounters threats along the way in the network. The user will be an animated person and the player will be able to control the movement of the animated person through the up, left, right, and down arrow keys on the keyboard. In the simulated network there will be security issues which will be faced by the player such as virus or malware issue and the user then needs to choose right solutions to fix the issue using solutions such as anti-virus, firewall, etc. A particular situation the user will encounter while travelling through the network is getting connected to different Wi-Fi networks. When the user is connected to a public Wi-Fi, then the user is required to choose right solutions such as VPN and firewalls to increase their security against cyber threats.

#### H. Security Incidents

This game is about training and raising awareness amongst the users about incident reporting practices by challenging the user to take right measures when faced with a cybersecurity incident. This is a puzzle based game that aims at educating players about cybersecurity incident reporting practices. In the game the user will be posed with a cybersecurity incident on the game screen to which the user needs to determine the

appropriate incident reporting practice. For a given incident shown on the screen, the user will be given four choices of incident reporting measure. The user needs to gauge the incident shown and choose an appropriate option to report the cybersecurity incident. There will be a time restriction on the game of 30 seconds within which the user needs to decide the right option.

#### I. Social Media

This game aims at enabling the users to decide what type of posts are appropriate to post on social media. This game would create awareness about the users to be cautious while using social media. Updating personal information on social media is very common these days and it is important that the users are cautious while posting such information. In this game, the users will be posed with different examples of social media posts and they are asked to decide if posting such posts is safe or not. As shown in the above figure, on the player's screen, an example social media post will be shown. The user needs to decide whether this post is appropriate to be posted on social media or not. The user swipes right to accept the example post or left to reject it.

#### J. Encryption

In this game, the user learns the importance of encrypting the files using encryption methods. The game trains the user about encrypting files which is an important security practice against cyber threats. This is a puzzle and action type of game. On the user screen, there will be a set of different files from which the user identifies the files that are flagged with an alarm notification that needs encryption. There will be an alarm icon on the screen that indicates that there is a file on the screen that needs encryption. The user needs to find the important file and drop it in the encryption box. An alarm will flash on the top of the screen when there is an important file in the page. The user then needs to find the important file and drop it into the encryption box. Once all the files are encrypted, the screen will move on to next page and then beep an alarm if there is a file that needs to be encrypted. For instance, files such as bank statements or company reports needs to be encrypted. The user needs to identify the important files within a given time constraint.

The remaining other two games are Password game and Malware game, which were selected for the design and development. These two games are selected for implementation because they form a part of most of the applications over the internet. For example, password is a common security aspect which can be related with most of the applications; whether it is social media applications, emails or any other customized applications over the internet. Similarly malwares/virus/worms etc. are major security problems that all the internet users are facing with. So, based on the larger scope for security concerns, and to create maximum security related awareness these two games are selected which have wider reach. The design aspects of these two games are discussed in detail in the next section.

IV. GAMES DESIGN

A. Password Protector

1) Intended Learning

The game aims at training the users to create complex, and strong passwords. An important and a basic measure for cybersecurity is having safe and secure passwords, and changing them frequently. This game would train the users to create complex as well as memorable passwords. It will be an important learning aspect for the user and might be able to use this game training to implement in real life.

2) Concept

This game is designed with an aim to train the players to hone their password creation skills by training them on the aspects of what constitutes a strong password. The game allows the players to practice strong password creation in a competitive context. The players are provided with a limited set of characters from which the players are required to create best possible strong, memorable passwords. The standard password strength principles of long passwords with diverse set of characters are used as guidelines to define the strength of the password. The game is timed and the password created are rated via a password meter. The game has a threshold score for the password to be acceptable. Further, the created password should not just be strong but also memorable to the players. The players are required to repeat the password, to test the repeatability of the passwords they created. Additionally, the game is time constrained. The players have to create the password in a limited time for game rewards.

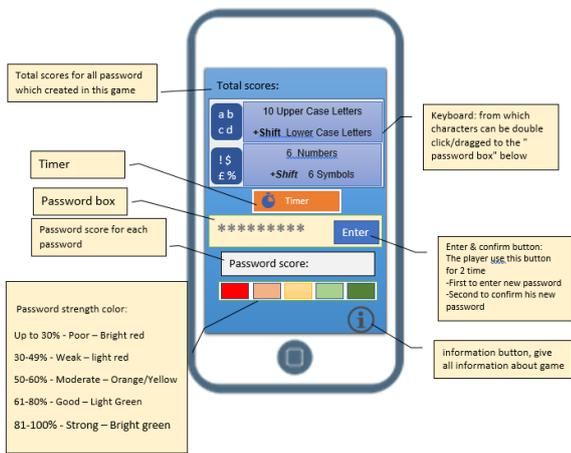


Fig 1. Password game Features

After creating the password, in the next screen, the user has to enter the password that was created. In order to get a high score, the user has to remember the password. Additionally, a timer option is provided in the game. The user can achieve high scores in creating and remembering the password in short time duration. The lesser the time is, the higher the score would be. There are currently seven levels proposed for the game, where the difficulty level increases as the game progresses to the next levels. A prototype of the game is presented in figure 2.



Fig 2. Password Game Prototype

The scoring technique is based on the complexity of the password created. The password strength factors are calculated as per the guidelines of CSCAN [14]. The complexity of the password is based on the password complexity rules that are set out in the game. In the game, the password complexity rules are inspired from password guidelines provided by IBM knowledge center and are defined based on the usage of characters from four characters, which includes the following

- Uppercase letters [A-Z]
- Lowercase letters [a-z]
- Number [0-9]
- Non alphanumeric characters (~!@#% ^&\* \_ - +=\()\{\}[]:;'"<>.,?)

The interface of the game is designed using vibrant colors and effective display of characters in order to effectively engage the players in the game.

3) Game Levels

Each level has different set of conditions that are needed to be fulfilled in order to move to the next level. In each level, the number of characters to be used for creating the password is different from other levels. The complexity level remains same at all levels, however, the user need to create strong password based on different combinations to score more. The available time decreases as the game proceeds to next levels. Level 1 (90 seconds), Level 2 (70 seconds), Level 3 (50 seconds), Level 4 (40 seconds). Different combinations that users may use include numbers and letters; mixed case letters; letters, numbers and Symbols; and mixed case letters, numbers and Symbols. Based on the combination and the length of the password, the score is determined in each level.

B. Malware Guardian

1) Intended Learning

This game aims at training the players with skills to understand and identify malicious software and creates awareness about the importance of using up-to-date anti-

malware protection systems and taking backups as a safety measure. In the game, the players have the task of defending their system by roleplaying as an anti-malware package that scans file as it approaches a computer device. The player blocks the files that are malicious before it reaches the computer device. The game includes different types of malware such as viruses, trojans, spyware, ransomware that can infect the computer devices in different ways. Only clean and uninfected files must reach the computer device in the game and the player is rewarded/penalized based on the number of malware reaches the device. The player tasks are also combined in parallel with the task of updating the malware software and taking regular backups of the device.

2) Concept

The malware game will be designed and developed with an aim of educating the players about various types of malwares, the harm they would cause, and the necessary techniques of destroying them before they attack the system. A narrative approach is adopted in the game, with action sequences to make it attractive and to create interest among the players. The concept of game is based on destroying the various malwares that would try to attack the personal computer. The malwares are represented as bugs in the game. The malwares can move towards from multiple directions towards the computer on the screen. The player has to select the right method to destroyed different malwares. The methods describe the techniques to be used to destroy different types of malwares in reality. Therefore every malware can have a different or a common method that can be used to destroy them. Additionally, there is back-up option for users in game, which they need to select in the event of an attack. An update bar on the screen, which notifies fully update or update available. The player needs to regularly update the anti-malware to destroy the new threats.



Fig 3. Malware Game Prototype

The score is calculated depending on the number of bugs destroy by the players. For example, if the player destroys all the five bugs, score would be 50, if he destroys 4; score would be 40, and so on. However, the time can be fixed to create more sporting spirit among the players. The game is a multi-level game, where the complexity and the toughness increase as the game progress to next levels. Depending on the number of malwares fixed, the player gets the score, and ends the game. The prototype model of the game is shown in the figure 3.

3) Types of Malware

TABLE I. TYPES OF MALWARE

Malware type	Explanation	Messages from pop up in the top screen
Adware	Adware (short for advertising-supported software) is a type of malware that automatically delivers advertisements.	A message indicating that the adware malware is destroyed and another message indicating the harm this type could cause to the system.
Bot	Software program to automatically perform specific operations.	A message indicating that the bot malware is destroyed and another message indicating the harm this type could cause to the system.
Bug	A fault that creates undesired outcome	A message indicating that the bug is debugged and another message indicating the harm this type could cause to the system.
Ransomware	It holds the computer on hold until an action dictated by it is not performed	A message indicating that the ransomware is destroyed and another message indicating the harm this type could cause to the system.
Rootkit	A software used by intruders to remotely control the computer without the realization of the actual user	A message indicating that the rootkit malware is destroyed and another message indicating the harm this type could cause to the system.
Spyware	It is a malware that monitors the user data on their device	A message indicating that the spyware is destroyed and another message indicating the harm this type could cause to the system.
Trojan Horse	Trojan is a common type of malware that disguises itself as a genuine software and infects the device	A message indicating that the Trojan horse malware is destroyed and another message indicating the harm this type could cause to the system.
Virus	A virus is a bug that copies itself into multiple times in the system effecting its performance	A message indicating that the virus is destroyed A message indicating the harm this type could cause to the system.
Worm	This type of malware exploits the vulnerabilities of the computer's operating system	A message indicating that the worm is removed from the device and another message indicating the harm this type could cause to the system.

4) Game Levels

The complexity of the game increases as the levels increase. As the game proceeds to the next levels, the available time decreases, number of malware attacks increases, and the number of safe files decreases.

TABLE II. LEVELS DESCRIPTION

Level	Discripe	Final scores
First level	90 seconfd, 35 random malware attack, 25 Safe gfile, Auto-scan each 5 seconds	1 Star: If total scores=> 150 2 stars: If total scores=>250 Hihg scores: If total scores=>300
Secound level	70 seconfd, 40 random malware attack, 20 Safe gfile, Auto-scan each 7 seconds	1 Star: If total scores=> 150 2 stars: If total scores=>250 Hihg scores: If total scores=>300
Third level	60 seconfd, 45 random malware attack, 15 Safe gfile, Auto-scan each 10 seconds	1 Star: If total scores=> 200 2 stars: If total scores=>300 Hihg scores: If total scores=>400
Fourth level	50 seconfd, 50 random malware attack, 10 Safe gfile, Auto-scan each 15 seconds	1 Star: If total scores=> 250 2 stars: If total scores=>350 Hihg scores: If total scores=>450

After developing the games, the next focus is on experimental evaluation, with at least 50 participants being involved for each game. In each case, participants will be asked to complete pre-test questionnaires that assess their current knowledge and appreciation of the topic (i.e. password selection or malware protection, depending upon the game they are due to trial). They will then have a period of two weeks in which the related game app is installed on their device, and their usage of it will be tracked (i.e. counting the number of times they play it). At the end of the trial period they will be asked to complete a further survey, this time re-assessing their views on the topic area (i.e. to see if it has changed from the initial survey), as well as their feedback on the game (assessing aspects such as playability, enjoyment, and awareness-raising value).

V. EXPERIMENTAL EVALUATION AND EARLY FINDINGS

A pilot of this process has already been conducted with nine participants, in order to validate the surveys and ensure that the game apps were meaningful to a wider audience. While it is not statistically meaningful to analyze the survey results from this stage in any depth, they did at least serve to give an illustration that the games themselves were considered effective by the participants. The themes of usability, learning content and enjoyment were each investigated via 4-5 related questions, with each rated on a 5-point scale (where 1 is most negative and 5 is most positive).

TABLE III. PILOT STUDY RESULTS

	Average ratings		
	Usability	Learning content	Enjoyment
<b>Password Protector</b>	3.8	3.9	2.9
<b>Malware Guardian</b>	3.6	3.8	3.7

Given the low number of respondents, rather than looking at each question individually, Table 3 presents the averages across all questions in each category for each of the games. This gives a broad indication that the overall results in all cases was skewed towards the positive side. The notably lower ‘enjoyment’ score for the Password Protector game is perhaps to be expected, as it is a time-based memory game, whereas the apparently more enjoyable Malware Guardian is an action game and does not make the participants feel that they are being explicitly tested and rated in the same way.

VI. CONCLUSION

Awareness programs are important to address the challenges posed by rapidly changing technological systems. The paper presents design aspects of the two mobile games being developed to spread awareness on cybersecurity. The games developed accommodate two important aspects of cybersecurity: strong password creation and protection against malware. Currently, the games are in the final stages of design process. The future work would focus on developing, implementing and evaluating the games.

REFERENCES

- [1] Klahr, R., Shah, J.N, Sheriffs, P., Rossington, T., Pestell, G., Button, M. and Wang, V. 2017. *Cyber Security Breaches Survey 2017*, Main report, April 2017, Department for Culture, Media and Sport, London, UK.
- [2] Hendrix, M., Al-Sherbaz, A. & Victoria, B., 2016. Game based cyber security training: are serious games suitable for cyber security training?. *International Journal of Serious Games*, 3(1), pp. 53-61.
- [3] Boyle, S., 2011. *An Introduction to Games based learning*, s.l.: UCD Dublin.
- [4] Trybus, J., 2014. *Game-Based Learning: What it is, Why it Works, and Where it's Going*, s.l.: New Media Institute.
- [5] Denning T., Lerner A., Shostack A., and Kohno T., “Control-Alt-Hack: the design and evaluation of a card game for computer security awareness and education,” in *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, pp. 915–928, 2013.
- [6] Gondree M., Peterson Z. N., and Denning T., “Security through play,” *Secur. Priv. IEEE*, vol. 11, no. 3, pp. 64–67, 2013.
- [7] Nyeste P. G. and Mayhorn C. B., “Training Users to Counteract Phishing,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 54, pp. 1956–1960, 2010.
- [8] Boyle, S., 2011. *An Introduction to Games based learning*, s.l.: UCD Dublin.
- [9] Trybus, J., 2014. *Game-Based Learning: What it is, Why it Works, and Where it's Going*, s.l.: New Media Institute.
- [10] Teed, R. "Game-Based Learning", *Games*, 2017. [Online]. Available: <http://serc.carleton.edu/introgeo/games/index.html>. [Accessed: 24-May- 2017].
- [11] Oblinger, D. 2006., "Simulations, games, and learning," *Educause Learning Initiative*.
- [12] Ke, F. 2008., "A case study of computer gaming for math: Engaged learning from gameplay?," *Computers & Education*, vol. 51, pp. 1609-1620.
- [13] Virvou, M., et al., "Combining software games with education: Evaluation of its educational effectiveness," *Educational Technology & Society*, vol. 8, pp. 54-65, 2005.
- [14] CSCAN. (2017). "Rate your Password", Centre for Security, Communications and Network Research, University of Plymouth, Available: <https://www.cscan.org/passwordstrength/>. Last accessed 15th Oct 2017.

### **Session 3:        Wireless Networking and Communication**

Title: On Blockchain-Based Authorization Architecture for Beyond-5G Mobile Services  
(Authors: Shinsaku Kiyomoto, Anirban Basu, Mohammad Shahriar Rahman, Sushmita Ruj)

Title: Hybrid Public Key Authentication for Wireless Sensor Networks  
(Authors: Daehee Kim, Jakeun Yun, Sungjun Kim)

Title: Intelligent Safety Management System for Crowds Using Sensors  
(Author: Norah Farooqi)

Title: A Hybrid Approach for Femtocell Co-tier Interference Mitigation  
(Authors: Abdullah Alhumaidi Alotaibi, Marios C. Angelides)

# On Blockchain-Based Authorization Architecture for Beyond-5G Mobile Services

Shinsaku Kiyomoto, Anirban Basu, and Mohammad  
Shahriar Rahman  
Information Security Laboratory  
KDDI Research Inc.  
Fujimino, Japan  
kiyomoto@kddi-research.jp

Sushmita Ruj  
Computer and Communication Sciences Division  
Indian Statistical Institute  
Kolkata, India  
sush@isical.ac.in

**Abstract**—This paper proposes a new conceptual architecture for authorization of mobile services based on blockchain technologies, and presents a design of procedures for heterogeneous mobile communication services. Furthermore, an extension of the procedures is considered in order to enhance privacy protection for users. The new architecture realizes the separation of mobile communication infrastructure and billing functions and multiple use of several mobile communication services under a single contract with a billing operator.

**Keywords**-Blockchain, Mobile Communication, Authorization, Heterogeneous Network, Privacy

## I. INTRODUCTION

Recently, all communication services including email, chat, voice communication, and multimedia content broadcasting are provided on the Internet, and the number of users of legacy mobile communication services has decreased. Systems providing legacy services are not cost effective and complicated architectures including legacy services and several types of communication media (such as WiFi, WiMax, 5G, unlicensed-band communication) lead to an increase in maintenance costs and result in serious communication failures. Furthermore, the authentication and authorization mechanisms become complicated, and these have to be separately constructed for each communication medium. ID federation techniques realize transparent use of different mobile communication services; however, management for each billing function may require complicated systems. Service areas of several communication media often get overlapped, thereby contributing to a waste of communication bandwidth and wasteful investment in communication infrastructure. With flat-rate communication services, there is often unfair use between light users, who use communication services sparingly versus heavy users who use such services very frequently. In addition, IoT devices do not have huge communication demands, but may be subject to the same level of authorization and billing. Thus, a simple authorization

architecture for mobile communication services that solves the above issues is required for future communication services. Blockchain technologies offer cost-effective solutions in several service domains. In our conceptual architecture, we use a blockchain as a ledger of service use in mobile communication systems. Our architecture: (1) is able to include several types of wireless communications, (2) reduces total costs for system operations by using blockchain and shares costs with the stakeholders, (3) provides billing systems that are independent from systems organized by mobile network operators, (4) provides several mobile communication services to users, based on a single contract with their own billing operators, and (5) reduces the risk of traceability of a user's service use by mobile operators.

The paper is organized as follows: section 2 introduces related work about blockchain applications, and section 3 presents our new architecture and basic authorization procedures for mobile communication. Security and privacy considerations are explained in section 4, and our business model is discussed in section 5. Finally, we conclude this paper in section 6.

## II. RELATED WORK

Blockchains gained popularity through their use in cryptocurrencies; however, blockchains have stirred up interest in other types of applications. Several network services are considered to be blockchain-based systems [4]. Tschorsch and Scheuermann in [21] refer to applications of blockchain technologies in a decentralized domain name system [19]. Those include preventing the abuse of cloud services [18], decentralizing cloud storage [23] and anonymous, distributed messaging [22]. Kosba et al. proposed smart digital contracts which allow anonymous parties to enforce complex agreements [12]. A blockchain-based anonymous delivery system has been constructed by AlTawy et al. [1]. MIT's Enigma [25] is utilizing blockchains to build a decentralized privacy-preserving computational platform based on secure multi-party computation. Kishigami et al. proposed, in [11], a decentralized

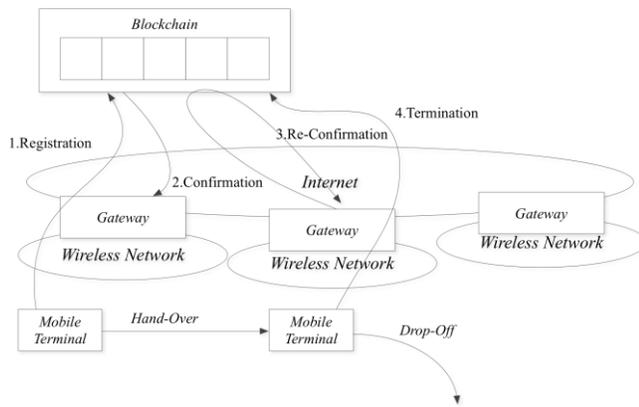


Figure 1. System Overview

blockchain-based digital content distribution system. Dennis and Owen presented a blockchain-based reputation system [8] applicable to e-commerce services. In order to verify product authenticity and ethical standards in a supply chain, [24] discusses a blockchain-based solution, and Counterparty [7] utilizes blockchain to create unalterable records of rights and transactions. A medical data management system [3] using blockchain technologies was presented by Azaria et al.. Sarr et al. discussed a blockchain-based model to handle transactions of social applications by using their access classes in [17]. Several blockchain applications to IoT environments were considered as well [16], [10]. Dorri et al. considered application of a blockchain [9] in order to improve the security and privacy of smart homes citing examples involving IoT devices.

### III. BLOCKCHAIN-BASED MOBILE SERVICE

In this section, we present our new conceptual architecture for authorization of mobile communication services. An overview of the system is shown in Fig. 1.

#### A. Entities

There are four entities in our architecture.

- Users are entities who have a mobile terminal and use mobile communication services. It is assumed that a user uses several types of mobile communication services in the beyond 5G world.
- Mobile Operators provide network services using facilities for wireless communication and gateway servers. We assume that mobile operators provide communication services including infrastructures and applications such as voice communication (VoIP) and email is independently provided on the Internet.
- Billing Operators manage billing information and have a contract with users in order to charge a communication tariff to users. The service is regarded as a collection agency service for mobile operators.
- Blockchain Service operates a blockchain on the Internet. The service is jointly managed by enterprise companies or individuals.

In this paper, we focus on Internet access services as communication services; all communication services and applications are assumed to be provided on the Internet by service providers in the future.

#### B. Basic Service Steps

The basic procedure of service use is shown schematically in Fig. 2. The procedure consists of four basic steps and one additional step as follows:

1) *Registration*: A gateway server publishes a transaction ID to a mobile terminal when a request from the mobile terminal is received, and the mobile terminal registers its service use information to a blockchain via a gateway server. The service use information includes the service status and a time stamp, and it names registration information. A user sends the registration information to a gateway server and stores it locally. Before authorization by the gateway server, the mobile terminal can only access the blockchain service. The gateway server confirms the validity of the information before sending it to the blockchain, computes its hash value and attaches the mobile operator's ID, the transaction ID, and a billing operator's ID stored in the registration information; then the tuple of the hash value, the mobile operator's ID, the transaction ID, the billing operator's ID and the service status are sent to the blockchain service.

2) *Confirmation by Blockchain*: Once a block including the registration information from the mobile terminal is confirmed by a consensus process of the block, the blockchain service sends a confirmation message to the gateway server. When the gateway server receives the confirmation message for the mobile terminal, the gateway server allows the mobile terminal to access other Internet services. The registration information is sent to a billing operator later. This step is similar to Wifi access services; thus it would be acceptable for users, even though we add the blockchain operations in the registration.

3) *Re-Confirmation*: The mobile terminal may re-connect to another gateway server when the mobile terminal moves to other areas managed by other gateway servers. In this situation, the mobile terminal sends a re-confirmation request to the blockchain service via the gateway server. The blockchain service searches previous blocks and sends a confirmation message if the service use information is found in the previous blocks. The gateway server allows the user to use the network services after receiving the re-confirmation message.

4) *Termination*: When the mobile terminal disconnects from the network service, the mobile terminal registers its service use information to a blockchain again. The service use information includes the service status as termination and a time stamp, and it names withdrawal information. A user sends the termination information to a gateway server and stores it locally. The hash value of the information is calculated and the same transaction ID, the mobile operator's

ID, and a billing operator's ID are attached to the information

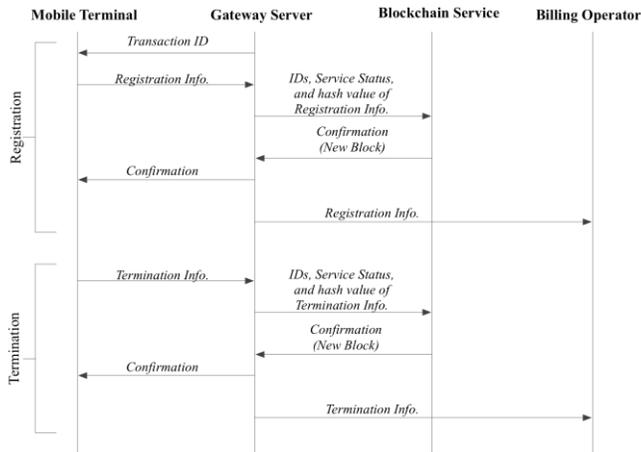


Figure 2. Basic Procedure of Registration and Termination

service status are sent to the blockchain service. After confirmation of a new block by the blockchain service, a confirmation message of the termination is sent to the mobile terminal. The withdrawal information is sent to a billing operator later.

5) *Re-Registration (Optional)*: This is an optional step and follows the same procedure as the registration step. Users may need to re-register for a certain interval in order to avoid uncertainty in relation to tariff charges. A mobile communication service is automatically terminated after a certain time period if re-registration is not executed. Even if the gateway server fails to send the withdrawal information to the blockchain, the service charge is not increased beyond the specified time period.

All entities satisfy the following conditions in the above basic scheme. All entities can be authenticated based on their own public key certificate and a corresponding private key.

- Mobile operators can calculate tariffs based on their mobile operator's ID and transaction ID by referring to the public ledger information, even though none of the users can be identified by the mobile operators from the ledger, and the user's history of service use across mobile operators is not traceable. This is because the ledger information does not include the user ID, and the transaction ID and hash values are one-time information. A billing operator's ID can be used to confirm the correct amount of a payment transaction from billing operators to mobile operators.
- Billing operators can confirm a user's service use information sent from gateway servers because the user's signature is included in the information. Hash values of the information are included in the blockchain; thus, it is impossible to alter it. The billing operators calculate tariffs for each user according to service use information including user ID and charge it to users. The total amount of tariffs for each mobile operator can be aggregated

based on the mobile operator's ID, and the total amount of tariffs for all users is transferred to each mobile operator.

- Users can verify whether the tariffs are correct based on the ledger information and service use information stored in their local device, and prove that altered service use information is incorrect because an adversary cannot compute the correct information and the information is confirmed when it is registered by the blockchain.

### C. Service Use Information and Certificate for Payment

Service use information consists of two types: registration information and withdrawal information. It denotes the usage time of each communication service and it is registered to the blockchain in order to ensure identifiability, integrity, and non-repudiability. The information consists of the ID of a user  $ID_u$ , transaction ID  $ID_m$ , time stamp  $T_s$  for a registration (or  $T_e$  for a termination), the mobile operator's ID  $ID_m$ , service status  $S = \{starting, termination\}$ , the user's public key certificate  $PKC_u$ , the digital signature of the user for these data  $SUI = (ID_m||ID_u||ID_m||S||T_s||PKC_u)$ . A hash value  $H_t$  sent to the blockchain service by a gateway server is  $H_t = h(SUI||Sig_u(SUI))$ , where  $h(x)$  is a computation by a hash function  $h$ . The time stamp protects the system and the user against replay attacks and the digital signature proves that the user generated the information for a third party. The service charge is calculated from transaction data on the blockchain using a pair comprising registration and withdrawal information which has the same transaction ID and the payment process proceeds according to the information contained in the certificate. Revocation statuses of certificates are checked on gateway servers based on certificate revocation lists or other methods. The issuer of the certificates, such as banks, provide revocation information on issued certificates; a certificate is revoked if the user concerned cannot pay the communication fees.

*Public Key Certificates*. The mobile terminal holds the private key for generating the digital signature and a corresponding public key certificate is issued to the user when the user signs a contract with a network service operator or other service providers (such as banking services) who are billing operators; they guarantee that the user will be able to pay the tariff. The certificate consists of general information including the public key certificate as an identifier of a user, issuer (billing operator), issue and expiration dates, public key, and the issuer's digital signature. Additional information that places restrictions on the user's use of a communication service is contained in the certificate as well. For example, a user may not be allowed to use a mobile communication service provided by certain mobile operators.

### D. Blockchain Service

The blockchain service holds records of service use and realizes a scalable and cost-effective platform. Ledger information received during a certain time period is recorded to a new block as a root hash. We use majority-consensus-based block-chain technologies and mobile operators cooperate to reach a consensus for each block of the blockchain; a Byzantine fault tolerant [6] scheme is a typical and reasonable

scheme for the consensus between authorized entities. Some efficient consensus techniques such as a fast Byzantine

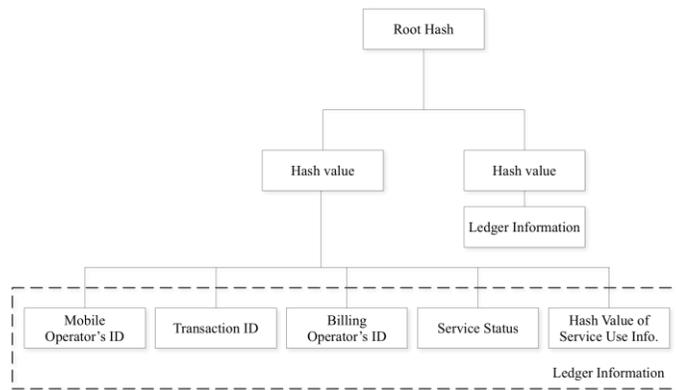


Figure 3. Ledger Information

Agreement protocol in [5] are applicable to our system, in order to improve latency of the confirmation process; a mobile operator would be selected as a signer of a new block according to a hash value of the previous block. The root hash value of the block is calculated from hash values of service use information and transaction ID by the Merkle tree hashes technique [14] as shown in Fig. 3. The hash value of service use information, service status, and transaction ID, which are recorded in an authorized block, are distributively stored in DBs of parties who participate in the consensus, as ledger information, and the DB can be accessed by mobile operators, billing operators, and users. Once the ledger information is recorded in the block of the blockchain, it has been authorized as a validated record of service use. A new block is generated for an interval such that it is generated every second; thus, users can start to receive a mobile communication service within a few intervals. If no consensus is reached, the result is notified to gateway servers; then users re-try registration or the gateway resends the withdrawal information. The restriction should be checked on the gateway servers before starting mobile services.

#### IV. SECURITY AND PRIVACY IN OUR ARCHITECTURE

In this section, we discuss the security and privacy aspects of our authorization architecture.

##### A. Security Analysis

The security analyses of our scheme are explained as follows.

1) *Invalid Use of Service*: Service use by invalid users should be prevented in order to ensure fair use of services. The digital signature of the user's private key is used as evidence of the user's consent to receive a communication service in the registration process. An attacker who has no valid key cannot use the mobile network services, and masquerading as users is difficult due to the impossibility of altering the digital signature. Replay attacks are avoided by a time-stamp and transaction ID included in data corresponding to the digital

signature. Hash values of all registration and withdrawal information are stored in the blockchain; thus, it is impossible to alter or remove the information. A certificate corresponding to the private key is issued by a party that ensures the user's ability to pay and the certificate is revoked if the user has no ability to pay; thus, the payment will be securely executed if the certificate is valid.

2) *Invalid Charge of Tariff*: Charging the correct communication tariff should be ensured to protect users against invalid mobile operators. The tariff is calculated using the transaction ID. Blockchain ensures the validity of transaction data and any entity can verify the correctness of a billing amount for the transaction ID. Thus, alteration of the billing amount is impossible. An invalid gateway may not send registration data or termination data in order to increase/decrease the total billing amount. Thus, authentication of gateway servers by users and the blockchain service is mandatory for secure transactions. Existing techniques such as TLS are applicable to the authentication procedure. Furthermore, the re-registration step avoids invalid charging by denial of termination by an invalid gateway server.

3) *DoS Attack*: Denial-of-Service attacks on the system should be considered as a potential security issue. This is a general issue and it is essentially impossible to fully protect the system against DoS attacks. Gateway servers should have established countermeasures for general DoS attacks such as filtering of traffic. Another critical part of our architecture is the blockchain service; countermeasures should be applied to the service such as parallelization of the blockchain service.

##### B. Privacy Issues

Transaction data may be sensitive, for example, it may include time of service use, network information, and billing information; and users are concerned about privacy breaches. Ledger information contains no private information including detailed billing information and no information that would identify users such as the ID of users, even though the hash value of the information is included. Thus, no privacy leakage from ledger information is expected. Mobile operators may trace a user's history using information from their own gateway servers; however, a user's history of service use across some mobile operators cannot be traced. A billing company receives service use information including information that could identify users; such information is traceable through some communication services and the time of service use is identified for each user when a user connects to several communication services. The billing operator should be trusted by users; but users may be concerned about privacy issues; the billing operator has to know details of the billing information, but time of service use is not mandatory information for billing operations. Thus, a remaining privacy requirement is to hide the time of service use from billing operators, even though tariffs are revealed to a billing operator. We will discuss an extension of our scheme that addresses the privacy issue in the next subsection.

C. Extension to Privacy-Preserving Scheme

In this subsection, we consider an extension of the scheme in order to realize a privacy-preserving scheme that reduces the

The design and implementation of a practical homomorphic signature scheme remains an open issue.

Table I. Modification of Service Use Information

1	Registration Info.	$ID_m$	$ID_m$	$ID_b$	$ID_u$	starting	0	$ID_u$	$T_s$	$PKC_u$
2	Withdrawal Info.	0	$ID_m$	0	0	0	termination	$ID_u$	$T_e$	0
1-2	Billing Info.	$ID_m$	0	$ID_b$	$ID_u$	starting	termination	0	$T_s - T_e$	$PKC_u$

risks of traceability by billing operators. There are some existing confidential transaction schemes [13], [15] for Bitcoin or other virtual coin systems; however, those schemes are not directly applicable to our architecture as it is not designed for peer-to-peer transactions. Thus, we design a privacy-preserving scheme to solve the remaining privacy issue described in the previous subsection. Intermediate nodes named Trusted Store Function are set up between the gateway server and the billing operator as shown in Fig. 4, and each intermediate node distributively stores service use information and provides a part of the service use information to billing operators. The trusted store function may be installed in the gateway servers if the gateway servers are trusted. There are several constructions for a privacy preserving scheme as follows:

- **Naive Construction.** If we can assume that the intermediate node is trusted, the node simply removes or modifies privacy sensitive information such as transaction ID and timestamps and then sends the modified information including a digital signature of the node to a billing operator on a daily basis.
- **Construction using Sanitizable Signatures.** Replacement of a user’s signature with a signature generated by a sanitizable signature [2] scheme is another possible solution for a privacy preserving scheme. Some privacy sensitive parts of service use information can be masked on the intermediate node and a user’s signature is still verifiable by a billing operator, when using a sanitizable signature scheme.
- **Construction based on a Homomorphic Signature.** We can construct a privacy-preserving scheme under the assumption that a practical homomorphic signature [20] scheme exists. We modify the original service use information as shown in Table 1. A pair of the service use information (registration and withdrawal) is transformed into concatenated bit strings, and calculate a subtraction between the Registration Information and Withdrawal Information is calculated on the trusted store function. The homomorphic digital signatures are attached to both pieces of information; thus, a digital signature for the subtracted information can be calculated. The subtracted information and its digital signature provided by a user would be transferred to a billing operator for a certain interval (e.g. daily).
- **Homomorphic Signature.** Let  $Sig(sk, x)$  and  $Sig(sk, y)$  be a digital signature of messages  $x$  and  $y$  by a private key  $sk$ . A signature scheme  $Sig(\cdot)$  is a homomorphic signature where:  $Sig(sk, x) \pm Sig(sk, y) = Sig(sk, x \pm y)$ .

To use the above schemes, the exact time of service use can be hidden from a billing operator, and the billing operator only knows the usage time used for calculating the tariff for a communication service. It is assumed that the intermediate nodes cooperatively manage trusted and distributed DBs for storing service use information, and an indeterminate node randomly selected when the information is sent to a billing operator; thus, the traceability of a user on an indeterminate node is reduced. A privacy-preserving scheme without intermediate nodes is an open issue; entities that store the service use information of users should be required to verify transactions.

V. BUSINESS MODEL CANVAS

A block in the blockchain described above is evidence of communication service use, and service operators can charge tariffs for service use according to the transaction data related to the ledger information authorized by the blockchain. Thus, a fair use communication service based on service use records is realized. The blockchain can be shared (i.e., operated) by mobile network operators in order to reduce the total cost of the system, and the communication services include several communication media, such as WiFi, WiMax and 5G. Our system allows for the case where a mobile network operator is not identical to the billing operator that charges the tariff. Mobile operators can receive service charges from the billing operators that sign the contracts with users. Users are only required to have a single contract with their appropriate billing operators such as banking services. In another way, we can use a public blockchain with volunteer miners such as the Bitcoin; in that case, contributors to a blockchain consensus obtain incentives based on the charges from mobile users. For example, if it is assumed that 50 million users pay 50 GBP per month and new block is generated every second, a miner obtains 10 % commission as  $(50,000,000 \times 50) / 2592000 \times 0.1 = 10$  GBP for each block created, even though we should consider parallel blockchains to solve scalability issues. In this model, a value of a digital coin (like Bitcoin) is ensured by communication tariffs from users. The coin is transferable to real money according to the commission rate. Thus, the value of the coin is endorsed by commission fees for real services in this new business model.

VI. CONCLUSION

In this paper, we presented a new conceptual architecture for future mobile services. It realized a transparent authorization mechanism for several types of mobile communication services using blockchain technologies. The architecture leads mobile

services to a new business model that separates a billing functionality from mobile network management by mobile operators.

#### REFERENCES

- [1] Riham AlTawy, Muhammad ElSheikh, Amr M. Youssef, and Guang Gong. Lelantos: A blockchain-based anonymous physical delivery system. *Cryptology ePrint Archive*, Report 2017/465, 2017. <http://eprint.iacr.org/2017/465>.
- [2] Giuseppe Ateniese, Daniel H. Chou, Breno de Medeiros, and Gene Tsudik. Sanitizable Signatures, pages 159–177. 2005.
- [3] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman. Medrec: Using blockchain for medical data access and permission management. In 2016 2nd International Conference on Open and Big Data (OBD), pages 25–30, 2016.
- [4] N. Bozic, G. Pujolle, and S. Secci. A tutorial on blockchain and applications to secure network control-planes. In 2016 3rd Smart Cloud Networks Systems (SCNS), pages 1–8, 2016.
- [5] Jing Chen and Silvio Micali. ALGORAND: the efficient and democratic ledger. *CoRR*, abs/1607.01341, 2016.
- [6] Allen Clement, Edmund Wong, Lorenzo Alvisi, Mike Dahlin, and Mirco Marchetti. Making byzantine fault tolerant systems tolerate byzantine faults. In the 6th USENIX Symposium on Networked Systems Design and Implementation, NSDI'09, pages 153–168, 2009.
- [7] Counterparty Foundation Community. Counterparty: <http://counterparty.io/>, 2014.
- [8] R. Dennis and G. Owen. Rep on the block: A next generation reputation system based on the blockchain. In 2015 10th International Conference for Internet Technology and Secured Transactions (ICITST), pages 131–138, 2015.
- [9] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram. Blockchain for IoT security and privacy: The case study of a smart home. In 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pages 618–623, 2017.
- [10] S. Huh, S. Cho, and S. Kim. Managing IoT devices using blockchain platform. In 2017 19th International Conference on Advanced Communication Technology (ICACT), pages 464–467, 2017.
- [11] J. Kishigami, S. Fujimura, H. Watanabe, A. Nakadaira, and A. Akutsu. The blockchain-based digital content distribution system. In IEEE BDCloud 2015, pages 187–190, 2015.
- [12] Ahmed E. Kosba, Andrew Miller, Elaine Shi, Zikai Wen, and Charalampos Papamanthou. Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In 2016 IEEE Symposium on Security and Privacy, pages 839–858, 2016.
- [13] Denis Lukianov. Compact confidential transactions for bitcoin, 2015.
- [14] Ralph C. Merkle. A certified digital signature. In *Proceedings on Advances in Cryptology, CRYPTO '89*, pages 218–238, 1989.
- [15] Tim Ruffing and Pedro Moreno-Sanchez. Mixing confidential transactions: Comprehensive transaction privacy for bitcoin. *Cryptology ePrint Archive*, Report 2017/238, 2017. <http://eprint.iacr.org/2017/238>.
- [16] M. Samaniego and R. Deters. Blockchain as a service for IoT. In 2016 IEEE International Conference on Internet of Things (iThings), pages 433–436, 2016.
- [17] Idrissa Sarr, Hubert Naacke, and Ibrahima Gueye. Blockchain-based model for social transactions processing. In 4th International Conference on Data Management Technologies and Applications, pages 309–315, 2015.
- [18] Jakob Szefer and Ruby B Lee. Bitdeposit: Deterring attacks and abuses of cloud computing services through economic measures. In *IEEE/ACM CCGrid 2013*, pages 630–635. IEEE, 2013.
- [19] Daniel Kraft et al. Namecoin: <https://namecoin.info/>, 2014. [20] Giulia Traverso, Denise Demirel, and Johannes Buchmann. Homomorphic signature schemes - a survey -. *SpringerBriefs in Computer Science*, 2016.
- [21] Florian Tschorsch and Bjoern Scheuermann. Bitcoin and beyond: A technical survey on decentralized digital currencies. *IACR Cryptology ePrint Archive*, 2015:464, 2015.
- [22] Jonathan Warren. Bitmessage: A peer-to-peer message authentication and delivery system. white paper (27 November 2012), <https://bitmessage.org/bitmessage.pdf>, 2012.
- [23] Shawn Wilkinson and Jim Lowry. Metadisk: A blockchain-based decentralized file storage application. 2014.
- [24] Reid Williams. How bitcoin's technology could make supply chains more transparent, 2015.
- [25] Guy Zyskind, Oz Nathan, and Alex Pentland. Enigma: Decentralized computation platform with guaranteed privacy. *arXiv preprint arXiv:1506.03471*, 2015.

# Hybrid Public Key Authentication for Wireless Sensor Networks

Daehee Kim

Department of Internet of Things  
 Soonchunhyang University  
 Asan, Korea  
 daeheekim@sch.ac.kr

Jakeun Yun, and Sungjun Kim

Department of Electrical and Electronic Engineering  
 Korea University  
 Seoul, Korea  
 {jkyoon, sjunii}@dsys.korea.ac.kr

**Abstract**—In this paper, we propose a hybrid public key authentication scheme for wireless sensor networks whose goal is to minimize energy consumption while providing public key authentication. To accomplish this goal, we employ the certificate based public key authentication scheme and the Merkle hash tree based public key authentication scheme depending on the number of sensor nodes. The simulation result shows that our proposed scheme outperforms the existing public key authentication scheme in terms of energy consumption

**Keywords**—hybrid public key authentication; digital certificates; Merkle hash tree; heterogeneous wireless sensor networks

## I. INTRODUCTION

Employing public key cryptography (PKC) in wireless sensor networks (WSNs) has been regarded as impractical due to its computation and communication overhead [1]. However, recent study has proved that PKC can be employed in WSNs by using elliptic curve cryptography (ECC) [2]. PKC is a best solution for establishing secret keys and authenticating broadcast messages in WSNs. One prerequisite for using PKC is to obtain a public key of the corresponding node securely. If the binding between the corresponding node and its public key is not trusted, PKC is not safe any more because an attacker can impersonate the corresponding node. Using digital certificates signed by the certificate authority (CA) is the most widely used

way to bind the identity of a node and its public key. However, the size of the digital certificate is at least 86 bytes when ECDSA-160 is used, thereby increasing the energy consumption in resource-constrained sensor nodes. To resolve this problem, [3] proposed Merkle hash tree based public key authentication scheme. It has the benefit of performing only one PKC operation by employing a Merkle hash tree, but it requires the auxiliary authentication information (AAI) to authenticate the public key, which increases energy consumption when the number of sensor nodes increases. In this paper, we propose a hybrid public key authentication scheme (HPKA) where we alternatively use the certificate based public key authentication scheme (CPKA) and the Merkle hash tree based public key authentication scheme (MPKA) according to the number of sensor nodes to reduce the energy consumption of the entire WSN.

## II. HYBRID PUBLIC KEY AUTHENTICATION

We assume a heterogeneous WSN which consists of a small number of powerful sensor nodes and a large number of resource-constrained sensor nodes. The heterogeneous WSN is assumed to form multiple clusters at deployment where each powerful sensor node plays a role of a cluster head (CH). Fig. 1 shows the example of the heterogeneous WSN with three clusters. To authenticate the base station (BS), every sensor node is

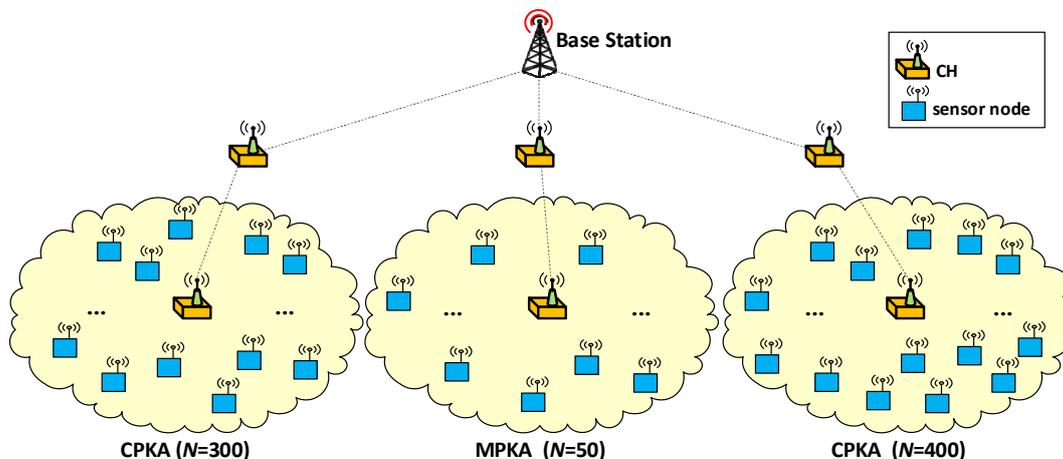


Figure 1. The hybrid public key authentication in the heterogeneous WSN.

TABLE I. THE VALUES OF PARAMETERS

Parameter	Value	Parameter	Value
$ cert_{ID} $	86 bytes	$E_{byte}$	59.2 $\mu$ J
$ h $	20 bytes	$E_{PKC}$	5.508 mJ
$ PK_{ID} $	20 bytes	$E_h$	1.89 $\mu$ J

equipped with the public key of the BS.

The goal of our proposed hybrid public key authentication scheme is to minimize the energy consumption by using CPKA and MPKA adaptively while providing public key authentication. In CPKA, each sensor node sends a message with the following format.

$$\langle ts, ID, m, sig_{ID}\{h(ts|ID|m)\}, cert_{ID} \rangle$$

where  $ts$  is a timestamp,  $ID$  is an  $ID$  of a sensor node,  $m$  is a message,  $sig_{ID}$  is a digital signature signed by the sensor node  $ID$ ,  $h$  is a hash function, and  $cert_{ID}$  is a certificate of sensor node  $ID$  signed by the BS. Upon receiving the message, a sensor node first verifies the certificate with the public key of the BS and obtains the public key of the sending node, with which the sensor node then verifies the message. Thus, each sensor node performs two PKC operations in CPKA.

In our proposed scheme, each CH constructs its own Merkle hash tree of the cluster whose leaf is  $h(ID|PK_{ID})$ . The value of the internal node is computed by hashing its two children, and in this way the value of the root node,  $h_{root}$ , is finally obtained.  $h_{root}$  is then broadcasted to all sensor nodes in the cluster signed by the CH. In MPKA, each sensor node sends a message with the following format.

$$\langle ts, ID, m, sig_{ID}\{h(ts|ID|m)\}, PK_{ID}, AAI_{ID} \rangle$$

where  $PK_{ID}$  is a public key of the sensor node  $ID$  and  $AAI_{ID}$  is an AAI of the sensor node  $ID$  which is used to authenticate the public key using the Merkle hash tree. When receiving the message, each sensor node verifies the public key of the sending node with  $AAI_{ID}$  and  $h_{root}$ . During the verification process, a chain of hash operations with  $h(ID|PK_{ID})$  and  $AAI_{ID}$  is performed. If the final value is equal to  $h_{root}$ , the public key is verified as an authentic one, with which the sensor node authenticates the message. It is important to note that the size of  $AAI_{ID}$  increases by  $(|h| * \log_2 N)$  where  $|h|$  is the size of the hash value, when the number of sensor nodes,  $N$ , grows.

In summary, CPKA has a communication overhead of  $|cert_{ID}|$  and a computation overhead of two PKC operations whereas MPKA has a communication overhead of  $|PK_{ID}|+|AAI_{ID}|$  and a computation overhead of  $\log_2 N$  hash operations and one PKC operation. Thus, CPKA consumes fixed energy regardless of the number of sensor nodes while MPKA

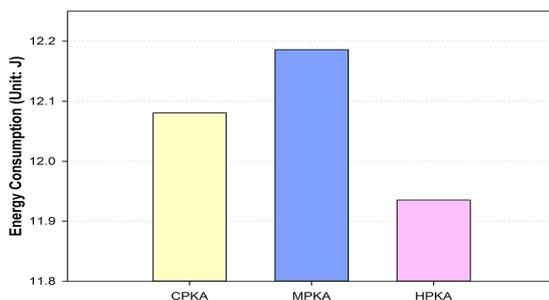


Figure 2. The energy consumption of CPKA, MPKA, and HPKA.

consumes more energy with larger  $N$  since  $|AAI_{ID}|$  and the number of hash operations depend on  $N$ . We now calculate the threshold of  $N$ , with which each CH selects the appropriate public key authentication scheme between CPKA and MPKA, as follows.

$$\underbrace{|cert_{ID}| \cdot E_{byte} + 2E_{PKC}}_{\text{Energy consumption of CPKA}} < \underbrace{(|PK_{ID}| + |AAI_{ID}|) \cdot E_{byte} + E_{PKC} + E_h \cdot \log_2 N}_{\text{Energy consumption of MPSA}} \quad (1)$$

$$\text{where } |AAI_{ID}| = |h| \times \log_2 N$$

where  $|\cdot|$  is the size of each field,  $E_{byte}$  is the energy consumed for transmitting and receiving one byte,  $E_{PKC}$  is the energy consumed for performing one PKC operation, and  $E_h$  is the energy consumed for performing one hash operation. We ignore the energy consumed for distributing  $AAI_{ID}$  and  $h_{root}$  in MPKA since it is performed only once at deployment.

To give an example of our proposed scheme, we assume that SHA-1 and ECDSA-160 is used for authentication, and TelosB motes are used and thus the value of each parameter is shown in Table I [4]. Using (1) and Table I, we can compute the minimum  $N$  to satisfy (1), which is 245. This implies that each CH employs MPKA when the number of sensor nodes in the cluster is less than or equal to 245, whereas CPKA is used in the cluster when the number of sensor nodes exceeds 245. By adapting the public key authentication scheme to the number of sensor nodes, our proposed scheme significantly reduces the energy consumption in the whole network. To compare our proposed scheme with CPKA and MPKA in terms of energy consumption, we assume that the WSN consists of three clusters, each of which has 300, 50 and 400 sensor nodes respectively and each sensor node sends one message to the neighbor node. Fig. 2 shows that our proposed scheme outperforms CPKA and MPKA.

### III. CONCLUSION

In this paper, we proposed a hybrid public key authentication scheme for heterogeneous WSNs where each cluster employs either CPKA or HPKA depending on the number of sensor nodes in the cluster. We also showed that our proposed scheme outperforms other schemes in terms of energy consumption through a simple simulation. In our future work, we need to implement the proposed scheme in the real sensor motes after which it should be evaluated in the real environments.

### REFERENCES

- [1] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," in *Proceedings of the 9th CCS*, 2002, pp. 41-47.
- [2] Liu and P. Ning, "TinyECC: A configurable library for elliptic curve cryptography in wireless sensor networks," in *Proceedings of the 7th IPSN*, 2008, pp. 245-256.
- [3] K. Ren, S. Yu, W. Lou, and Y. Zhang, "Multi-user broadcast authentication in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, pp. 4554-4564, 2009.
- [4] D. Kim, D. Kim, and S. An, "Source authentication for code dissemination supporting dynamic packet size in wireless sensor networks", *Sensors*, MDPI, vol. 16, no. 7, pp. 1063:1-22. July 2016.

# Intelligent Safety Management System for Crowds Using Sensors

Norah Farooqi

College of Computer and Information Systems  
Umm Al-Qura University  
Makkah, Saudi Arabia  
nsfarooqi@uqu.edu.sa

**Abstract**— Crowd management is an important research area due to its impact on significant numbers of people in society. It has been studied from different perspectives including movement, tracking, monitoring, identification, and behaviour. However, this topic still suffers from limitations that require further developments. Safety concerns are a vital part of crowd management that necessitates more investigation. The proposed paper presents a safety management system for crowds using sensors to improve the safety level and to reduce dangerous situations. The proposed system handles external environmental safety issues and internal body safety effects. The system manages safety factors and evaluates situations as individuals' behaviours and group behaviour in the same location in order to identify hazards and avoid disaster in crowds.

**Keywords**- crowd management; crowd behaviour; safety management ; intelligent systems; wireless sensors

## I. INTRODUCTION

A crowd consists of a massive number of people who occupy the same area to participate in specific tasks in a similar manner. Such groups may form for different purposes according to various situations. One complex type of crowd is related to religious events during which millions of people want to engage in the same task in the same place at the same time, such as Hajj in Makah. Unfortunately, this may lead to disaster, with the potential for thousands of pilgrims to be killed or injured. Crowds are also found in sports stadiums for events such as at the World Cup and other mass events. Managing the crowd is a quite challenging task that requires a significant investment in money and effort [1,2]. Consequently, crowd management is an active and important research area that demands additional investigation.

Hundreds of research studies focusing on crowd management have developed different approaches to address this topic. The majority of them have studied the movement of crowds, tracking people, navigation, and monitoring. Many studies present algorithms of computer vision in crowd management including image and video processing [1,3]. In addition, some recent studies proposed improved crowd management by applying network concepts including RFID, wireless sensors, and mobile sensing. Researchers applied RFID for crowd management as an inexpensive and easy solution [1,2,4,5]. Both studies in [1, 2] proposed identification

management systems using RFID technology to monitor crowds and find lost people. The system connects to a database module to find the medical records for those who need help. The authors in [6] used a wireless sensor network and RFID for monitoring and evacuation techniques in crowds. In [5], a location-based service for use in crowds was evaluated that depends on RFID and smartphone. Some recent studies have taken advantage of the speed of mobile phones and applied mobile sensing methodology to conduct sensing tasks and to share data in crowds [4,7,8]. Mobile sensing can be used to improve the management of crowds through tracking, detection, counting, and monitoring behaviour [7]. The authors in [8] applied mobile sensing to analyse crowd behaviour for safety assessment. However, this study only considered specific behaviours of movement such as queuing, walking, and group formation. The study in [3] handled the behaviour aspect using image and video processing to detect atypical situations such as escape scenarios. Other studies have dealt with crowd management from different perspectives including big data analytics [9], cloud-computing services [10], and the Internet of Things [11].

To date, the literature on crowd management has focused mainly on dealing with the directions and movement of crowds. It has generally addressed the safety issues in crowd management and usually handled these potentially fatal problems by traditional methods of storing and tracking medical records. However, one of the most important points in managing crowds is providing acceptable levels of safety to protect lives from accidents. The mass disasters that have happened at religious events that killed and injured thousands of pilgrims due to safety concerns are the main motivation behind this paper. The proposed system considers basic safety issues and provides a secure environment for crowds using sensors. Safety problems occur in crowds by either external environmental factors or internal body factors.

Many smart watches and smart devices have been developed to provide some safety services and to check the movements, steps, and heartbeats of their users. However, the merging among all of these internal features for groups in nearby areas and the addition of new external environmental factors to improve the safety in crowds has not yet been applied. The paper proposes a safety management system for crowds to detect both internal and external safety issues. The

system monitors individuals' behaviours and then studies the behaviour of groups of people in the same location to detect any risky situations and to identify any dangerous actions.

The research paper is organised as follows. Section 3 highlights the contributions of the present research. Section 4 shows the structure of the proposed system. Section 5 explains the processes run and the detections rules followed in the safety management system for crowds. The main points are reviewed in Section 6.

## II. CONTRIBUTIONS

The developed system aims to:

- Monitor the safety situations in crowds.
- Predict dangerous situations before they occur and warn the nearby areas in crowds.
- Improve the safety level for people as individuals and as members of groups in crowds.
- Connect users with six sensors that are used to detect fire, air poisoning, sunstroke, and crush.
- Provide recordings to permit additional investigations after accidents.
- Be interactive by using a messaging system to receive accident calls and to send warnings to crowds.

Adding these features to the dynamic systems makes it usable for many environments of crowds. It can be used effectively during religious occasions such as Hajj and sports occasions such as World Cup and Olympic Games, or any other crowded events. Many parties may adopt it as a system to improve the safety levels in crowds, to protect lives, and to provide new safety services for their customers such as safety companies, crowd management companies, and government entities.

## III. THE PROPOSED SYSTEM

The proposed system aims to provide safe, secure, and healthy environments for crowds. It is developed to be intelligent, sensitive, and dynamic to manage several safety aspects in crowds. The system connects with different wireless sensors and consists of three main modules: storage module, control module, and presentation module. The structure of the intelligent safety management system for crowds is illustrated in Fig. 1, and additional details are provided in the following sections.

### A. Wireless Sensors

The system consists of six connected sensors that can be categorised into two groups according to their functionality: external safety sensors and internal safety sensors. The external safety sensors monitor external factors in the environment that include an outside temperature sensor, air sensor, and fire sensor. The internal safety sensors monitor internal factors in the human body and include a heart sensor, body temperature sensor, and movement sensor. The outside temperature sensor (S1) captures the temperature in the environment. If the outside

temperature is high or otherwise not normal, it sends warning data to the system in order to avoid sunstroke and fire accidents. The air sensor (S2) measures the oxygen percentage in the air to discover if there are dangerous gases or air poisoning situations. The fire sensor (S3) detects any fire situations to send data to the system. The heart sensor (S4) measures the heartbeat rate and sends alerts if there are any unexpected changes in the heart rate. The body temperature sensor (S5) detects fever to assist in any dangerous health situation. The movement sensor (S6) detects the walking, step movement, and hand movements of the user to avoid any crushes or stampedes. All of these sensors are used to study situations and predict certain dangerous cases before they occur. These sensors are located in either a watch or bracelet that is used by each individual person in the crowd.

### B. Storage Module

The first module focuses on handling all forms of required data to manage the crowd properly. Part of these data are entered previously, which include information about the individual including his or her personal information, health record, and contact information. New data can also be received from related connected sensors. The module captures live data from all individuals in the crowd. All collected data are stored in the designed database. The database contains temporary data beyond the basic data. In addition, all actions are recorded during the entire event, which is useful to determine the causes if accidents do occur. Finally, the module processes and analyses collected data from sensors and works with the control module to understand and assess the dangerous situations.

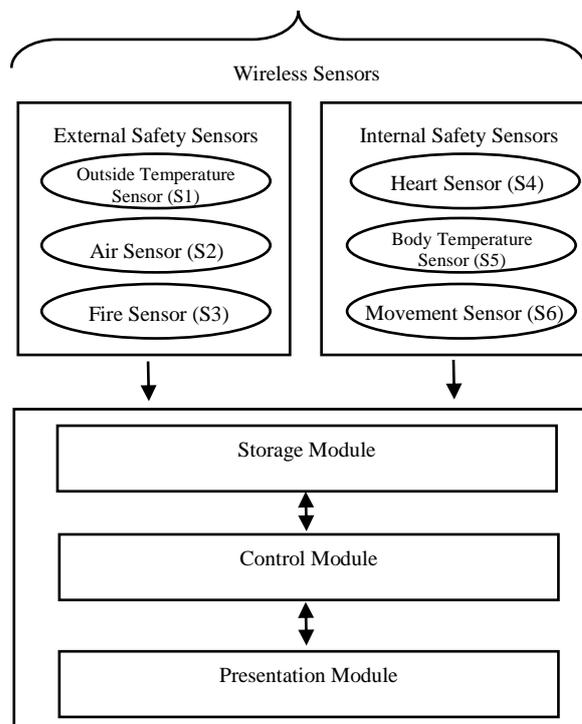


Figure 1. The structure of the developed safety management system for crowds.

### C. Control Module

The second module is responsible for checking defined rules, evaluating the processed data in the storage module, and recommending the proper decision. The control module monitors and evaluates human situations as individuals and as groups in the same area according to their locations to avoid any kind of disaster in the crowd. It studies each safety factor depending on the data received from the related sensor and stores them in the storage module. Subsequently, the module investigates the situation for a specific person and for a group of people in the same area. The control module monitors the movements of groups of people in specific areas and evaluates their heartbeats to detect any negative crowd interactions. It also measures air temperature degree and detects air purity for groups of people in the same location. Control module works as a real-time decision support system that helps people to avoid crowd-related incidents and disasters such as crushes, stampedes, sunstroke, gas poisoning, and fire.

### D. Presentation Module

The final module is associated with data visualisation. It presents smart diagrams for all data related to individuals and groups of people in the same area, leading to accurately express dangerous situations. It also connects to other emergency systems and warns related nearby areas in order to save lives during accidents. In the case that an accident has happened, it tries to limit the damage by sending alerts. Exported data from the presentation module can be used as a guide to handle dangerous situations by avoiding disaster or reducing accidents.

## IV. SYSTEM FUNCTIONALITIES

As mentioned in previous sections, the system consists of three modules that work to gather data for improving the safety levels in crowds. The section explains the system processes and highlights the general detection rules that are followed in the development of this system. The system process detects hazards and evaluates risky situations for individuals or groups in crowds via the following steps:

- The system is connected to a movement sensor and a heart sensor to measure the person's steps and heartbeats individually and then evaluate people's behaviour as a group in the same area according to their locations, to detect any crush and stampede.
- The system is connected to an outside temperature sensor, air sensor, and fire sensor to evaluate the environment for groups of people in the same place that lead to detecting the following situations: sunstroke, air poisoning, and fire.
- The system is connected to a body temperature sensor and a heart sensor to measure the person's fever and heartbeats individually and then evaluate situation to detect any health problems or side effects.
- The system records all data in the entire event like a black box in an airplane, which is useful to determine the causes if accidents do occur. The black box is used

for security purposes in order to discover hidden information.

The detection rules can be easily defined according to organisations' needs and requirements. The rules are used by the control module to evaluate the processed data and to understand whether the situation is safe or dangerous. Some basic algorithms for rules are used that can be extended to cover additional features and situations. Algorithm 1 explains the rules applied in a specific sensor for a group of people that measures the number of affected people in a specific area to make decisions. Algorithm 2 illustrates the followed rules for an individual that count the time to make a decision. Algorithm 3 shows the emerging rules between two sensors to analyse disaster situations.

```

/* Checking safety for group of people regarding
collected data from sensor S1*/

If (degree in S1 > normal degree)
{
  Loop count NumofIndividuals ++
}
Else
{
  NumofIndividuals =0
}

If (NumofIndividuals > 10 & area <= 10 meters)
{
  Decision = risky situation
}
Else
{
  Decision = safe situation
}

```

Algorithm 1. The rules for groups of people.

```

/* Checking safety for individual regarding collected
data from sensor S1*/

If (degree in S1 > normal degree)
{
  Loop Count Time
}
Else
{
  Time =0
}

If (Time > 5 minutes)
{
  Decision = risky situation
}
Else
{
  Decision = safe situation
}

```

Algorithm 2. The rules for individuals.

```

/* Checking safety for group of people regarding
collected data from sensors S1 & S2*/

If (degree in S1> normal degree & degree in S2>
normal degree)
{
  Loop count NumofIndividuals ++
}
Else
{
  NumofIndividuals =0
}

If (NumofIndividuals>10 & area<= 10 meters)
{
  Decision= disaster situation
}
Else
{
  Decision= safe situation
}

```

Algorithm 3. The rules for merging two sensors.

## V. CONCLUSION

Crowd management is a very daunting task and still requires more improvements in all related research aspects. Safety issues that occur in dense crowds can lead to humanitarian catastrophe and panic in society. The paper presents a safety management system for crowds to improve the level of safety for individuals and groups. The system is developed to detect both internal and external threats of safety using multiple sensors. Internal safety sensors relate to the reflections of the human body, and external safety sensors capture environmental concerns. The developed system consists of storage, control, and presentation modules. Future work will evaluate the proposed approach practically and test it from the perspectives of performance and cost.

## REFERENCES

- [1] A. Abidin, D. Lai and S. Chowdhary, "Advanced crowd management using ICT for just in time monitoring and real time alerts", 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, India, 2017, pp. 708-713.
- [2] M. Yamin and Y. Ades, "Crowd management with RFID and wireless technologies", First International Conference on Networks & Communications, India, 2009, pp. 439-442.
- [3] M. Szczodrak, J. Kotus, K. Kopaczewski, K. Lopatka, A. Czyzewski and H. Krawczyk, "Behavior analysis and dynamic crowd management in video surveillance system", 22nd International Workshop on Database and Expert Systems Applications, France, 2011, pp. 371-375.
- [4] R. Estrada, R. Mizouni, H. Otrok, A. Ouali and J. Bentahar, "A crowd-sensing framework for allocation of time-constrained and location-based tasks", in IEEE Transactions on Services Computing, vol. PP, no. 99, 2017, pp. 1-14.
- [5] R. O. Mitchell, H. Rashid, F. Dawood and A. AlKhalidi, "Hajj crowd management and navigation system: People tracking and location based services via integrated mobile and RFID systems", International Conference on Computer Applications Technology (ICCAT), Sousse, 2013, pp. 1-7.
- [6] S. Hashish and M. Ahmed, "Efficient wireless sensor network rings overlay for crowd management in Arafat area of Makkah", IEEE

International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), India, 2015, pp. 1-6.

- [7] M. Irfan, L. Marcenaro and L. Tokarchuk, "Crowd analysis using visual and non-visual sensors, a survey", IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, DC, USA, 2016, pp. 1249-1254.
- [8] D. Roggen, M. Wirz and G. Troster, "Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods", in Networks and Heterogeneous Media Journal, vol. 6, no. 3, 2011, pp. 521-544.
- [9] S. Awaghad, "SCEM: Smart & effective crowd management with a novel scheme of big data analytics", IEEE International Conference on Big Data (Big Data), Washington, DC, USA, 2016, pp. 2000-2003.
- [10] M. F. Ali, A. Bashar and A. Shah, "Smartcrowd: Novel Approach to Big Crowd Management Using Mobile Cloud Computing", International Conference on Cloud Computing (ICCC), Riyadh, 2015, pp. 1-4.
- [11] Y. Kawamoto, N. Yamada, H. Nishiyama, N. Kato, Y. Shimizu and Y. Zheng, "A feedback control based crowd dynamics management in IoT system", in IEEE Internet of Things Journal, vol. PP, no. 99, 2017, pp. 1-11.

# A Hybrid Approach for Femtocell Co-tier Interference Mitigation

Abdullah Alhumaidi Alotaibi and Marios C. Angelides  
 Department of Electronic and Computer Engineering  
 College of Engineering Design and Physical Sciences  
 Brunel University London  
 Uxbridge UB8 3PH  
 United Kingdom  
 {abdullah.alotaibi, marios.angelides}@brunel.ac.uk

**Abstract**— this paper presents a hybrid technique for mitigating co-tier interference between Femtocell Access Points (FAPs). This technique combines indoor deployment and cluster-based resource allocation. This requires estimation of the number of femtocell access points and their deployment locations to achieve the optimum coverage and distributing the spectrum fairly across the resulting cluster in relation to the number of users. The cluster head is selected in relation to the number of neighbouring FAPs and the distance between FAPs. Our simulation results show that both the coverage and capacity are significantly increased, the outage probability is significantly reduced, and most importantly co-tier interference is mitigated.

**Keywords** — Femtocell co-tier interference, femtocell deployment plan, clustering, resource allocation

## I. INTRODUCTION

FAPs are superior to macrocells for indoor coverage due to high penetration loss. In addition, they are self-configurable to minimise the effects of a dramatic increase in the number of users which in turn may increase coverage and decrease the probability of outage. One disadvantage of this technology is reduced communication quality because of unplanned FAP deployment [1]. Another disadvantage is co-tier interference between FAPs and cross-tier between a FAP and a macrocell [2,3] both types of which may also be significantly affected by the FAP access mode, i.e. open, closed or hybrid [4,5,6]. FAPs may use Interference Management and Radio Resource Management to mitigate interference [7] or adjust the Down Link transmission power [8]. Increasing the transmission power may maximise the FAP coverage but may also introduce co-tier or cross-tier interference [9,10,11,12]. This paper proposes a new hybrid technique which combines an indoor deployment plan and cluster-based resource allocation to mitigate co-tier interference and optimise coverage. The rest of this paper is organised as follows: section II presents related research with a focus on cluster-based resource allocation, section III presents our new mitigation technique, section IV presents our simulation results and section V concludes.

## II. RELATED WORKS

A cluster-based technique is considered as one of effective techniques to manage interference among FAPs or between FAPs and macrocell. In [13] researchers propose a scheme with which to choose a spectrum sharing mode either underlay or overlay to minimise interference with consideration of the energy and throughput of users based on femtocell clustering. Their aim is to improve the system performance and assure the power consumption is kept at a low level. Complexity and signal

overhead issues are addressed based on distributed algorithms. This study not only considers the minimisation of energy consumption but it also considers the trade-off between energy consumption and transmission efficiency. It is suggested that small cells are clustered and the one that uses the underlay spectrum sharing boosts its power at the minimum level to prevent cross-tier interference. The small cell in overlay spectrum sharing boosts its transmission power at the maximum level as it is not vulnerable to interference with the macrocell. Simulation results prove that switching between underlay and overlay enhance the system performance in terms of avoiding interference especially for small cell users. Moreover, the energy consumption is minimised. In [14] the authors propose an algorithm based on a clustering technique with consideration of guaranteeing a high QoS. Their algorithm involves three phases: forming the cluster of femtocells, allocating resources for each cluster, and resolving resource contentions in the cluster. In the first phase, each Femtocell User Equipment (FUE) sends feedback to its FAP to enable the FAP to estimate the FAPs that are vulnerable to interference. A FAP that interferes with a high number of FAPs in comparison to other FAPs is considered as the cluster head. If there is more than one cluster head, the cluster head is chosen randomly. However, if there is no cluster head for such cluster, all FAPs attached to a neighbouring cluster consider the threshold of FAP numbers in each cluster. In the second phase, the cluster head allocates resources to all FAPs based on reports received from each FAP with consideration of guaranteeing the quality of service for high priority users. After allocating the spectrum portion for high priority users, the rest of the available spectrum is set to the rest. Finally, users at the edge may suffer from severe interference because each cluster head allocates spectrum with no cooperation with other cluster heads. Therefore, interfering users send feedback reports to inform their FAP of the issue so the FAP could resolve this using a Bernoulli distribution otherwise the allocated resource is removed from the reporting FUE. The results show that the interference level is minimised and the performance of femtocell is enhanced. In [15], a scheme is proposed to minimise the complexity of resource allocation based on clustering using a coalitional game. This method is applied in three steps: femtocells are in open access mode, cooperative femtocells receive a payoff from playing the game, and resources are allocated based on weighted water falling. This method suggests that femtocells can be part of a coalition depending on the availability of subcarriers i.e. cooperative femtocells can be allocated to a subcarrier as a reward if it is assigned for public users. The simulation results show that having an open access mode leads to enhancing the total

throughput. Moreover, users of a cooperative game may exploit extra resources when femtocells draw advantages from being allocated to available subcarriers. In addition, it is proven that the total performance of the system is enhanced. In [16], an interference management based on clustering using graph-based resource is proposed. Researchers suggest that not only small cells should be clustered but also users themselves. After both small cells and users are clustered, each small cell cluster is assigned a specific value of transmission power and each user cluster is allocated to a specific portion of the spectrum. The proposed algorithms provide a solution for both co-tier and cross-tier interference. A weighted graph is used to estimate the interference among users and to alleviate it, it divides users inside cell clusters into user clusters with a low level of interference. The results show that the performance is enhanced and there is a dramatic increase in the spectral efficiency. Moreover, this method outperforms other clustering methods in terms of low complexity. In [17], researchers propose a solution that addresses the interference with consideration of the network throughput. The method used in this research consists of two steps: clustering femtocells and allocate resources to each cluster to minimise interference, and considering each cluster as resource allocation unit to use sub channels which are allocated equally to a specific level of power. Femtocells elect their cluster head based on the interference graph. Femtocells are clustered based on the interference level. Femtocells which cause high levels of interference to each other are grouped together in one cluster. To prevent co-tier interference femtocells inside each cluster should be assigned to different sub-channels. Femtocells which cause low interference to each other are also grouped together in the same cluster and they allowed to use the same sub-channels. The cluster head takes the responsibility to allocate both sub-channels and power. Femtocell users are allocated to sub-channels by estimating the difference between the experienced achievable rate and the desired rate. Simulation results illustrate that the throughput is enhanced. Both interference co-tier and cross-tier is demonstrably minimised. In [18], the authors propose a new femtocell clustering technique called semi clustering victim cell which promises to outperform the typical clustering mechanism. This technique is used for femtocells deployed close by based on identifying the victim femtocell and its aggressors to manage the frequencies among them. The same frequencies are reused among two of them if there is no possibility to interfere with each other. The authors suggest that interfering femtocells are classified into four types: victim FAPs which do not affect another FAP's users although one or more of their own users may be affected by another FAP, victim-aggressor FAPs which affect another FAP's users and their users are also affected, aggressor FAPs which affect other FAPs users but not their own and neutral FAPs which do not affect another FAP's users and their own users are safe from interference. Simulation results illustrate that co-tier interference is minimised and spectral efficiency is increased. Moreover, it is proven that this scheme has superiority over the classic femtocell clustering schemes due to its ability of increasing the capacity and utilising the resource. In [19], path loss is shared among neighbouring FAPs in LTE. Information on path loss among FAPs and the usage Component Carriers is thus modified. FAPs exchange such information using either an HeNB Femtocell gateway or an Over-The-Air method. The

HeNB gateway manages co-ordination information exchanges between FAPs and serves as intermediate node between FAPs and the mobile core network. The Over-The-Air method connects FAPs and the Mobile Base Station by a direct link. Each FAP estimates co-tier interference based on path loss and the availability of carriers. Each FAP utilises a carrier that is available for use: a carrier that is not in use by other FAPs, a carrier that is occupied by the furthest neighbour or a carrier that is occupied by the least number of neighbours. The results illustrate that co-tier interference is minimised significantly.

### III. PROPOSED SYSTEM MODEL

In our previous research, outdoor and indoor deployment plans are proposed [20, 21]. Indoor deployment plan is driven from outdoor, it is modified to fit with indoor environment and it is applied inside a building to find out its validity. The results show a significant improvement in the performance of femtocell technology in terms of the coverage, outage probability and the capacity. However, the presence of co-tier interference is still a major drawback. In this research, we propose a new hybrid technique that combines indoor deployment and cluster-based resource allocation.

#### A. Indoor deployment plan

Before start deploying the femtocells we need first to predict the required numbers. It has been shown that the optimum number of femtocells that need to be installed indoors is expressed as follows:

$$FAPn = \frac{A}{F^2} + FAPc \quad \text{where } A = L \times W$$

where FAPn is the optimum number of FAPs to be deployed, A denotes the area that requires coverage in metres, L is the length of the building, W denotes the width of the building, F is the distance threshold radius and FAPc is the centre FAP.

The indoor deployment plan involves deploying FAPc at the centre of the area and then deploying the rest of femtocells at different angles from FAPc taking into accounts two conditions: Firstly, all FAPs are deployed at a specific distance far from FAPc. Secondly, the FAPs except FAPc is deployed at a fixed angle distance from one another. The purpose of using this deployment plan is to increase the coverage area and address the probability of outage whilst increasing the capacity and offering a better service for users. The following equation explains the deployment process:

$$FAPc \rightarrow \{L/2 \mid W/2\}$$

$$\text{for all } FAP1 \dots FAPx \text{ where } x \geq 1$$

$$FAPx \rightarrow F \& (45^\circ + (G \times x - 1)) FAPc$$

where FAPc is the central FAP,  $\{L/2 \mid W/2\}$  refers to the centre of the area,  $\rightarrow$  refers to the location of installation, FAPx denotes all required FAPs to be deployed except the central FAP, F refers to the coverage radius for a FAP,  $45^\circ$  FAPc refers to a  $45^\circ$  angle from FAPc, G refers to the fixed value of the degree between FAPs deployed and x denotes the number of a FAP. G is calculated as follows:

$$G = R / FAPn - FAPc$$

where  $G$  is the distance between  $FAPx$  in degrees and  $R$  is the total FAP coverage angle set at  $360^\circ$ . The distance between  $FAPx$  is  $360^\circ$  over the number of FAPs excluding  $FAPc$ .

Figure 1 shows an example of deploying 9 FAPs inside a building based on our indoor deployment plan.  $FAPc$  is deployed at the centre of the building and then the rest of FAPs are installed at  $F$  distance from  $FAPc$  and the angle between these FAPs is equal to  $G$ .  $FAP_1$  is usually installed at a  $45^\circ$  angle from  $FAPc$ .  $FAP_2$  is installed at  $F$  distance and at  $(45^\circ + G)$  angle from  $FAPc$ .  $FAP_3$  is installed at  $F$  distance and at  $(45^\circ + (G \times 3 - 1))$  angle from  $FAPc$  and the process continues for all remaining FAPs.

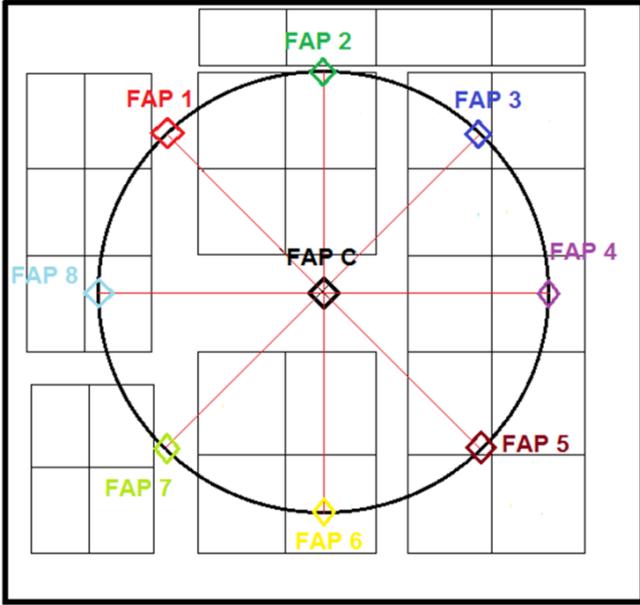


Figure 1: An example of deploying 9 FAPs

**B. Cluster-based resource allocation**

Following completion of deployment, the resulting cluster formation evolves. The next step involves the selection of the cluster head using two conditions: firstly, the number of neighbour FAPs interfered with, and secondly, the distance value to these FAPs. The candidate head is the FAP that interferes with the highest number of FAPs and is at the closest distance to these FAPs. The following algorithms explain the process of selecting the cluster head:

**iff one  $FAP_i \cap \ll FAP_n$**   
**Then  $FAP_i = CH$**   
**else iff  $\langle$  one  $FAP_i \cap \ll FAP_n$**   
**Then check  $D$**   
**iff  $FAP_i - \gg D \rightarrow \ll FAP_n$**   
**then  $FAP_i = CH$**

where  $FAP_i$  denotes the candidate FAP to be selected as CH,  $\cap$  denotes interference,  $\ll FAP_n$  denotes the highest number of FAPs,  $-\gg D \rightarrow$  denotes the closest distance to the highest number of FAPs.

Table 1 presents a case whereby  $FAPc$  and  $FAP_1$  interfere with the highest number of FAPs and the distance between  $FAPc$  and  $FAP_1$  to the rest of FAPs  $F = 40$  m. Comparing  $FAPc$  to  $FAP_1$  in terms of distance  $FAPc$  is closer to 5 FAPs ( $FAP_3, FAP_4, FAP_5, FAP_6, FAP_7$ ) and  $FAP_1$  is closer to only 2 FAPs ( $FAP_2, FAP_8$ ).

Table 1:  $FAPc$  and  $FAP_1$  distances to all FAPs

FAP <sub>n</sub>	FAP <sub>c</sub>	FAP <sub>1</sub>
FAP <sub>c</sub>	-	40 m
FAP <sub>1</sub>	40 m	-
FAP <sub>2</sub>	40 m	32 m
FAP <sub>3</sub>	40 m	53 m
FAP <sub>4</sub>	40 m	76 m
FAP <sub>5</sub>	40 m	80 m
FAP <sub>6</sub>	40 m	76 m
FAP <sub>7</sub>	40 m	53 m
FAP <sub>8</sub>	40 m	32 m

$FAPc$  is selected to be the cluster head as the device that interferes with most FAPs. However, in some circumstances, there could be a huge penetration loss between  $FAPc$  and other FAPs due to the nature of modern buildings so a cluster head may be selected from among the other FAPs in the cluster. The responsibility of the cluster head is to assign the available spectrum to these FAPs in consideration of the fairness for sharing. FAPs sense the environment so all FAPs are aware of the cluster head as each FAP generates its own report about the neighbouring FAPs and based on this report they assign themselves as follower to the cluster head.  $FAPc$  has an important role to play in dividing the spectrum fairly. Therefore,  $FAPc$  assigns a portion of the spectrum to the remaining FAPs in the cluster based on their number of users.  $FAPc$  usually assigns to itself the highest amount of spectrum for due to its location which is ideal in serving the highest number of users. The following equation explains how  $FAPc$  divides the spectrum among FAPs:

$$Sp_i = \frac{FAP_i_{UE}}{FUEs} \times BW$$

where  $Sp_i$  denotes the bandwidth assigned to each FAP,  $FAP_i_{UE}$  represents the number of users connected to  $FAP_i$ ,  $FUEs$  denotes the total number of the users that require service from their FAPs, and  $BW$  denotes the total bandwidth that is assigned by the operator. In our research we assume that the femtocells and macrocell do not use the same channels. This eliminates cross-tier interference between FAPs and macrocell.

After dividing the bandwidth between FAPs, the head reviews regularly the available spectrum and each FAP's number of users to update the bandwidth allocation among FAPs. In our research, we have given priority to fairness over equality so the bandwidth assigned to FAPs may not satisfy quality of service.

IV. SIMULATION OF THE PROPOSED HYBRID

The parameters used in the simulation are shown on table 2. We assume there is no cross-interference between FAPs and the macrocell. The cluster head candidate is usually the central FAP which is then becomes responsible for dividing the bandwidth among FAPs. However, we check to ensure that there is no alternative candidate for being cluster head. We test our proposed technique inside a building that contains 27 rooms. Despite macrocell coverage, we need to cover the whole area with FAPs.

Table 2: Simulation parameters

Parameters	Abbreviation	Values
Length of the building	<b>L</b>	80 m
Width of the building	<b>W</b>	80 m
FAP radius	<b>F</b>	40 m
Radius angle	<b>R</b>	360°
Total number of users	<b>FUEs</b>	60 users
Bandwidth	<b>BW</b>	10 MHz
Frequency	-	2.6 GHz

A. Indoor deployment plan

Firstly, we predict of the number of the FAPs that need to be deployed based on our equation from the former section:

$$FAPn = \frac{A}{40^2} + 1 \quad \text{where } A = 80 \times 80$$

$$FAPn = \frac{6400}{1600} + 1 = 5 \text{ FAPs}$$

The next step involves deploying these FAPs. The optimum coverage of the five FAPs is achieved based on this formula:

$$FAPc \rightarrow \{40/2 \mid 40/2\}$$

FAP<sub>c</sub> is deployed in the middle of the building and the rest of that FAPs are deployed at different angles to that. Before deploying the FAPs we need to calculate the factor *G* which is the fixed angle distance between FAPs:

$$G = 360^\circ / 5 - 1 = 90^\circ$$

FAP<sub>1</sub> is installed at 40m and a 45° angle from FAP<sub>c</sub>:

$$FAP1 \rightarrow 40 \text{ \& } (45^\circ + (90^\circ \times 1 - 1)) \text{ FAPc}$$

$$FAP1 \rightarrow 40 \text{ \& } (45^\circ + 0) \text{ FAPc}$$

FAP<sub>2</sub> is installed at 40m and a 135° angle from FAP<sub>c</sub>:

$$FAP2 \rightarrow 40 \text{ \& } (45^\circ + (90^\circ \times 2 - 1)) \text{ FAPc}$$

$$FAP2 \rightarrow 40 \text{ \& } (45^\circ + (90^\circ)) \text{ FAPc}$$

$$FAP2 \rightarrow 40 \text{ \& } (135^\circ) \text{ FAPc}$$

FAP<sub>3</sub> is installed at 40m and a 225° angle from FAP<sub>c</sub>:

$$FAP3 \rightarrow 40 \text{ \& } (225^\circ) \text{ FAPc}$$

FAP<sub>4</sub> is installed at 40m and a 315° angle from FAP<sub>c</sub>.

$$FAP4 \rightarrow 40 \text{ \& } (315^\circ) \text{ FAPc}$$

Figure 2 shows the deployment of all FAPs and how they are clustered. FAP<sub>c</sub> interferes with all FAPs. Therefore, FAP<sub>c</sub> is selected as a cluster head. Although FAP<sub>c</sub> is the FAP that interferes with the highest number of FAPs, in some circumstances, FAP<sub>c</sub> may compete with other FAPs to be a cluster head due to multipath. For instance, if a FAP located elsewhere inside the building increases its transmission power it may also interfere with the rest of the FAPs. Figure 2 shows the resulting cluster. Despite the high level of interference among FAPs, coverage is optimal and the outage probability is low. Moreover, the number of the users that may be served is significantly increased.

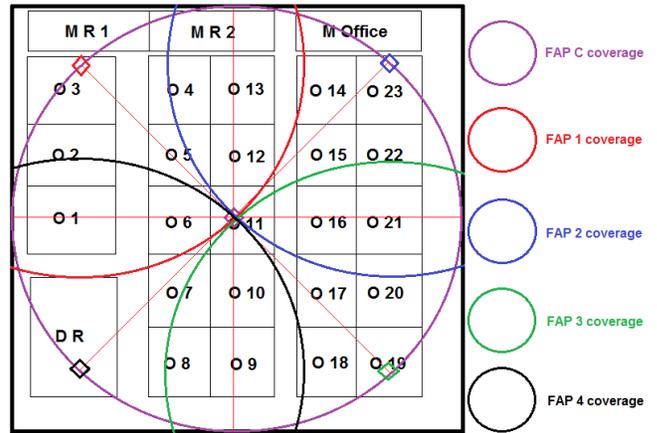


Fig. 2: FAP deployment

Figure 3 shows the coverage area divided equally among the FAP<sub>c</sub> and the remaining four FAPs. The number of users who may receive a good service simultaneously is 80 as each FAP support 16 users.

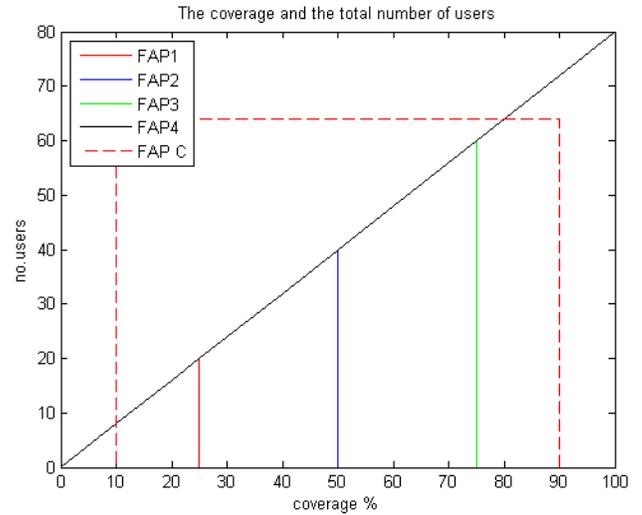


Figure 3: Coverage and the total number of users

*B. Cluster-based resource allocation*

A certain portion of frequencies are assigned to each FAP based on the number of its users. We assume the total number of users is equal to 60, BW is 10 MHz and the frequency is 2.6 GHz. We also assume that the penetration loss is very low in the upper right corner of the building so FAP<sub>2</sub> interferes with all FAPs. Before dividing the available spectrum to FAPs, we need to select the cluster head using our algorithm:

*iff*  $FAP_{c,2} \cap FAP_5$   
**Then check D**  
*iff*  $FAP_2 - \gg D \rightarrow \ll FAP_5$   
**then**  $FAP_2 = CH$   
*else iff*  $FAP_c - \gg D \rightarrow \ll FAP_5$   
**then**  $FAP_c = CH$

Table 3 includes the distance from FAP<sub>C</sub> and FAP<sub>2</sub> to the rest of the FAPs.

Table 3: FAP<sub>C</sub> and FAP<sub>2</sub> distances to all FAPs

FAP <sub>n</sub>	FAP <sub>C</sub>	FAP <sub>2</sub>
FAP <sub>C</sub>	-	40 m
FAP <sub>1</sub>	40 m	53 m
FAP <sub>2</sub>	40 m	-
FAP <sub>3</sub>	40 m	53 m
FAP <sub>4</sub>	40 m	80 m

FAP<sub>C</sub> is selected as the cluster head as it is the closest FAP to the rest. This FAP is then responsible for dividing the available spectrum among all FAPs. We assume that the number of FUEs that are connected to all FAPs is as follows:

$FAP_{cUE}$  is 16,  $FAP_{1UE}$  is 14,  $FAP_{2UE}$  12,  $FAP_{3UE}$  is 10, and  $FAP_{4UE}$  is 8. The BW among these femtocells is assigned as follows:

- FAP<sub>c</sub>  

$$Sp_c = \frac{16}{60} \times 10 \text{ MHz} = 2.67 \text{ MHz}$$
- FAP<sub>1</sub>  

$$Sp_1 = \frac{14}{60} \times 10 \text{ MHz} = 2.33 \text{ MHz}$$
- FAP<sub>2</sub>  

$$Sp_2 = \frac{12}{60} \times 10 \text{ MHz} = 2 \text{ MHz}$$
- FAP<sub>3</sub>  

$$Sp_3 = \frac{10}{60} \times 10 \text{ MHz} = 1.67 \text{ MHz}$$
- FAP<sub>4</sub>  

$$Sp_4 = \frac{8}{60} \times 10 \text{ MHz} = 1.33 \text{ MHz}$$

The bandwidth is divided into 5 frequency bands based on the number of users of each FAP. For instance, FAP<sub>c</sub> is assigned the highest value of BW as it serves the highest number of users. It is assumed that FAP<sub>c</sub> and FAP<sub>1</sub> utilise half of the BW as both of these serve half of the number of the total users who seek FAP service. 2MHz of the bandwidth is assigned for FAP<sub>2</sub> which is connected to 20% of FUEs. Finally, both FAP<sub>3</sub> and FAP<sub>4</sub> use about 3MHz of the bandwidth. Table 4 shows the FAPs and their frequencies. All users experience the same quality of service which serves our objective and co-tier interference is totally mitigated. The results show that, applying this new technique offers great advantages such as predicting the required number of FAPs instead of deploying FAPs randomly to cover the required area, achieving optimum coverage, mitigating the outage probability, increasing capacity, and sharing the spectrum fairly.

Table 4: FAPs, FUEs, and frequencies values

FAP <sub>i</sub>	FUEs	Start frequency	End frequency	BW
FAP <sub>c</sub>	16	2.600 GHz	2.6026 GHz	2.67 GHz
FAP <sub>1</sub>	14	2.6027 GHz	2.605 GHz	2.33 GHz
FAP <sub>2</sub>	12	2.6051 GHz	2.607 GHz	2.00 GHz
FAP <sub>3</sub>	10	2.6071 GHz	2.6087 GHz	1.67 GHz
FAP <sub>4</sub>	8	2.6088 GHz	2.610 GHz	1.33 GHz

V. CONCLUDING DISCUSSION

In this research, we propose a new hybrid technique which combines indoor deployment plan and cluster-based resource allocation. Our technique outperforms other clustering-based resource allocation techniques in terms of forming the cluster, selecting the cluster head, and the level of transmission power. Past research suggests that all FAPs should be grouped together in clusters without any consideration being paid to calculating the required number of FAPs and without considering the probability and level of outage. In turn, achieving both optimum coverage and high capacity is not a conscious objective. In our research, we cluster over the desired area after predicting the required number of FAPs that are necessary to achieve the optimum coverage and aiming at mitigating the probability of outage. Past research clusters FAPs separately according to the level of interference they cause or low and the spectrum to which each FAP is thereby assigned is suitably chosen. However, in our research we cluster together all FAPs required for cover over the target area. Moreover, in past research the cluster head is selected based on the number of FAPs that interfere with it without any considered approach being suggested, other than a random selection, to address the possibility of coexistence of other FAPs that may also be subject to the same level of interference by other FAPs. With our technique we consider distance in deciding which FAP should be the cluster head in those cases where there are multiple candidates. Our simulation

results suggest that our cluster head selection process is much accurate in comparison to other techniques. Naturally, there are many other factors to consider in drafting such a selection technique and that is our current focus of our research. With respect to transmission power, past research suggests adjusting transmission power levels for each FAP which in turn yields a mixed service quality and coverage. However, in our research we address the uneven levels of quality and coverage by keeping the transmission power at the same high level across all FAPs. An alternative approach that we are considering is to set thresholds for service and coverage and thereby consider small power adjustments that would yield even battery drain across all FAPs. Uneven battery drain across FAPs is a rising concern. One of the most significant challenges, however, is whether to grant macrocell users priority over femtocell users in providing service. Our initial approach is to consider all FUEs equally and not offer preferential access to macrocell users over FAP users. Finally, our technique is one of few cluster-based techniques that apply fairness among FAPs and users in terms of sharing the available spectrum, providing even service experience across users and keeping transmission power high and constant. Our simulation results suggest an improved performance of femtocell technology in terms of increasing coverage and capacity, reducing the probability of outage, and most importantly mitigating co-tier interference.

## REFERENCES

- [1] N. Kafai, "Is it a Picocell or a Femtocell?," 2013. [Online]. Available: <http://www.ubeeairwalk.com/2011/08/is-it-a-picocell-or-a-femtocell/>.
- [2] K. Rajesh, "What are Femtocells and what are their advantages and disadvantages," 2009. [Online]. Available: <http://www.excitingip.com/182/what-are-femtocells-and-what-are-their-advantages-and-disadvantages/>.
- [3] T. Zahir, K. Arshad, A. Nakata, and K. Moessner, "Interference Management in Femtocells," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 1, pp. 293–311, 2013.
- [4] G. de la Roche, A. Valcarce, D. Lopez-Perez, and J. Zhang, "Access control mechanisms for femtocells," *IEEE Commun. Mag.*, vol. 48, no. 1, pp. 33–39, 2010.
- [5] A. Khalifah, N. Akkari, and G. Aldabbagh, "Dense areas femtocell deployment: Access types and challenges," *2014 3rd Int. Conf. e-Technologies Networks Dev. ICeND 2014*, pp. 64–69, 2014.
- [6] N. Diaa El-Din, E. A. Sourour, K. G. Seddik, and I. A. Ghaleb, "Coordinated partial co-channel deployment in two-layer networks," *2013 Int. Conf. Comput. Netw. Commun. ICNC 2013*, pp. 1162–1167, 2013.
- [7] A. Zalonis, N. Dimitriou, A. Polydoros, J. Nasreddine, and P. Mahonen, "Femtocell downlink power control based on Radio Environment Maps," *IEEE Wirel. Commun. Netw. Conf. WCNC*, pp. 1224–1228, 2012.
- [8] H. Su, L. Kuang, and J. Lu, "Interference avoidance in OFDMA-based femtocell network," *Proc. - 2009 IEEE Youth Conf. Information, Comput. Telecommun. YC-ICT2009*, pp. 126–129, 2009.
- [9] V. U. Sankar and V. Sharma, "Subchannel Allocation and Power Control in Femtocells to Provide Quality of Service," in *2012 National Conference on Communications (NCC)*, 2012.
- [10] N. Saquib, E. Hossain, L. B. Le, and D. I. Kim, "Interference management in OFDMA Femtocell networks: issues and approaches," *IEEE Wirel. Commun.*, no. June, pp. 86–95, 2012.
- [11] K. Vadivukkarasi and R. Kumar, "Enhancement of indoor localization by path-loss reduction using modified rssi technique," *Int. Rev. Comput. Softw.*, vol. 9, no. 5, pp. 865–871, 2014.
- [12] K. Vadivukkarasi and R. Kumar, "A novel algorithm to improve indoor localization accuracy and path-loss reduction using real time RSSI," *Int. Rev. Comput. Softw.*, vol. 10, no. 3, pp. 332–339, 2015.
- [13] P. Mach and Z. Becvar, "Energy-Aware Dynamic Selection of Overlay and Underlay Spectrum Sharing for Cognitive Small Cells," vol. 66, no. 5, pp. 4120–4132, 2017.
- [14] A. Hatoum, R. Langar, N. Aitsaadi, R. Boutaba, and G. Pujolle, "Cluster-Based Resource Management in OFDMA Femtocell Networks With QoS Guarantees," *Veh. Technol. IEEE Trans.*, vol. 63, no. 5, pp. 2378–2391, 2014.
- [15] K. Rohoden, R. Estrada, H. Otrok, and Z. Dziong, "A Coalitional Game for Femtocell Clustering in OFDMA Macro-femtocell Networks," pp. 221–226, 2016.
- [16] L. Zhou, X. Hu, E. C. Ngai, H. Zhao, S. Wang, J. Wei, and V. C. M. Leung, "A Dynamic Graph-Based Scheduling and Interference Coordination Approach in Heterogeneous Cellular Networks," vol. 65, no. 5, pp. 3735–3748, 2016.
- [17] J. Dai and S. Wang, "Clustering-based interference management in densely deployed femtocell," vol. 2, no. April 2016, pp. 175–183, 2017.
- [18] I. Shgluof, M. Ismail, and S. Member, "Semi-Clustering of Victim-Cells Approach for Interference Management in Ultra-Dense Femtocell Networks," vol. 5, 2017.
- [19] L. Zhang, L. Yang, and T. Yang, "Cognitive interference management for LTE-A femtocells with distributed carrier selection," *IEEE Veh. Technol. Conf.*, no. Icic, pp. 0–4, 2010.
- [20] A. A. Alotaibi and M. C. Angelides, "Wireless femtocell coverage on the go: A case in the Kingdom of Saudi Arabia," in *IEEE UEMCON 2016-The 7th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference*, 2016.
- [21] A. A. Alotaibi and M. C. Angelides, "Femtocell Deployment Plan: Moving Indoors," in *Intelligent System Conference 2017*, 2017.

## **Session 4: Internet Application and Technology**

Title: Entity Identity, Performance, and Storage  
(Author: Aspen Olmsted)

Title: Automated Course Management System  
(Authors: Ashik Mostafa Alvi, Md. Faqrul Islam Shaon, Prithvi Ranjan Das, Manazir Mustafa, Mohammad Rezaul Bari)

Title: Change Management and the Integration of Information Technology: Research Notes from Selected African Universities  
(Authors: Omotayo Kayode Abatan, Manoj Maharaj)

Title: Web Service Injection Attack Detection  
(Authors: Victor Clincy, Hossain Shahriar)

Title: Mobile Business Performance Metrics: Framework and Case Study  
(Authors: Ahyoung Kim, Junwoo Lee)

# Entity Identity, Performance, and Storage

Aspen Olmsted

Department of Computer Science  
College of Charleston, Charleston, SC 29424  
olmsteda@cofc.edu

*Abstract— When a database designer needs to model an entity that exists in a domain that does not have a global name authority, the designer must resort to using surrogate identifiers. Traditionally, these entities were referred to as Weak-Entities in Entity-Relationship models. There are many choices available to a database modeler, but each choice comes with sacrifices. This paper documents an experiment that evaluates several of these choices and hypothesizes the best solution from both a performance and a storage perspective. We then apply the suggested best algorithm to a production system and discover tremendous savings in disk space requirements and execution time on certain queries.*

**Keywords-** Database Identity, Modeling Weak Entities

## I. INTRODUCTION

A natural key is a unique identifier for an object in a database that is derived from application data itself. An example of a natural key is a social security number for a student in a college database. A second example of a natural key is a vehicle registration ID to represent cars on the road. A surrogate key is a unique identifier for an object in the database that is not derived from application data. We tend to use surrogate keys when there is not a global naming authority for the entity we are representing in the database. Many algorithms have been used by database designers over the years to generate surrogate identifiers. Each algorithm has strengths and weaknesses. We explore several of these algorithms including identity, Universally Unique Identifiers (UUID) and the Hi/lo algorithm in the related work section of this paper.

This research documents our experiments with different surrogate key algorithms and their impact on the performance and the storage space requirements. The organization of this paper is as follows. Section II describes related work trying to improve the storage space requirements or performance of one of the surrogate key algorithms mentioned above. In Section III, we give an overview of our different test cases for our experiment. Section IV describes the performance and storage space results from the experiment. Section V contains the conclusion and possible future work in generating surrogate identifiers.

## II. RELATED WORK

James [1] documents a technique used by the TokudB storage engine to create binary, sequential UUIDs. Appigatla [2] documents the same technique along with adding a "bigint" primary key. In both tests, the authors are only concerned with insert performance. We use a similar algorithm in four of our five test cases and add to the work by considering both insert performance and storage space requirements.

Garcia-Molina, Ullman, and Widom [6] spend several chapters in their database textbook discussing the latency caused by the organization of data on disk. Several of our test cases try to organize the data so new inserts of tuples will be on the same data page as previous inserts to ensure that disk caching is maximized and disk head movements are minimized.

Many college-level database textbooks document the entity-relationship (ER) model. The ER model is used to communicate the design of a database with stakeholders before the development team begins to code the database. Additionally, in [6], the authors describe a Weak Entity as an entity that relies on a relationship with a Strong Entity to uniquely identify the tuples in the entity set. These Weak Entities are the entities that need a surrogate identifier.

In our previous work [7], we developed modeling techniques for expressing database constraints in distributed systems. A major limitation of traditional identity algorithms is that they do not distribute well. Our work here builds on this by providing the best surrogate indemnity algorithm for distributed systems.

Relational Database Management Systems (RDBMS) have implemented several common algorithms for handling the most column identity algorithms include:

- Identity Columns – Most commercial RDBMS products support the concept of an identity column. The identity column is an attribute on the primary key for a specific table. The RDMS service will assign the next identity on an insert of a new tuple. During the definition of the table, a starting value and increment value can be set for the identity column. Typically, the database designer can use any of the integer data types for the identity column. If the RDMS supports unsigned integers, then the use of an unsigned attribute will increase the range of the identifiers. The primary benefit of identity columns lies in the simplicity of use. Once the tables are created, the tuples are inserted without the identity value, and the server will assign the proper value. There are three primary flaws with identity columns:
  - Distribution of inserts of tuples to a single table is not possible because the next identity value is controlled by a specific RDMS server.
  - If two independent databases need to be merged after they already contain tuples, then one set of tuples will need to have their primary key reassigned. Databases typically are merged when two business conglomerate operations and attempt to combine their operations databases.

<sentence tying the two together, how one business has to lose their primary key>

- Inserts of parent records in a hierarchical relationship require retrieving the assigned value so that it can be added to the foreign key columns of children. The extra request can significantly complicate the code used for inserting new records and hold locks for longer durations reducing concurrency.
- **UUID – UUID or Globally Unique Identifiers (GUIDs)** solve the problem the three problems with identity columns. A UUID can be generated by the client before an insert on the RDMS server. The client-side creation allows for distribution and post-creation database consolidation as well as allowing the client to assign the foreign key values for children inserts from the same function. UUID are standardized [5] by The Internet Engineering Task Force (IETF). The standard defines five different types of UUIDs. The difference between the types is how the unique identifiers are generated. All different types generate a 36-character string that is made up of hexadecimal characters and dashes. The challenge with UUIDs is in their size. They are difficult for humans to communicate and remember and they take more storage space on disk.
- **Hi/Lo -** The Hi/Lo algorithm divides the sequences domain into “hi” groups. A “hi” value is assigned synchronously. Every “hi” group is given a maximum number of “lo” entries, which can be assigned off-line without worrying about concurrent duplicate entries. The only commercial implementation of the Hi/Lo algorithm is in the hibernate framework [6]. The identifiers generated by the Hi/Lo algorithm are unique only to the server that generates the “Hi” sequence.
- **Sequence –** The American National Standards Institute (ANSI) add the create sequence command to the standards for SQL:2003 [3]. The "CREATE SEQUENCE" command allows a sequence to be created with a starting value and an increment value. This sequence can then be used on any table. The challenge of distribution found with identity columns is solved because a different sequence can be used per server as long as the start values are unique.

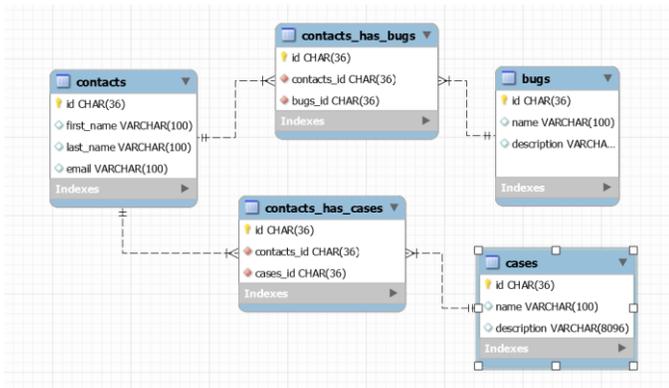


Figure 1.Test Case 1

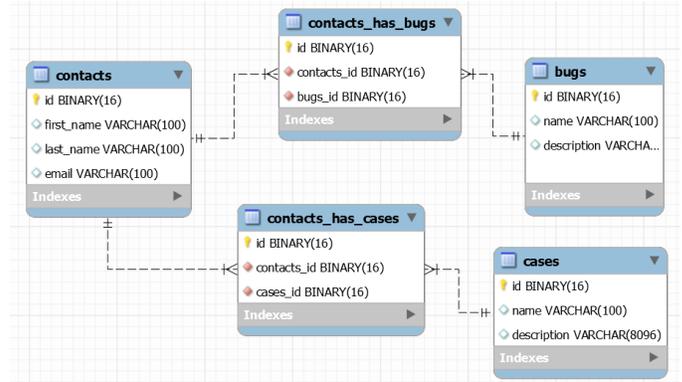


Figure 2.Test Case 2 and 3

Unfortunately, the other challenges found in identity columns also exist for sequence columns.

### III. TEST CASES

To test the impact on the different surrogate key generation algorithms, we chose a subset of the database design from the SuiteCRM [8] product. SuiteCRM is an open source customer relationship management (CRM) software solution. SuiteCRM is a fork of SugarCRM [9] commercial CRM. Both solutions use UUIDs as unique identifiers for all entity objects stored in the database. Figure 2 shows the subset of the design we use in all of our tests. Test case 1 uses the MySQL [10] RDMS as the storage tier for our testing. With MySQL, SuiteCRM uses a fixed-length character field to represent the identifier and foreign identifiers. The length of the identifiers is thirty size characters.

Test case 2 and 3 use a binary encoding of the hexadecimal string into a binary field of length 16 bytes. The binary conversion is done by stripping the dashes from the string and encoding sequential pairs of hexadecimal digits as binary bytes. Figure 2 shows the modified entity-relationship (ER) diagram. The difference between the two cases is in the generation of the UUID. Test case 2 uses a standard type 1 UUID converted to a binary value. Algorithm 1 shows how the binary value is created for test case 2. Test case 3 modifies the type 1 UUID to shift the time segment to the beginning of the string. This shift ensures sequential UUIDs will be increasing in value. Increasing inserts will be on the same data page or sequential data pages instead of randomly distributed. Algorithm 2 shows how the binary sequential value is created. Writing data to the

#### Algorithm 1 Generate Binary UUID

```

INPUT:
OUTPUT: @Buuid Binary(16)

SET @UUID = uuid();
SET @Buuid = UNHEX(@UUID);
RETURN @Buuid
    
```

**Algorithm 2 Generate Binary Sequential UUID**

**INPUT:**

**OUTPUT:** @BSuuid Binary(16)

```

SET @UUID = uuid();
SET @BSuuid = CONCAT(SUBSTR(@uuid, 15, 4),
    SUBSTR(@uuid, 10, 4),SUBSTR(@uuid, 1, 8),
    SUBSTR(@uuid, 20, 4),SUBSTR(@uuid, 25));
SET @BSuuid = UNHEX(@BSuuid);
RETURN @BSuuid
    
```

same page will be faster because the data page can be cached and the hard disk write head does not need to move around the disk drive because the data will be sequential in the write operation.

Test case 4 and 5 use a binary encoding of the hexadecimal string into a binary field of length 16 bytes as a unique identifier but insert a 4 byte unsigned integer as the primary key. The primary key is used to reduce the size of the foreign keys and also to ensure that writes are added to the end of the data file. The identity column will allow the data to be cached and reduce the disk head movement latency as our modified version of the UUID did.

**IV. PERFORMANCE RESULTS**

We ran two test sizes: small and large. In the small test, 1 thousand tuples were inserted into the leaf node of the design. These include the contacts, bugs, and cases tables. For the large test, 10 thousand records were inserted into the leaf table. The joining tables randomly were assigned three records for each parent records. So, in the small test, the joining tables had 3 thousand records inserted, and in the large test, they had 30 thousand records inserted. Figure 5 shows the cumulative timing results. In each experiment, case 1, where we used the type 1 UUID consisting of a 36-character string, was the slowest. Forcing the generated identifiers to be sequential reduces the execution time of certain queries. For the join tables, there was a significant difference with the time required to insert tuples with a string identify representation vs. a binary representation.

**Algorithm 3 Test Data Generation**

**INPUT:**

**OUTPUT:**

```

Loop through leaf nodes in one-to-many relationships
Loop 1000 times
Replicate data in table replacing surrogate ID
Loop through all tables where current table is one side
Replicate many records using new surrogate ID
    
```

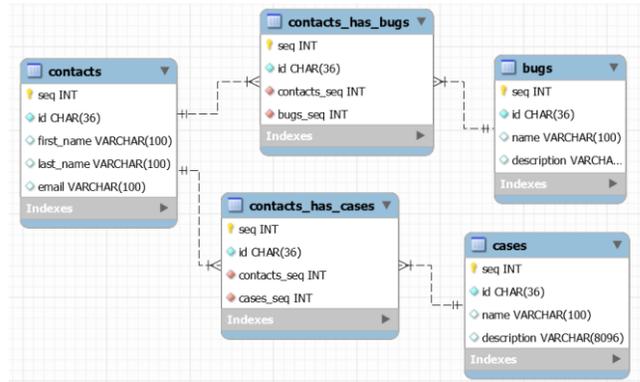


Figure 3. Test Case 4 and 5

**V. STORAGE RESULTS**

To compare the storage space required by each surrogate identity algorithm, we aggregated the storage requirements for the data inserted in both the small and large test runs. Figure 4 shows the data requirements for each of the five test cases. Test case 4 uses the least amount of data because of the use of the non-public primary key as the foreign key for the join tables. The challenge with the solution of using a non-public identity primary keys is that merging of data becomes difficult and the distribution of writes among several servers requires allocations of identities. In our testing, we also did not set a unique constraint on the public identifier. The absence of the unique constraint saved disk space by not creating an index. In reality, if the public identity is used for searches, the indexes would still be required; and the storage space savings would be depleted. Test case 3 is the next best algorithm and only has a 13 percent storage penalty over the use of non-public primary keys. The storage benefit of Algorithm 2 over Algorithm 1 comes from the increasing order of the keys generated that allows the data to be stored more compactly.

**VI. DISTRIBUTION**

Figure 5 shows the time required to insert records on a single server. The last 2 cases use a server generated identity, and the first 3 test cases use a globally unique identifier. The use of the globally unique identifier means the first three tests can be scaled linearly by allowing inserts to a single table over multiple servers. There are many reasons to distribute writes to multiple servers simultaneously, including partition tolerance, performance, geographic distribution, and sharding. If an application needs to support a partition from the client network to the server network, a local copy is often used for reads and writes, and the data is synchronized with the server when a partition does not exist. In some application domain spaces, millions of inserts happen in a short time span. The heavy insert pattern is often the case with streaming data. In this scenario, it is better to write to many servers and later synchronize the data to a central repository. Like the scenario with partition intolerant applications, geographic distribution is often used when offices are geographically dispersed. For example, if the latency may be too high for an application that is run from both

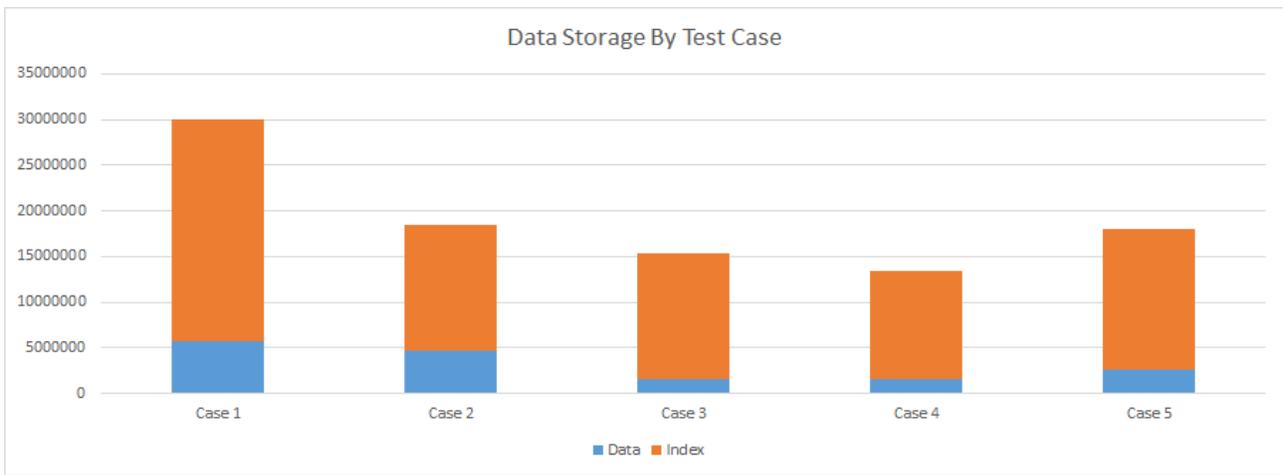


Figure 4.Storage Requirements

an office in Europe and an office in California to share a single MySQL database. In this case, each office would have a separate copy for local write operations that would later be replicated to each instance. The last use case where distribution is important is with sharding. Sharding horizontally partitions the database across multiple servers. There are tools such as SpockProxy [11] that sit between the client application and a set of MySQL servers. Based on a partitioning function, the SpockProxy will send the requests to different servers. Sharding is important in use cases with many concurrent users and large datasets that will not fit in memory.

### VII. PRODUCTION TEST

All 5 test cases above used a very small set of the tables that represented an abstraction from the data models used by a SuiteCRM or SugarCRM production system. In a real production system, there are many more tables, relationships, and search indices. As we saw in our test systems, the usage of

the type 1 UUID as a primary key causes a large amount of index space to be used. We examined a production system used by a large performing arts center and modified the source code to use Algorithm 2 in a test environment. The system was developed on top of the SugarCRM platform with many tables specifically used to store data for the performing arts market [12]. We saw a storage space reduction of 36 percent. The vertical market solution, built on top of the SugarCRM system, had only a handful of users. We wondered how our solution would impact the horizontal market in general. SugarCRM reports their solution allows has more than 2 million users in over 120 countries [13]. To experiment on a system with a database like a typical customer’s production system, we replicated the demonstration data that is installed with a clean installation of SugarCRM. Algorithm 3 shows how we generated a 20-gigabyte database based on the small sample data provided with a new installation. The test generation algorithm reads the foreign key relationships from the metadata and then walks through the tables on the one side of the one-to-

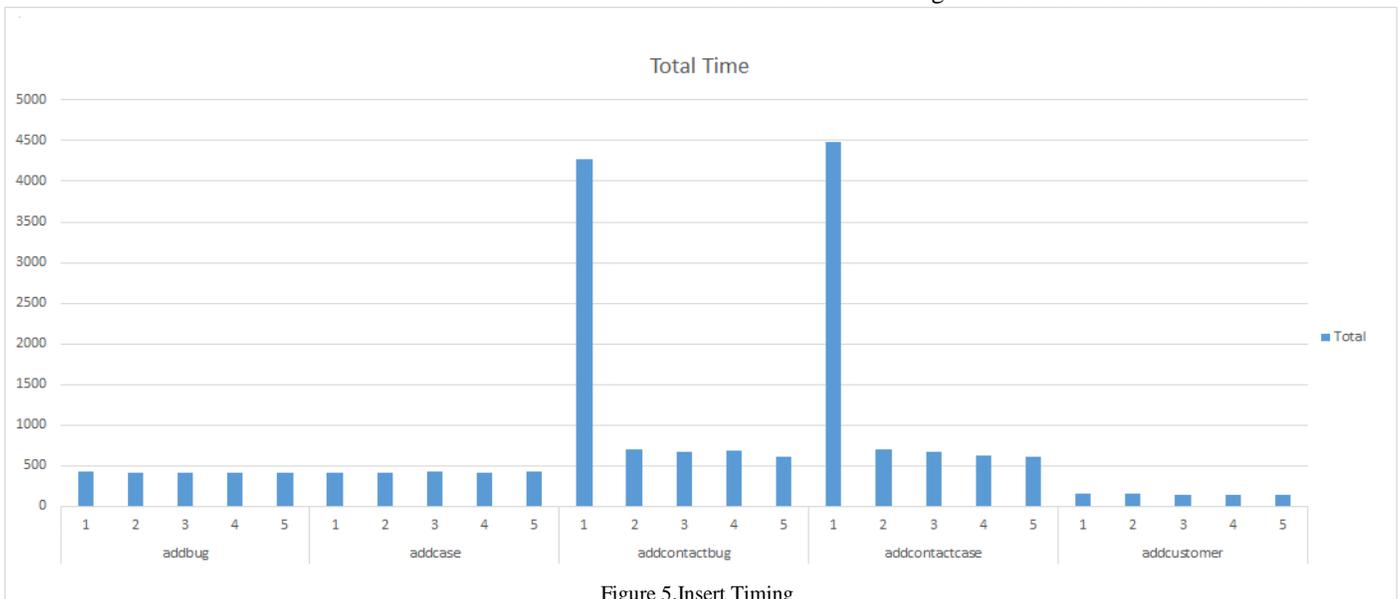


Figure 5.Insert Timing

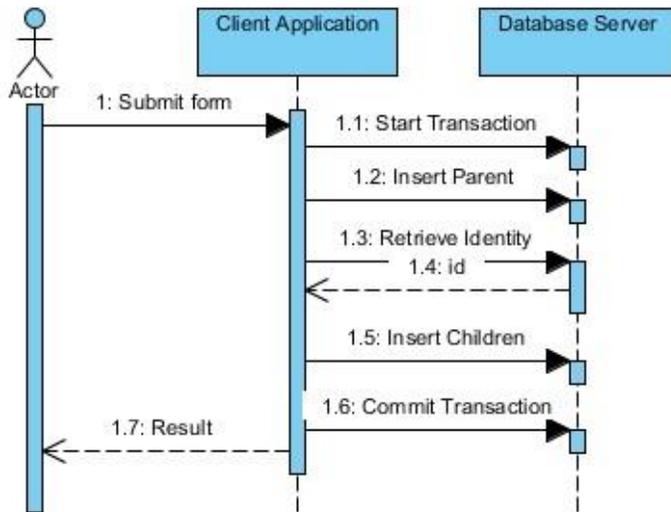


Figure 6. Sequence Diagram of Server Identity Generation

many relation. For each tuple in the original test data, 1 thousand new tuples are generated with new surrogate identifiers. For each of these tuples, the algorithm then walks all the many records that are related and generates new records with the new identifier as the parent record. This new dataset allowed us to measure the storage space requirements difference from Algorithm 1 and Algorithm 2. Algorithm 2 stored the same dataset using 59 percent less space. To test performance improvement, we measured the table scan time of the contacts table when searching on a non-index field. The compactness of the data allows for the query to execute with a 62 percent reduction in time. The scan test was performed when no data has been read from the database to prevent machine caching from influencing our results.

#### VIII. SERVICES GRANULARITY AND IDENTITY

We have already discussed the challenge with the distribution of inserts across multiple database servers with identity columns. Identity columns also have another natural impediment to concurrency. In traditional database programming, a client application would issue an insert command of a parent record, followed by a separate command to insert records for the related child records. With identity values assigned by the server, an additional round trip is required between the insert of the parent and the insert of the children to retrieve the identity assigned to the parent tuple. Figure 6 shows the communication flow between a client application and a database server using fine-grained service calls. The main challenge with this architecture is that locks will be held on shared resources for the duration of the transactions. The latency of the many round trips lowers the concurrency of the application. If the client could generate the ID before inserting the parent, then the latency could be reduced, thus allowing higher concurrency.

#### IX. CONCLUSIONS AND FUTURE WORKS

Based on our research, we believe the use of Algorithm 2 with a design similar to the ER model shown in Figure 2 provides the maximum performance and storage optimizations along with several other benefits that come with using UUIDs as surrogate identifiers. This solution provides the ability to partition a single table across multiple machines to increase the performance of insert operations or to reduce the latency from geographic distribution. We applied our work to a larger data implementation that was similar to the data of millions use in their production implementations of SugarCRM [9]. With the larger dataset size, we saw a tremendous improvement in both the storage space requirements and the performance of some queries. We also showed theoretical, additional improvements and gained with fine-grained service architectures. Future work will consider implications of identity in distributed hashes used for caching.

#### REFERENCES

- [1] R. James, "GUID/UUID Performance Breakthrough," 09 2016. [Online]. Available: <http://mysql.rjweb.org/doc.php/uuid>. [Accessed 01 April 2017].
- [2] K. Appigatla, "Store UUID in an optimized way," 19 12 2014. [Online]. Available: <https://www.percona.com/blog/2014/12/19/store-uuid-optimized-way/>. [Accessed 01 April 2017].
- [3] H. Garcia-Molina, J. Ullman and J. Widom, Database Systems: The Complete Book, Pearson, 2008.
- [4] A. Olmsted and R. Stalvey, "Service Constraint Guarantees," *International Journal of Intelligent Computing Research*, vol. 5, no. 1, pp. 430-437, 2014.
- [5] The Internet Engineering Task Force (IETF), "A Universally Unique Identifier (UUID) URN Namespace," 07 2005. [Online]. Available: <https://tools.ietf.org/html/rfc4122>. [Accessed 01 April 2017].
- [6] RedHat Inc, "org.hibernate.id Class SequenceHiLoGenerator," 2010. [Online]. Available: <https://docs.jboss.org/hibernate/orm/3.5/api/org/hibernate/id/SequenceHiLoGenerator.html>. [Accessed 01 April 2017].
- [7] A. a. M. J. a. K. K. a. M. J.-E. a. Z. F. Eisenberg, "SQL:2003 Has Been Published," *SIGMOD Rec*, vol. 33, no. 1, pp. 119--126, 2014.
- [8] SalesAgility, "SuiteCRM – Open Source CRM for the world," 2016. [Online]. Available: <https://suitecrm.com/>. [Accessed 01 April 2017].
- [9] SugarCRM Inc, "Transformative CRM. More sales and less force.," 2017. [Online]. Available: <https://www.sugarcrm.com/>. [Accessed 01 April 2017].
- [10] Oracle Corporation, "The Worlds Most Popular Open Source Database," 2017. [Online]. Available: <https://www.mysql.com/>. [Accessed 01 April 2017].
- [11] Spock, Inc, 2016. [Online]. Available: <http://spockproxy.sourceforge.net/>. [Accessed 10 April 2017].
- [12] A. Olmsted, "Building Vertical Market Applications with SugarCRM," *php/architect*, vol. October, 2007.

- [13] SugarCRM Inc., "All we do is CRM. It's in our name.," [Online]. Available: <https://www.sugarcrm.com/about>. [Accessed 23 April 2017].

# Automated Course Management System

Ashik Mostafa Alvi, Md. Faqrul Islam Shaon, Prithvi Ranjan Das, Manazir Mustafa, Mohammad Rezaul Bari

Department of Electrical and Computer Engineering  
North South University

Plot – 15, Block – B, Bashundhara, Dhaka 1229, Bangladesh

{ashik.mostafa, shaon.islam, prithvi.das, manazir.mustafa, rezaul.bari}@northsouth.edu

*Abstract*— Recent developments in computer networks and communications have brought a great change in the E-learning and course management systems. This change has brought the teachers and the learners closer than they were ever before. The traditional teaching methods have been getting replaced by the new technologies and methods. In E-learning, in particular, the course management system plays an important role. As the availability of internet has increased, people have become even more dependent on it. In that context of development, the course management system should be designed in such a way so that the users experience less complexity while using the same. In our work, a hassle-free technique to design a course management system where all the facilities and services needed by the stakeholders are accessible at one place, has been proposed. Stating from the classroom creation to manage and maintain the classroom have been made even more easier and secured.

*Keywords* - online course management system; e-learning; learning management system; online classroom; course content management

## I. INTRODUCTION

In recent years, the rapid development of Computer Aided Instruction (CAI) has been playing an important role in modern teaching management system. Traditional teaching methods (e.g., face-to-face teaching) is not providing enough support for the learners to engage towards the course. Consequently, many online social networking sites and some other course management sites that are based on the platform of WWW applications are getting popular among students and teachers [1]. Social networking sites such as Facebook, Google+, Twitter, etc., in addition to hosting marketing, personal sharing, and/or social events have started to host E-learning services. There are over 1.86 billion active Facebook users according to the report of monthly active users in Facebook till January 2, 2017 [2]. As these social networking sites allow the users to do multiple activities other than E-learning, these sites are not an ideal place for the teachers and students to interact with each other for academic purpose.

In 2013, Student Experience and Expectation of Technology stated in its survey that the twenty-first century students maintain a significant digital engagement. In that survey, it was found that 96% of the students had access to a laptop or a desktop computer at home, and 82% had access to a smartphone [3]. These data stipulate that students utilize these digital devices in almost every aspect of their lives including their learning. It, therefore, is a better and safer undertaking in general to design and develop an online course management system for all educational institutions. Most of the top

universities are using their own course management system. It is not a good idea to share the university data (e.g., student and instructor information) to a third-party web application for a course management service. An institution has its own user (e.g., students, teachers, employees) data and those are private and exclusive in nature. Using these data, an institution can develop and maintain its own online course management system for its users.

An existing online education system in general exhibits integrated network sharing of teaching resources, teaching resource management, video courses, and teacher-student interaction. As of today several virtual learning models have been designed and implemented in different environments [4-8], and these models are being constantly upgraded and improved when used in real-world situations. A course management system is easy to use and is equipped with appropriate teaching aids, and thus, to a large extent, caters for the requirements of both the instructors and students in actual learning environment. For example, it has the capability to introduce syllabus, update bulletin board, preset courses, search for available data and information using search engine, provide abundant reference materials, traditional textbooks and multimedia data, offer simple editing facilities, emerging tools for exhibition and exchange such as built-in blog, and real-time communication tools such as online communication, message board, and BBS, so as to facilitate timely communication between the instructors and students. In addition, it can support teaching assistance modules such as exercise module, homework module, download module, etc. However, the existing models reveal some of the following difficulties:

- An instructor must create a classroom first and generate a corresponding unique classroom code,
- All of the students of a particular class of a particular instructor are required to be enrolled to the classroom using student email address or an unique id, and
- Students need to use their registered email addresses to login and use that unique classroom code to join the class.

To address the above difficulties, a scheme to pursue automation of the classroom creation part has been proposed. Teaching methods, interactions between instructors and students, and course management methodologies vary university to university. In this paper, a design of an online automated course management system (ACMS) is developed for North South University, the first private university in Bangladesh. Most of the private universities have open credit

system for taking courses. The system using which a student takes courses from among all the offered courses in a semester is known as the Advising System. A proposal has been introduced by the authors as a novel hassle-free way to design a course management system which will be integrated with the Student Portal and Teacher Portal accounts so that all the services can be found at one place. The system's structure is based on Windows, Apache, MySQL, and PHP (XAMPP). In addition, jQuery, Ajax, and JavaScript are also used in the front end.

In brief, the major contributions of this work towards developing a ACMS are:

- Introduction of an automated classroom creation methodology
- Integration of all essential classroom management modules
- Establishment of easier communication between instructors and students

The remaining of the paper has been arranged in four distinct sections. Section II briefly discusses about the related previous works in this area. Section III showcases the proposed design in details. Section IV presents the details of implementation of the proposed work and Section V presents the final thoughts on the topic and the suggestions for future work in the related field.

## II. RELATED WORKS

Learning and Management System (LMS) is a software application that enable lecturers and system administrator to carry out administration, documentation, tracking, reporting and delivery of E-learning courses [9]. In some literatures, interaction is also described as an essential function that must be provided in LMS [10-12]. In the area of online social networking, some researchers shift their focus from traditional e-learning to social networking based e-learning, due to the belief that it can provide better interaction [13-15].

In a recent work, online learning and course management system have been designed using social networking sites, i.e., Facebook. The teachers and the students interact directly through Facebook without login or registration as the authors used the Facebook API. Users do not need to provide their personal information as the biography is pulled from Facebook [1].

Moodle and Blackboard are two totally different learning management systems when compared to the course management approaches [16]. The full form of Moodle, i.e., Modular Object-Oriented Dynamic Learning Environment, indicates that the course management approach is modeled, and is dynamic and object-oriented. The course manager can add/remove the model of learning activities, which makes it is a very flexible learning management system. Compared to Moodle, most of the functions in Blackboard system are not that flexible [17]. Moodle is a General Public License (GPL) open source learning platform whereas Blackboard is a business package system.

Engrade is an online gradebook and record keeper that allows teachers to manage their classes online as well as post grades, assignments, attendance, and upcoming homework for parents and students to see. The Engrade suite provides a gradebook that automatically calculates grades and provides tools for custom grading scales and weighing assignments, an attendance book that automatically emails parents with absences, a homework calendar for students and parents, and online reports where students can view their grades, homework and attendance in real time [18]. Engrade was free until 2016 when it became a paid service.

## III. DESIGN OF ACMS

The database that has been used, was acquired from the Information and Technology Department of North South University (NSU). Only the Advising, Student, Teacher, Course table's schema diagram have been used to prepare the Advising Data.

No individual sign-up system has been kept in the proposed system. Since all the active teachers and enrolled students have their student and teacher portal accounts, no further process step is introduced for classroom creation. The proposed system is designed to integrate the existing students' and teachers' portals of North South University. The concept of one account for all services has been implemented, like Google uses for its services.

An Admin panel has been incorporated for creating new courses, and adding new teachers and students. Admin Panel is responsible for giving the new teachers and students their username and password. Fig. 1 shows the basic system structure.

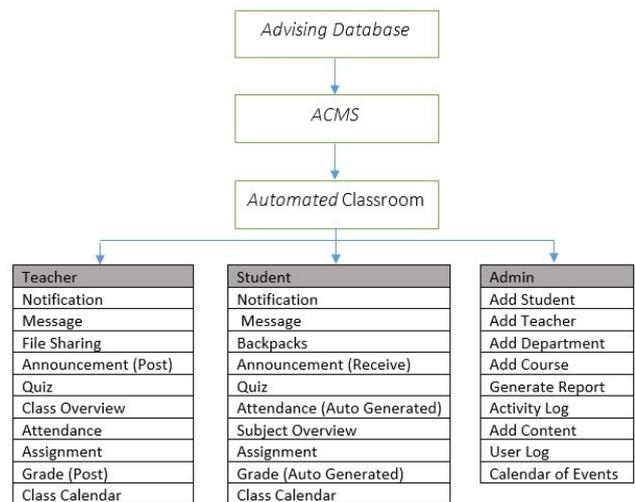


Figure 1. System design flow diagram.

### A. The Database

The system is designed based on the Advising Table as one of the major objective was to make the classroom management smoother. At the beginning of a new semester, each student advises his or her own courses by himself of herself. After this advising process, these data are saved to Advising Table of the database. In the developed system, it is named as

teacher\_class\_student table. Fig. 2 shows the attributes of the teacher\_class\_student table and also the constrains.

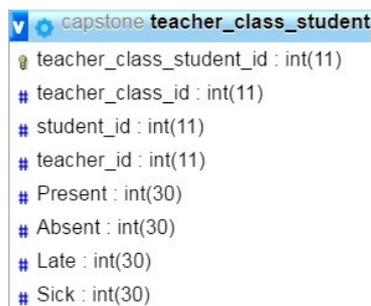


Figure 2. A snapshot of the teacher\_class\_student table's structure.

The aforementioned teacher\_class\_student table is created by the joining three other tables. They are: teacher\_class, student and teacher tables. The structure of these three tables are shown in Fig. 3.

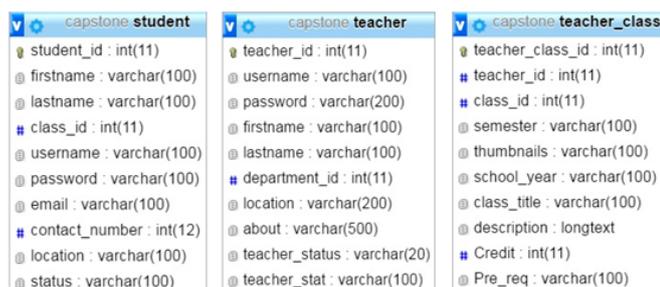


Figure 3. A snapshot of the student, teacher, teacher\_class table's structure.

### B. Classroom Creation

The after advising data have been used to make the classroom creation automated. From the aforesaid table (teacher\_class\_student, student and class table), an SQL query is run to generate classrooms. Fig. 4 is the SQL query for creating the classroom. Each classroom has a dynamic classroom page. This is the unique part of the proposed work which gives the users more relaxation in classroom creation.

```
SELECT * FROM teacher_class_student
LEFT JOIN student ON student.student_id = teacher_class_student.student_id
INNER JOIN class ON class.class_id = student.class_id
where teacher_class_id = '$get_id' order by lastname
```

Figure 4. A snapshot of the SQL query that creates the classroom.

No individual sign up is required to enroll in a classroom as a student or to create any classroom. Teachers will view and manage their classrooms after the student's advising is done. A teacher will be the admin for their each of the courses for posting assignments, grades, announcement etc. After login, teachers have notification, message, backpack, add downloadable, announcement, assignment and quiz module in their home page.

From the student's end, students will have access to their classroom after they complete their advising for the course.

Without entering a classroom students can have the access to notifications, messages, backpacks etc. modules.

### C. Core Modules

#### a) Notification Module

This module is common in both teacher's and student's end. When a teacher posts an announcement, assignment, quiz or share any file all the students within the class get notified. In the student's end, when a student submits an assignment, or a quiz, the teacher of the corresponding class gets the notification.

#### b) Message Module

A teacher can send message to any class or any individual student. Communication between active teachers is also possible. Student can send message to his or her classmates and teachers only.

#### c) File Sharing Module

Teachers have only the privilege to share any type of documents (.pdf, .doc, .ppt etc.) and media file (.mp4, .flv, .avi, .wmv etc.). Only students within the class can view and download those materials. This module can be accessed within the classroom and in the teacher's homepage.

#### d) Announcement Module

Only teachers can post announcements and students can only view them. A teacher can send same announcement to multiple classrooms if the announcement is sent from the home page. If the announcement is sent within a classroom, then the announcement is visible to only that class students.

#### e) Quiz Module

Teachers can post same quizzes to multiple classrooms if the quiz is created from the home page. For creating a quiz, teachers need to set the quiz title, marks, duration. Teachers are to set questions and answers for each question. Students quizzes count down will start from the time when they open the quiz question.

#### f) Subject Overview Module

This module is available within each classroom. This module is for the detailed information about each course. Teachers have multiple options to provide the detailed information about a course. They can directly share a course outline file or they can type the course information. Student can view or download the information about the course that from there, they will have the information about the course.

#### g) Attendance Module

This module is available within each classroom. Four attendance options have been included in the present system (namely, present, absent, late, and sick). These four options ensure the details information about each student. These options have

impacts in the Grading Module too. Attendance is saved in the database on the current daytime. Student can only view their attendance. And there is a count system that counts the number of times each student is absent or present or late etc.

h) Assignment Module

This module is available within each classroom and also in the homepage of the teachers. While posting an assignment to any class a teacher has to provide the deadline, marks, title, and description (if any). Students will get notification for any newly posted assignments. The assignment submission window will remain open until the deadline and there is a countdown counter in the assignment submission page.

i) Class Calendar/ Event Module

A teacher can create any event, such as, Midterm Exam, Quiz etc. and specify a timeline. A calendar system has been incorporated in this module which marks the events on that appointed date. This acts as a reminder for the teachers and students for any particular event of the course. Students within that class can view the event. And prepare themselves for the event.

j) Grade Module

This module can only be accessed within a classroom page. Two other options have been added for this module for the teachers. Teachers can create new poll (Mid 1, Mid 1, Final, Quiz1, Quiz 3, Assignment 2 etc.) with full marks and the poll weight (%) or they can view and update previous polls. Some hand-written quizzes or assignments can be taken by the teachers. That is why assignment, quiz, etc. are also kept as poll items. After creating a new poll, students list is shown with a blank text box against each student where a teacher has to give the mark. Teachers can edit the poll weight, full marks and the students' marks. Students can only view their poll's marks. All the polls are listed in a table in the student's end. Finally, a teacher can generate the final grade according to the grading policy of the university. Attendances, quizzes and assignments marks are taken from their corresponding module.

**IV. IMPLEMENTATION AND THE SYSTEM**

In this section, the implementation of the system is illustrated with some representative screenshots of the automated course management system. This approach is taken to explain the features of the system.

The desired implementation of the system is based on Codeigniter framework, which is a PHP framework [19], the programming language is PHP for web server-side programming and JavaScript, CSS and JQuery for web front-end programming. The database used in the system is capstone, which is a MySQL database.

The main interface for all type of users is shown in Fig. 5. After successful login, different types of users will be redirected to their own homepages. No sign-up option is available for the users. The reason is already discussed regarding ACMS.



Figure 5. Main interface of ACMS.

Fig. 6 is visible after successful login. In the middle, all the courses of the current users are shown and in the left side, notification, message, backpack, file sharing, announcement, assignment, quiz etc. modules are listed.

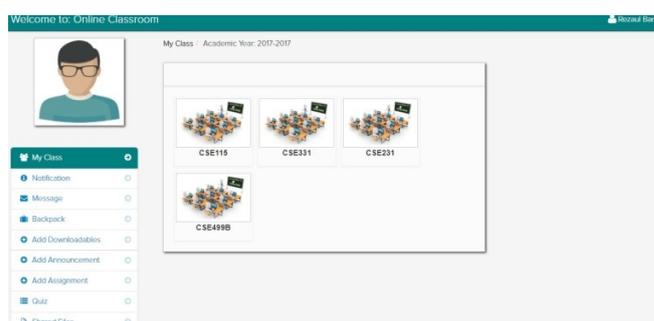


Figure 6. The instructor's homepage.

Fig. 7, is a dynamic page. Students list will be shown after clicking a classroom from the teacher homepage. Subject overview, attendance, file sharing, assignment, announcement, class calendar, grade modules are presented here.



Figure 7. The classroom homepage from the instructor end.

Fig. 8 illustrates the homepage of a student after a successful login. All the courses of the students are listed in the middle and the notification, message, backpack modules are itemized in the left sidebar.

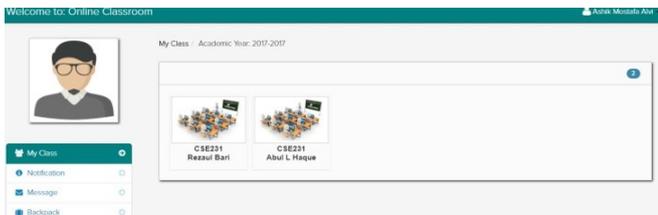


Figure 8. The student homepage.

The classroom homepage is shown in Fig. 9 from the student's end. Classmates list will be shown in the middle with subject overview, attendance, file sharing, assignment, announcement, event, grade modules in the left sidebar.

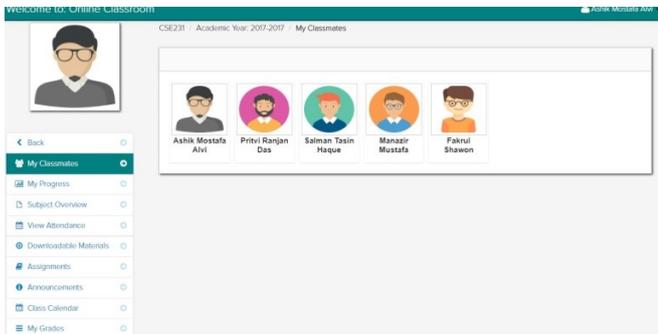


Figure 9. The classroom homepage from the student's end.

Fig. 10 and Fig. 11 demonstrates the attendance module: Fig. 10, is from the teacher's end and Fig. 11, is from student's end.

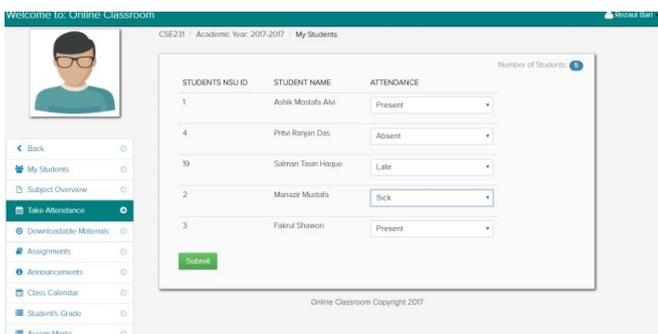


Figure 10. The attendance page from the teacher's end.

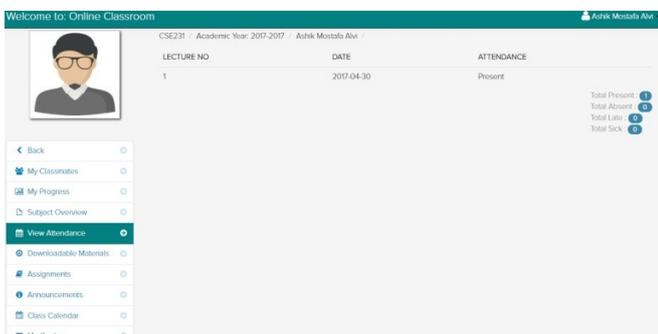


Figure 11. The attendance page from the student's end.

Screenshots (Fig. 12 - Fig. 17) of some of the core modules are given below.

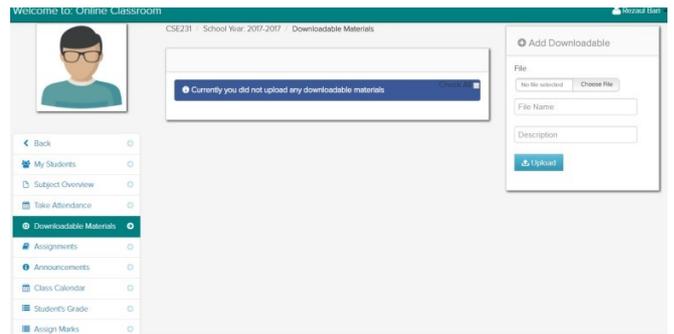


Figure 12. The file sharing page from the teacher's end.

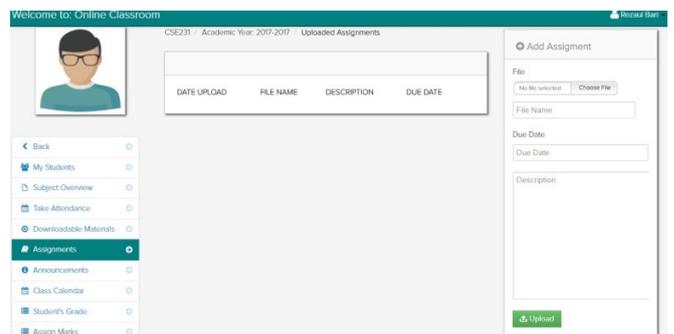


Figure 13. The assignment page from the teacher's end.

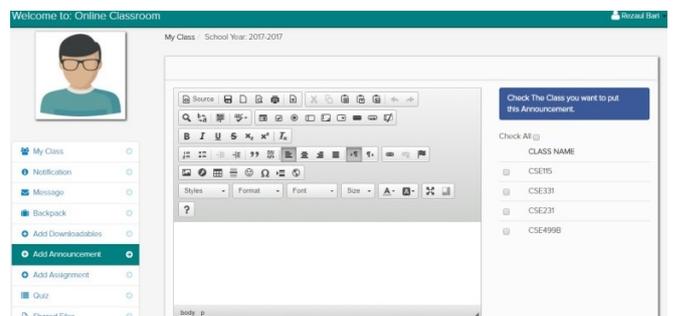


Figure 14. Screenshot of the announcement from teacher's end.

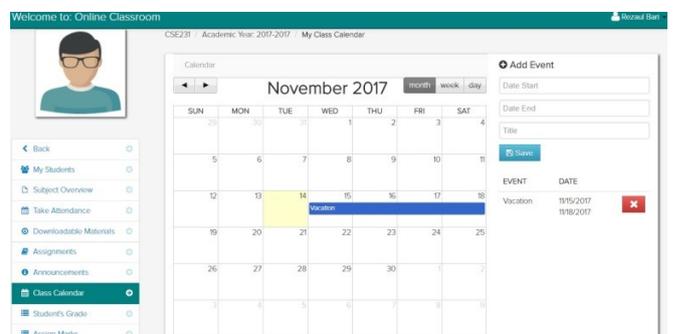


Figure 15. The event creation page from the teacher's end.

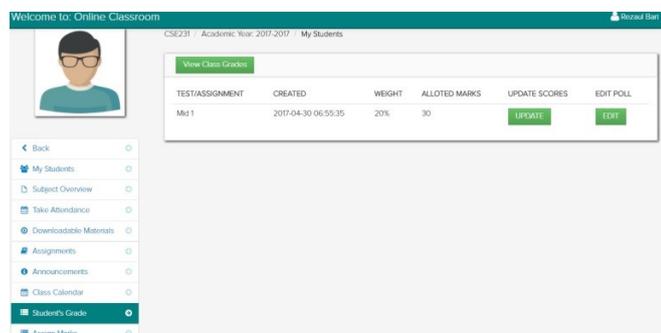


Figure 16. The poll viewing and editing page from the teacher's end.

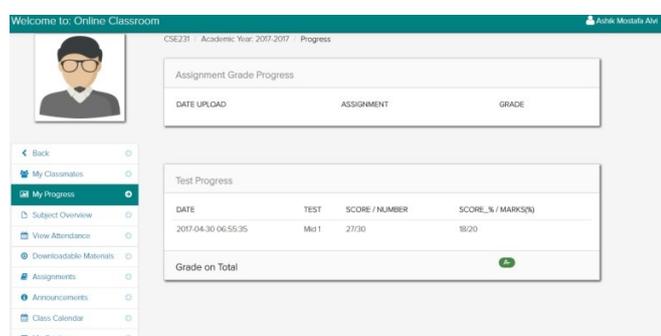


Figure 17. The final grade page from the student's end.

## V. CONCLUSION AND FUTURE WORK

From the description of the above designed and developed system, it can be concluded that the proposed system is a hassle-free automated classroom management system with all the LMS features. It is a highly secured system as there is no outside or third-part involvement in the system. It uses the central database of the university and, as such, creation and maintenance of an additional database is not required.

Most of the world's top universities are using their own classroom management system to address concerns of security and privacy. An idea has been introduced in this paper to design an automated but simple course management system for universities offering open credits without condoning the security and privacy issues.

In the future, the central Advising System can be integrated in the existing system so that the course management system can circumvent dependence on the advising database. The online tutorial sharing and online examination taking modules can also be integrated in the system. These works would make the designed system even more complete.

## ACKNOWLEDGMENT

The authors would like to thank the faculty members of the Department of Electrical and Computer Engineering of the North South University for their guidance and assistance during survey and requirements analysis. The authors also acknowledge the support of the Information Services

Department of the same university in providing access to the required test and validation data.

## REFERENCES

- [1] Ting IH., Wu WJ., Kao HT., Wang D. (2015) An Implementation of Online Learning and Course Management System Based on Facebook. In: Uden L., Liberona D., Welzer T. (eds) Learning Technology for Education in Cloud. LTEC 2015. Communications in Computer and Information Science, vol 533. Springer, Cham.
- [2] "Top 20 Facebook Statistics - Updated April 2017," Zephoria Inc., 11-Apr-2017. [Online]. Available: <https://zephoria.com/top-15-valuable-facebook-statistics/>. [Accessed: 23-Apr-2017].
- [3] Young, Sherman, and Ian Solomonides. "PACE and Online Learning and Engagement". Learning through Community Engagement, 2017, pp. 263-273, Springer, Singapore.
- [4] McAndrew, Patrick, Peter Goodyear, and James Dalziel. "Patterns, Designs and Activities: Unifying Descriptions of Learning Structures". International Journal of Learning Technology 2.2/3 (2006): 216. Web. 23 Apr. 2017.
- [5] Gonzalez-Barbone, Victor, and Luis Anido-Rifon. "Creating the First SCORM Object". Computers & Education 51.4 (2008): 1634-1647. Web. 23 Apr. 2017.
- [6] Ceylan, Beril, Birim Balci, and Mustafa Murat Inceoglu. "An Application of Creating and Packaging Learning Objects". Procedia - Social and Behavioral Sciences 1.1 (2009): 2051-2056. Web. 23 Apr.
- [7] Müller, Christian. "Experiences and Evaluation of a Blended Learning Concept for Learning Chinese in Higher Education". Procedia - Social and Behavioral Sciences 34 (2012): 158-163. Web. 22 Apr. 2017.
- [8] Adeyinka, Tella, and S. Mutula. "A Proposed Model for Evaluating the Success of Webt Course Content Management System". Computers in Human Behavior 26.6 (2010): 1795-1805. Web. 20 Apr. 2017.
- [9] "Learning Management System". En.wikipedia.org. N.p., 2017. Web. 11 Apr. 2017
- [10] Du Z., Fu X., Zhao C., Liu Q., Liu T. (2013) Interactive and Collaborative E-Learning Platform with Integrated Social Software and Learning Management System. In: Lu W., Cai G., Liu W., Xing W. (eds) Proceedings of the 2012 International Conference on Information Technology and Software Engineering. Lecture Notes in Electrical Engineering, vol 212. Springer, Berlin, Heidelberg.
- [11] Davidson-Shivers, Gayle V. "Frequency and Type of Instructor Interactions in Online Instruction". Journal of Interactive Online Learning 8 (2009): 23-40. Web. 21 Apr. 2017.
- [12] Tang, Alicia. (2013). LEARNING MANAGEMENT SYSTEM USING MULTI-AGENT TECHNOLOGY: A PROPOSED IMPLEMENTATION STRATEGY. International Journal of Asian Social Science. 3. 1878-1886.
- [13] F. Smith, "How to Use Social-Networking Technology for Learning," Edutopia, 20-Apr-2007. [Online]. Available: <https://www.edutopia.org/how-use-social-networking-technology>. [Accessed: 23-Apr-2017].
- [14] Navarrete, Cesar C., and George Veletsianos. "Online Social Networks as Formal Learning Environments: Learner Experiences and Activities". The International Review of Research in Open and Distributed Learning 13.1 (2012): n. pag. Web. 22 Apr. 2017.
- [15] Wilson, Christo et al. "User Interactions in Social Networks and Their Implications". Proceedings of the fourth ACM european conference on Computer systems - EuroSys '09 (2009): n. pag. Web. 21 Apr. 2017.
- [16] "Blackboard | Education Technology & Services". Anz.blackboard.com. N.p., 2017. Web. 22 Apr. 2017.
- [17] "Best Learning Management Systems 2017 - Reviews & Comparison". Lms.softwareinsider.com. N.p., 2017. Web. 19 Apr. 2017.
- [18] "Engrade". McGraw Hill Education. N.p., 2017. Web. 16 Apr. 2017.
- [19] "Codeigniter Web Framework". Codeigniter.com. N.p., 2017. Web. 17 Apr. 2017.

# Change Management and the Integration of Information Technology

Research Notes from Selected African Universities

Omotayo Kayode Abatan  
School of Computing  
University of South Africa  
Johannesburg, South Africa  
eabatao@unisa.ac.za

Manoj Maharaj  
School of Management, IT and Governance  
University of KwaZulu-Natal  
Westville, South Africa  
maharajms@ukzn.ac.za

**Abstract**— The aim of this article is to examine change management awareness (CMA) by evaluating the perceptions of academics regarding information technology integration in higher education in Africa. Data were collected from 592 academics at three prominent higher education institutions in Africa. These included two direct contact (or on-site) institutions, namely Lagos State University in Nigeria and the University of KwaZulu-Natal in South Africa, and an open distance learning (ODL) institution, the University of South Africa. Data were analysed using descriptive statistics. The requirements of effective integration of information technology and change management awareness in higher education were identified using a change management model. The findings indicated that academics are self-aware of the need for change in order to enhance technology integration in higher education. The study identified that universities should take responsibility in providing the strategies needed to implement change in the use of information technology.

**Keywords**- academics; change management; higher education; information technolog; information and communication technologies

## I. INTRODUCTION

This Information and communication technologies (ICTs) play a significant role in education, both in formal and informal settings. The evolution of ICTs has produced numerous information technology (IT) tools and resources used for various educational activities, such as conventional teaching and online instructions and/or web-based education (electronic learning/e-learning) as an alternative approach to “chalk and board” teaching and learning. The creation and integration of new IT through ICTs into higher education is progressively changing and transforming the field of education in terms of teaching and learning, and the creation of ideas. IT has changed the conventional methods of discussing and disseminating ideas within the educational environment. It has completely transformed the way in which academics transfer new knowledge and research findings to students [1]. IT may provide powerful learning opportunities to academics and students, nevertheless, both lecturers and students need to learn how to take full advantage of the opportunities that learning technologies provide, in order to improve their profiles in the globalised educational market [2].

The globalisation of ICTs, which has escalated the use of IT for educational purposes, does not imply that it is possible to claim that the presence of IT has been fully exploited [3]. This study does, however, reflect on the current state of IT integration in African higher education, by evaluating related historic trends and the underlying reasons for such integration. The next section of the article presents the literature study. Next is a discussion of the change management model, as well as the research design and methodology. In what follows, the article focuses on academics’ perceptions of change management awareness, significance testing and summary of the findings. Finally, the article concludes by proving that there is a strong awareness of the need for change management as it deals with IT integration in African universities.

## II. LITERATURE REVIEW

There are different types of learning technologies, each with different features and functions. Included are blended, mobile, ubiquitous, e- and online learning [1]. All these learning technologies are increasingly in use, as several higher education institutions (HEIs) are enthusiastic about integrating them. The creation and utilisation of learning technologies in pedagogy is conceivably the most effective approach to technology integration. Garnham and Kaleta claim that learning technologies focus more on information delivery than on student learning [4]. Hybrid/blended learning is described as the type of learning where a significant part of the learning activities are carried out online, thus not completely eliminating the time spent in traditional classrooms but radically reducing it. The use of IT for educational purposes can enhance the quality of education by increasing academics’ motivation and engagement, facilitating the acquisition of basic skills, and increasing students’ learning outcomes. IT can greatly enrich teaching and learning in higher education if it focuses on the basic objectives of education. IT also offers an opportunity to alleviate higher education challenges [5].

HEIs represent a country’s skills base as they facilitate the exchange of information, provide knowledge sources and transform the economy through university–industry networks [6]. To generalise, higher education is a means of improving economic growth and mitigating poverty in any country [7]. In contrast to what is generally accepted, some international

development community (IDC) members have indicated that higher education has little or no impact on reducing poverty in Africa [8]. Some African countries today are still struggling to match student enrolment levels with institutional capacity. In addition, technology integration and academic research output in Africa is among the lowest in the world [9] and these are among the challenges facing higher education in Africa. These challenges can be categorised into technological advancement, social progress and economic development. One major means through which higher education in Africa can enhance economic development is through technological catch-up.

According to Ndoye, the first African ministerial forum indicates that the innovative integration of ICTs into teaching and learning has prompted the creation and implementation of plans and policies in respect of the following: the development of student-centered approaches; guidance towards and planning for change; the up-skilling of students, academics and institutional management [10]. These plans and policies can improve the use of digital devices which are not limited to interactive whiteboards, computers, cell phones and tablets. The policies aim to enhance the promotion of logical curricular reforms for the computer age in terms of skills, knowledge and value [10]. The implementation of these plans and policies will prompt change among stakeholders (i.e. academics, students and management). However, for these plans and policies to have a long-term effect and be successful, some African countries may need to be supported/funded by partnering with international organisation such as the World Bank, the United Nations Economic Commission for Africa (UNECA) or the United Nations Educational, Scientific and Cultural Organisation (UNESCO). In Nigeria, for instance, the Connect Nigeria Initiative (CNI) by the Federal Ministry of Nigeria connected over 1.4 million higher institution students to the Internet across 27 federal universities. The initiative received outside assistance by partnering with the World Bank's STEP-B project [11].

Another World Bank project, launched in partnership with Kenyatta University in Nairobi, is aimed at establishing education networks that offer computer studies as a subject in schools. The outcome of this initiative has made it much easier for educators to integrate ICTs into their higher education systems [12]. There are still imbalances in access to ICTs within South African HEIs, due to several developmental challenges and the digital divide in the country [13]. With reference to the slow pace in which ICTs are integrated into the teaching and learning environment, the South African Department of Education called for ICTs to be integrated into school curricula. The purpose was not only to improve the quality of education, but also to increase educational skills, and enhance teaching and learning outcomes.

Today, it is difficult to identify many operations and activities in African HEIs that do not involve/support the use of ICTs. The Internet and other relative technologies have had an extensive impact on the way people live and work in Africa. According to Alemneh and Hastings, "digital technology advancements are shaping the way people create, use, preserve and access information resources that the traditional methods of accessing or organizing information resources are no longer effective" [14]. Internet applications and developments in digital library initiatives are making information resources easier to

access, and providing academics in Africa with access to more diverse information sources and services which will allow ICTs in education to continue to grow [15]. Therefore, to facilitate the integration of ITs in higher education, it is important to create strategies to implement guidelines and plans for change.

The next section presents the model that underpins the construct of this article, which examines change management awareness by evaluating academics' perceptions regarding IT integration in higher education and the underlying reasons for change.

### III. CHANGE MANAGEMENT MODEL

Managing changes in higher education does not necessarily impose the introduction of new technology. Rather it is about encouraging the people involved in the delivery of instruction or education to change the way they do things [16]. The process of managing change begins with individuals' or people's understanding that change is actually needed in an institution. What follows is that people must understand and accept that they must change. Finally, actually embracing change – a process which may take several years to achieve, but will eventually enhance the integration of technological innovation into higher education. To better understand this process, Kershaw indicates that the strategies for implementing change in any institution should involve clarifying the need for educational technology, creating suitable institutional/organisational structures, providing adequate support, training, and promoting technology use for different academic purposes [16]. The institution must be prepared to reallocate limited resources to support academics, learners and other staff who use the technology, otherwise there will be no change.

### IV. RESEARCH DESIGN AND METHODOLOGY

A quantitative research methodology was applied. Self-administered and structured questionnaires were distributed to participants using a simple random sampling technique. The sample population was academic staff members which included tutors/teaching assistants, junior lecturers, senior lecturers, associate professors and professors. The study performed a quantitative analysis on the collected data by means of the Statistical Package for the Social Sciences (SPSS) and Microsoft Excel. An analysis of the findings was used to draw sound conclusions and offer appropriate recommendations. A reliability test was performed to validate the integrity of instrument used. In addition, this exploratory study furnished new insights into the effective integration of IT and change management awareness in HEIs.

The study was conducted in Nigeria and South Africa. Questionnaires were distributed to participants at two prominent direct contact (on-site) HEIs in each country, and one open and distance learning (ODL) institution in South Africa. The institutions are 1) Lagos State University (LASU) in Lagos, Nigeria; 2) the University of KwaZulu-Natal (UKZN) in Durban, South Africa; and 3) the University of South Africa (Unisa) in Pretoria, South Africa. In total, 193 questionnaires were obtained out of the 450 handed out at LASU, another 198 were obtained out of the 437 handed out at UKZN, and 201 were obtained out of the 555 handed out at Unisa. Thus, in total 592 academics participated in this study.

V. ACADEMICS' PERCEPTIONS OF CHANGE MANAGEMENT AWARENESS

To probe the construction of opinions about change management awareness, the researcher required participants to indicate the opinion that best represents how they feel about the imperative to use ITs. This question was intended to offer insight into academics' awareness of change management, their understanding of how participants are encouraged to get involved in using technology to deliver instructions and how they viewed their respective roles in the institution. Each participant was required to indicate, on a scale of 1–4, their perceptions of change management, with the possible answers being “strongly disagree”, “disagree”, “agree” and “strongly agree”. The researcher took cognizance of the opinions that may have a direct impact on participants' perceptions and understanding of change management. The assumption is that the more strongly they agree, the more likely they are to understand that change is actually needed and the more they tend to accept change in the use of IT in higher education.

A. Perceptions of Change Management Awareness – LASU

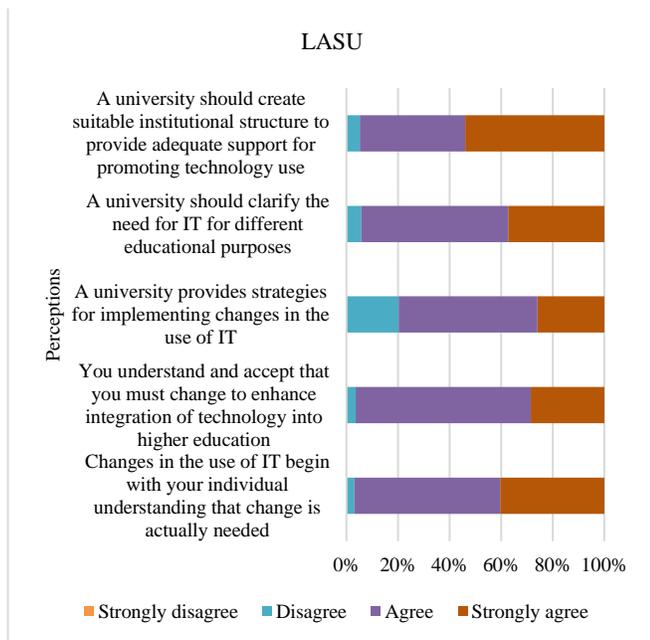


Figure 1. Change Management - LASU

A significant number of participants in LASU indicated that “change in the use of information technology begins with their individual understanding that change is actually needed” with 56.48% (agree) and 40.41% (strongly agree). The same can be said about participants who understand and accept that they must change to enhance the integration of technology into higher education: 67.88% (agree) and 28.50% (strongly agree). The third highest rating in terms of awareness to change management is the perceptions that university should create suitable institutional structure to provide adequate support for promoting technology use, to which 40.93% of participants indicated they agree and 53.89% indicated they strongly agree. What follows is the proposition that university should clarify the need for information technology for different educational purposes, to

which 56.99% indicated that they agree and 37.31% indicated they strongly agree. 20.21% of participants indicated that they disagree that university should provide strategies for implementing changes in the use of information technology while 53.89% and 25.91% indicated they agree and strongly agree respectively.

B. Perceptions of Change Management Awareness – UKZN

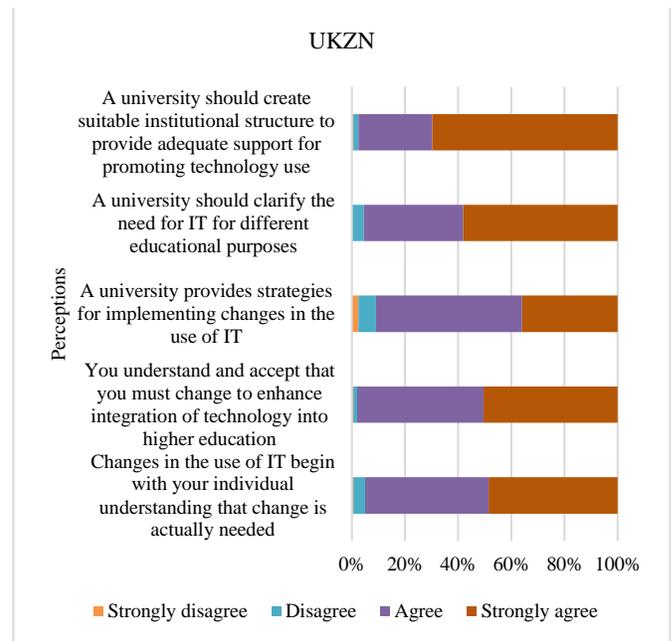


Figure 2. Change Management - UKZN

In order to understand the formation of the opinions of participants about change management awareness, participants in UKZN were required to specify the choice that best represents how they felt about the imperative for use of information technology in higher education on a scale of 1-4 with possible answers being strongly disagree, disagree, agree and strongly agree. To achieve the research objective that seek to examine the nature, rationale, strength and weaknesses in the adoption of information technology, the responses presented in the Fig. 3 below offer understanding of the perceptions of academics regarding change management in the adoption of information technology. Participants' responses presented in Fig. 2 show that each opinion/perception is imperative to change management as a high percentage of participants (over 90%) in UKZN specified agree and strongly agree to the propositions about change management. This is a clear indication that most academics who participated in UKZN survey agreed to change, and perceived these opinions as critical to the adoption and integration of IT into higher education.

C. Perceptions of Change Management Awareness – UNISA

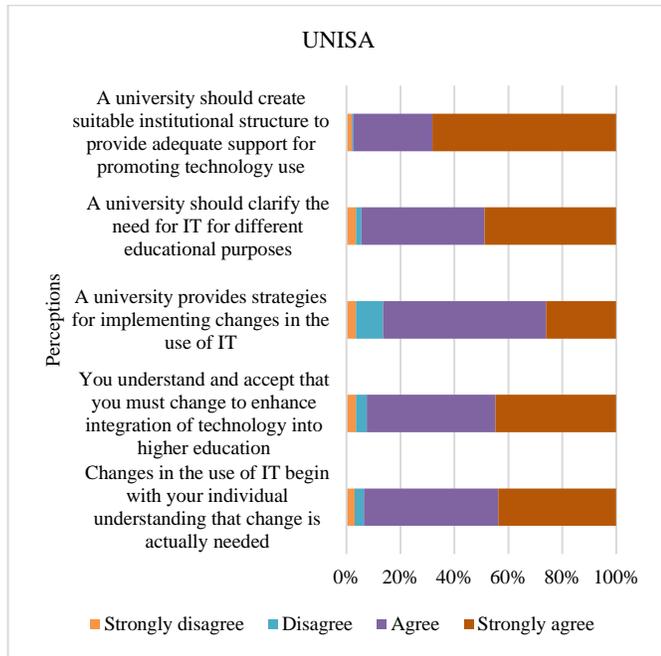


Figure 3. Change Management - UNISA

From Fig. 3, academics in UNISA who specified agree and strongly agree to the five perceptions about change management self-awareness constituted an overwhelming number of over 86% combined. In actual fact, the findings suggest that participants (49.80%) agree and (43.80%) strongly agree that change begins with their individual understanding that change is needed in order to integrate information technology in higher education. Majority of participants (47.80%) agree and (44.80%) strongly agree that they have to accept change to enhance the integration of technology in higher education. Another significant number of participants agree and strongly agree that the university should provide strategies for implementing changes in the use of information technology at 60.50% and 26.00% respectively. The same can be said of the participants who agreed that university should clarify the need for information technology for different educational purposes, in which case 45.80% and 48.80% agree and strongly agree respectively. “University should create suitable institutional structure to provide adequate support for promoting technology use” had the highest rating in which 29.40% agree and 68.20% strongly agree to the proposition on change management. It can be seen from Fig. 3 that participants generally perceived change management as a model to enhance the integration of IT in higher education.

VI. SIGNIFICANCE TESTING

Cronbach’s Alpha analysis was developed by Lee Cronbach to provide the measure of internal consistency of a scale, item or test, and it can be interpreted in numbers, between 0 and 1. The purpose of presenting internal consistency is to identify the extent to which all items in a scale or test, measure the same construct or concept. According to Tavakol and Dennick, Cronbach Coefficient Alpha analysis is necessary in a study of

this nature (with over 300 participants), and is a sufficient condition for measuring unidimensionality or homogeneity in a sample of test items [17]. The test is used to understand the variance in Cronbach Coefficient Alpha. There have been different reports about the acceptable values of Alpha, which ranges from 0.70 to 0.95. Tavakol and Dennick (2011) note that a Cronbach’s Alpha value near 0.7 is acceptable, but a value lower than 0.7 could be due to a low number of participants or redundancy in the number of questions.

A reliability test was conducted on the five items thought to be useful for examining the strategies for the integration of IT in African HEIs. These items are thought to be the underlying reasons which motivate changes in the use and integration of technology by academics in HEIs. Table I shows that the resulting value of the Cronbach’s Alpha reliability test on the five items was 0.779 for the case processing summary obtained from 592 participants at LASU, UKZN and Unisa.

TABLE I. RELIABLE TEST ON CHANGE MANAGEMENT AWARENESS (CRONBACH’S ALPHA)

Case Processing Summary			
		N	%
Cases	Valid	591	99.8
	Excluded <sup>a</sup>	1	.2
	Total	592	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability statistics	
Cronbach's Alpha	N of items
.779	5

If the least acceptable Cronbach’s Alpha value is equal to or greater than 0.7, but less than 0.95, the reliability of the items measured on change management is said to be acceptable, internally consistent and not redundant. A Cronbach’s Alpha of 0.779 means that there is a 0.39 error variance or random error, calculated as follows:

$$(0.78 \times 0.78 = 0.61; 1.00 - 0.61 = 0.39)$$

Table II shows the results of the total statistics/measurements conducted on each item/variable used to measure the reliability of the instrument. The Cronbach’s Alpha of the five items measured are within the acceptable 0.7 to 0.95 Alpha value. This is an indication that the items/variables used to measure change management awareness among academics in HEIs are acceptable, internally consistent, not redundant and significantly reliable.

TABLE II. ITEMS – TOTAL STATISTICS

	<i>Cronbach's Alpha if item deleted</i>
Changes in the use of IT begin with your individual understanding that change is actually needed	.773
You understand and accept that you must change to enhance integration of technology into higher education	.710
A university provides strategies for implementing changes in the use of IT	.755
A university should clarify the need for IT for different educational purposes	.710
A university should create suitable institutional structure to provide adequate support for promoting technology use	.739

F-test was also conducted on the change management awareness instruments thought to be useful to examine the perception of academics with regards to the integration of IT in HEIs in Africa. The test results are presented in Tables III and IV.

To test if the variances are the same or equal and if the variances are not the same or equal, we consider H0 and H1 where:

H0: The variances are the same or equal

H1: The variances are not the same or equal

Test for ANOVA in Table III and the intraclass correlation coefficient in Table IV indicates that there is a difference in means across the factors. The associated p value is .000, where the value of p – when it is less than or equal to 0.05 ( $p \leq 0.05$ ) – is an indication that the hypothesis is statistically significant.

TABLE III. ANOVA TEST ON CMA (F-TEST)

ANOVA <sup>a</sup>					
	Sum of Sq.	df	Mean Sq.	F	Sig.
Between people	629.662	590	1.067		
Within people					
Between items	65.896	4	16.474	69.813	.000
Residual	556.904	2360	.236		
Total	622.800	2364	.263		
Total	1252.462	2954	.424		

Grand Mean = 3.38

Therefore, H0 is rejected, since the p value is less than 0.01, a very strong evidence to reject the null of no difference in means across the group of variables. This means that the variances used in the study for measuring the change management awareness and the effective integration of IT among academics in higher education are significantly different.

TABLE IV. INTRACLASS CORRELATION COEFFICIENT ON CMA

Intraclass correlation coefficient							
	Intraclass correlation <sup>a</sup>	95% Confidence interval		F-Test with true value 0			
		Lower bound	Upper bound	Value	df1	df2	Sig
Single measures	.413 <sup>b</sup>	.374	.454	4.523	590	2360	.000
Average measures	.779 <sup>c</sup>	.749	.806	4.523	590	2360	.000

Two-way mixed effects model where people effects are random and measures effects are fixed.

- a. Type C intraclass correlation coefficients using a consistency definition-the between-measure variance is excluded from the denominator variance.
- b. The estimator is the same, whether the interaction effect is present or not.
- c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

## VII. SUMMARY

We have examined the change management awareness that can be used to evaluate the integration of IT into higher education. The study suggests that the majority of participating academics thought that change management awareness would help to facilitate and improve technology integration in HEIs. The article concludes that university management should provide guidelines for managing change (change management) for any innovation and/or emerging IT developed and introduced into teaching and learning. Institutional management should be responsible for clarifying the need for change to users or stakeholders (i.e. academics, students and faculty), by providing compelling reasons for the importance of embracing change. Institutional management is also responsible for creating suitable institutional structures (e.g. policies) that will enhance change management and the integration of IT.

The creation and/or development of strategies for implementing policies to accomplish the vision, mission and mandate of the institution, as well as accommodating stakeholders' requirements and needs, should be guided by the institution. This will give direction and motivate the use and integration of technology. Finally, it is the responsibility of the institution to create change management awareness among academics. In doing so, it will be easy to manage both the academics and the emerging technology. It will thus meet the fundamental requirements for the integration of IT in higher education.

### VIII. CONCLUSION

This study has examined change management awareness among academics at three prominent HEIs in Africa, namely Lagos State University in Nigeria, the University of KwaZulu-Natal in South Africa, and an open distance learning (ODL) institution, the University of South Africa. The result generally indicated that academics are not only aware of change management, but believe changes are needed to integrate technology into higher education. However, groups of participants in Nigeria disagreed, arguing that it is not the responsibility of the institution to provide strategies to implement changes in the use of IT. This could mean that changes in the use and integration of IT in higher education should begin with individuals' understanding and their acceptance of change. The study has created a platform for HEIs to understand that academic staff are aware of change, and it can serve as a useful tool for initiating the process of change. Academic change management awareness will enhance the integration as well as the introduction of technology into teaching and learning processes. Change management awareness could serve as a motivational tool to assist HEIs in obtaining funds to acquire more technology to enhance learning outcomes.

### ACKNOWLEDGMENT

Gratitude goes to the academics at LASU, UKZN and UNISA who participated in the study. Further thanks goes to Abiola Abatan for technical support.

### REFERENCES

- [1] J. Joseph, "The Barriers of Using Education Technology for Optimizing the Educational Experience of Learners," *Procedia - Social and Behavioral Sciences*, Vol 64, 2012, p. 427–436.
- [2] D. Euler, "Organisational eCompetence Development," Annual AHD Congress, Conference Presentation, 2004.
- [3] K. A. Bingimlas, "Barriers to the Successful Integration of ICT in Teaching and Learning Environments: A Review of the Literature," *Eurasia Journal of Mathematics, Science and Technology Education*, Vol 5, No. 3, 2009, p. 253–245.
- [4] C. Garnham, and R. Kaleta, "Introduction to Hybrid Courses," *Teaching with Technology Today*, Vol 8, Issue 6, 2002, [Online] <http://www.uwsa.edu/ttt/articles/garnham.htm> [Accesses Date: 10 January, 2017].
- [5] S. Jaffer, D. Ng'ambi, and C. Czerviewicz, "The Role of ICTs in Higher Education in South Africa: One Strategy for Addressing Teaching and Learning Challenges," *International Journal of Education and Development using Information and Communication Technology*, Vol 3, Issue 4, 2007, p. 131–142.
- [6] D. Kapur, and M. Crowley, "Beyond the ABCs: Higher Education and Developing Countries," *Center for Global Development*, Vol 109, 2008.
- [7] D. E. Bloom, D. Canning, and K. Chan, "Higher Education and Economic Development in Africa," *Human Development Sector, African region*, 2006, p. 1-83.
- [8] D. E. Bloom, D. Canning, K. Chan, and D. L. Luca, "Higher Education and Economic Growth in Africa. *International Journal of Higher Education*, Vol 1, No.1, p.23-27, 2014, [Online] <https://ejournals.bc.edu/ojs/index.php/ijahe/article/view/5643/4974> [Accesses Date: 5 January, 2017].
- [9] T. Yizengaw, "Challenges of Higher Education in Africa and Lessons of Experience for the Africa - U.S. Higher Education Collaboration Initiative, National Association of State Universities and Land-Grant Colleges (NASULGC), Washington D. C., 2008, p. 1–21.
- [10] M. Ndoye, "First African Ministerial Forum on ICT Integration in Education and Training," *Association for the Development of Education in Africa*, 2014 [Online] [http://www.adeanet.org/en/system/files/report\\_on\\_ict\\_forum\\_dec\\_2013\\_en.pdf](http://www.adeanet.org/en/system/files/report_on_ict_forum_dec_2013_en.pdf) [Accesses Date: 15 December, 2015].
- [11] A. Opoku-Mensah, "The State of eLearning Readiness in Africa," 10th e-Learning Africa International conference: Enriching Tomorrow. Addis Ababa: ICWE GmbH, 2015, p. 6–9.
- [12] S. Ndege, "Teaching Teacher Trainers to Teach Online," 10th e-Learning Africa International Conference: Enriching Tomorrow. Addis Ababa: ICWE GmbH, 2015, p. 38–39.
- [13] T. Assan, and R. Thomas, "Information and Communication Technology Integration into Teaching and Learning: Opportunities and Challenges for Commerce Educators in South Africa". *International Journal of Education and Development using Information and Communication Technology*, Vol. 8, Issue 2, 2012, p. 4-16.
- [14] D. G. Alemneh, and S. K. Hastings, "Develop The ICT Infrastructure for Africa: Overview of Barriers to Harnessing the Full Power of the Internet. *Journal of Education for Library and Information Science*, Vol. 47, No. 1, 2006 p. 4-16. Retrieved from [http://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=1036&context=libsci\\_facpub](http://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=1036&context=libsci_facpub) [Accesses Date: 10 January, 2017]
- [15] M. Kossai, and P. Piget, "Adoption of Information and Communication Technology and Firm Profitability: Empirical Evidence from Tunisian SMEs," *Journal of High Technology Management Research*, Vol 25, 2014, p. 9–20.
- [16] A. Kershaw, "People, Planning, and Process: The Acceptance of Technological Innovation in Post-Secondary Organizations," *Educational Technology*, 1996, p. 44–48.
- [17] M. Tavakol, and R. Dennick, "Making Sense of Cronbach's Alpha," *International Journal of Medical Education*, Vol 2, 2011, DOI: 10.5116/ijme.4dfb.8dfd. [Accesses Date: 10 January, 2017].

# Web Service Injection Attack Detection

Victor Clincy<sup>1</sup> and Hossain Shahriar<sup>2</sup>

<sup>1</sup>Department of Computer Science

<sup>2</sup>Department of Information Technology

Kennesaw State University, USA

{vclincy, hshahria}@kennesaw.edu

**Abstract-** Injection attacks on web services can expose valuable information resources. To protect deployed web services against injection attacks, it is important to have defense techniques. Intrusion Detection Systems (IDS) are popular defense techniques to mitigate network layer attacks. This paper proposes an IDS for mitigating injection attacks on web services. We apply Genetic Algorithm (GA) as part of new attack signature generation for web services. The approach has been applied to a prototype web service and was found effective in generation of new attack signatures.

**Keywords:** Web service, SOAP injecton, XPath, RESTful, DoS.

## I. INTRODUCTION

Web services are deployed to deliver data and content for online stores. Simple Object Access Protocol (SOAP) is used to transmit data between a web service consumer and provider over the network [20, 21]. Unfortunately, web services suffer from code injection vulnerabilities such as XPath and SOAP injection. There are some available mitigation approaches (e.g., [1, 2, 4, 6, 7, 9, 12, 18]) intended to mitigate code injection attacks. However, there is little or no effort in building Intrusion Detection System (IDS) for web services. An IDS can match input of web services for suspected attacks and prevent the payload to be processed further.

In this paper, we develop an Intrusion Detection Systems (IDS) for web services. We leverage Genetic Algorithm (GA) to generate new attack signatures from a set of initial signatures. We analyze SOAP messages, define a set of population and apply various steps of GA to generate new attack signatures. Our approach shows the application of GA to generate common attack signatures on web services including Xpath injection, XML bomb, SQL Injection and Remote File Inclusion. We evaluate the approach for a prototype web services and the initial results show the approach can generate new attack signatures.

Our proposed approach allows to add defense-in-depth for defending against web service attacks by complementing existing web scanner tools.

The paper is organized as follows: Section II discusses some examples of injection attacks on web services. Section III discusses related work. Section IV discusses the approach in details. Section V provides initial results and evaluation. Section VI presents the conclusions.

## II. INJECTION ATTACKS ON WEB SERVICES

There are several injection attacks that takes advantage of implementation vulnerabilities. These vulnerabilities include SQL Injection and XPath Injection [1]. These types of injection attacks take advantage of implementation vulnerabilities where malicious inputs may change SQL commands (while accessing data from database tables) or alter XPath queries (while retrieving data from XML documents). We briefly discuss the XPath injection attack in this section.

**XPath Injection:** In this attack, inputs that are not validated and can modify XPath queries [5]. Attackers may gain access to information from XML documents [4]. An Example of an XPath injection attack is shown in Fig. 1.

```

1. XmlDocument XmlDoc = new XmlDocument();
2. XmlDoc.Load("...");

3. XPathNavigator nav = XmlDoc.CreateNavigator();
4. XPathExpression expr =
  Nav.Compile ('string(//user[name/text()="'+input1
  and password/text()="'+ input2 + "']/account/text());

5. String account = Convert.ToString(nav.Evaluate (expr))
6. if (account==""){
    ... ..
    //name and password not found in XML, login failed.
7. } else {
    ... ..
    //login succeeded, proceed to application.
}

```

Figure 1: Example of an XPath vulnerable application

Here, Lines 1-3 are used to open an XML document. Line 4 generates an XPath expression where two inputs are added with the query through input1 and input2 variables. Both inputs are not filtered. An attacker may provide values of input1 as ' or 1=1 or "=" to change the semantic of the original XPath. The resultant query, in this case is `string(//user[name/text()=' or 1=1 or "=" and password/text()='foobar']/account/text())`. This is similar to `string(//user/account/text())`. The resultant query always returns the first account number and attacker would be logged in without providing a valid user name and password (Line 7).

**SOAP action overriding:** SOAP action field is used as a service operation identification which can be modified by an attacker.

```

POST/... HTTP/1.1
...
1. SOAPAction: "getStudentRoles"
2. <Envelope>
3. <Body>
4. <getStudentName>
5. <StudentID>432</StudentID>
6. </getStudentName>
7. </Body>
8. </Envelope>

```

Figure 2: Example of SOAP action overriding

Fig. 2 shows an example of SOAP action overriding. The SOAP message is intended to return a student's name if the envelope's action is invoked. However, HTTP header's action overrides the envelope to expose the student's roles.

### III. RELATED WORK

Over the last decade, there are more than a dozens of web scanner tools developed for security testing (e.g., Wapiti [10], W3af [11], WebScarab [13], Wfuzz [14]). These web scanner tools, however, are not directly useful for testing of web services.

A number of literature works have recently focused on testing of web services [1-9]. Vieira *et al.* [1] applied web application security scanner tools for detecting web service related vulnerabilities. The results revealed higher false positive rates and lower rate for attack coverage for the tested tools. Jan *et al.* [4] proposed a fuzz testing tool SOLMI. The tool injects XML meta-characters (e.g., >, <) into the message attempting to alter the structure of the message. Tiwari *et al.* [2] discussed mitigation approaches against web service attacks such as schema validation where all messages are validated with the XML schema. Any SOAP message that deviates from web services specification are rejected. Our approach is complementary to these earlier approaches.

Antunes *et al.* [18] proposed an automatic approach for the detection of XPath Injection vulnerabilities. Their approach instruments web services to intercept all SQL/XPath Commands executed, learn XPath queries issued by services, followed by generating attack loads based on a large set of XPath Injection attacks. In contrast, we generate various types of SOAP messages to generate attack signatures for an IDS. Rosa *et al.* [6] proposed XID, to mitigate zero-day web service attacks. They created an ontology that was queried by SPARQL (SQL like query engine). XID would detect a possible attack in a network packet using snort rules followed by examining packets against the ontology using SPARQL to query to look for an instance of exactly the set of attack actions. If no identical instance is found, the prototype would infer a new attack found.

Patel *et al.* [7] applied schema validation as a self-adaptive schema hardening. They created a prototype extension for IIS server that scans WSDL files to identify SOAP ports and then associated operations with the value of SOAPAction. It generates a XML Schema definition for validation. In

contrast, we cover an extensive number of attack types and apply GA to generate new attack signatures.

Havrikov *et al.* [9] applied search algorithm-based approach to generate XML documents to test web applications that process them. However, their focus was not testing of web service security vulnerabilities. Najjar *et al.* [19] develop a IDS that applies both anomaly and signature-based IDS for web services. It intercepts XML messages between web servers and firewall and examines them with signatures and anomaly patterns. The approach was shown for .NET IIS environment. The approach does not focus on generating new attack signatures.

## IV. PROPOSED DETECTION APPROACH

### A. IDS framework

To show our approach, we first consider a simple web service where the client provides an employee ID (int) and receives a response of detailed employee information (in the form of array of strings having employee name and telephone number). Fig. 3 shows an example of SOAP request, where an employee's name (John Doe) and age (32) is provided (bold line). The inputs are of string and integer (int) type. Figure 4 shows an example of SOAP response, where information on employee's ID (101) and phone number (1-888-888-8888) is received (bold lines).

```

<SOAP-ENV:Envelope
... xmlns:xsd="http://www.w3.org/1999/XMLSchema">
  <SOAP-ENV:Body>
    <ns1:getEmployeeDetails
      xmlns:ns1="urn:MySoapServices">
      <param1 xsi:type="xsd:string">John
Doe</param1>
      <param1 xsi:type="xsd:int">32</param1>
    </ns1:getEmpDetails>
  </SOAP-ENV:Body>
</SOAP-ENV:Envelope>

```

Figure 3: Example of SOAP request

```

<SOAP-ENV:Envelope
...
  <SOAP-ENV:Body>
    <ns1:getEmpDetailsResponse
      <return...
        xsi:type="ns2:Array"
        ns2:arrayType="xsd:string[2]">
        <item xsi:type="xsd:int">101</item>
        <item xsi:type="xsd:string">1-888-888-8888</item>
      </return>
    </ns1:getEmployeeDetailsResponse>
  </SOAP-ENV:Body>
</SOAP-ENV:Envelope>

```

Figure 4: Example of SOAP response

If the input is not validated, an attacker may provide XPath input if the information is stored in an XML document and if one of the nodes in XML tree is known to the attacker (e.g., XPath expression, age > 0, where all employees information would be returned, as shown in Figure 5).

We propose a signature-based IDS for web services. Fig. 5 shows the framework. Here, service consumer sends SOAP requests to a service provider, which in return sends SOAP response to the consumer. We intercept SOAP requests and responses and save them to a log repository. The logs are used by GA-based approach to generate attack signatures.

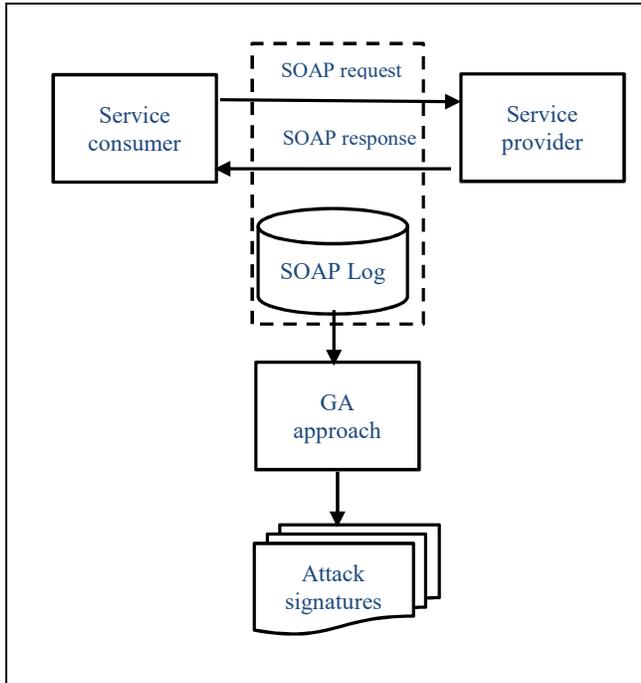


Figure 5: IDS Framework for Web service

```

<SOAP-ENV:Body>
  <ns1:getEmpDetails
    xmlns:ns1="urn:MySoapServices">
    <param1 xsi:type="xsd:string"> Foo </param1>
    <param1 xsi:type="xsd:int"> age > 30 </param1>
  </ns1:getEmployeeDetails>
</SOAP-ENV:Body>
</SOAP-ENV:Envelope>
    
```

Figure 6: Example of SOAP request with XPath injection attack

### B. Steps of the Genetic Algorithm

We apply a GA-based approach to generate these kinds of diverse signatures automatically. It takes a set of population elements as input and generates another set of population as output. The output population includes new and diverse elements compared to the original input population. We briefly outline them as follows.

**Step 1:** Create an initial population. This population consists of elements or chromosomes represented as bit strings.

**Step 2:** Each chromosome in the initial population is then evaluated for fitness. Choose chromosomes based on highest fitness values to perform genetic operation.

**Step 3:** Apply genetic operators such as selection, crossover and mutation to generate new chromosomes.

**Step 4:** Repeat Step 2-3 until the population size reaches to maximum or iteration reaches to maximum.

We discuss chromosome representation and fitness function in the next section.

### C. Chromosome representation

The GA accepts a set of chromosomes as input, and provides a set of chromosomes as outputs after certain number of iterations while following fitness evaluation, cross over and mutations. We convert each of the SOAP message parameters to a chromosome using bit string representation.

Figure 7 shows an example of chromosome representation, which is divided into five segments having total length of 14 bits. Here, the first segment has three bits to represent parameter type: 000-string, 001 – string array, 010-int, 011-int array. Similarly, other data types can be defined based on the need. The second segment represents the counting of injection attack related meta characters (2 bits) and keywords (2 bits) followed by query type (SQL is 0, XPath is 1). For example, based on Figure 7, there are 3 meta characters from SQL injection (“, =, -) and one keyword (or). The chromosome representation is 11010. The third segment is used to capture attack status on privilege escalation using 1 bit. If the request is not a privilege escalation (no mismatch between requested service in SOAP body and HTTP header), it is represented as 0. If the request is a privilege escalation, it is represented as 1. The fourth segment is for disclosure attack represented using 3 bits for presence of file name (1 bit), path (1 bit), encrypted/unencrypted payload (1 bit). The fifth segment is for DoS type attacks and has 2 bits. The first bit represents where recursive data is present in the SOAP header (1 bit) or not. The second bit represents if recursive data is present in the SOAP body (1 bit) or not.

Param_type	Injection	Privilege	Disclosure	DoS
3 bit	5 bit	1 bit	3 bit	2 bit

Figure 7: Example of chromosome representation

An example of chromosomes based on Figure 7 is 00011010000000. Each of the chromosomes length remain same across different types of attacks present in each of the parameters. Before we discuss fitness function, we elaborate our dataset generation.

### D. Dataset generation

As our goal is to apply GA for a signature-based IDS, we start with an attack dataset by deploying a sample web service in Apache and Axis servers. The test services are intended to

catalog products, where a number of services are implemented to query product based on id, retrieve all products, add/delete products. The services were accessed using clients with malicious parameters. The SOAP messages were logged during interaction. We generated malicious inputs based on a number of sources from OWASP security cheatsheet [15, 16, 17]. These attack inputs are applied randomly within SOAP requests from a browser. Table 1 shows the number of samples collected as attack requests.

Table I. Distribution of attack inputs

Attack type	# of samples
SQL Injection	240
XPath injection	120
SOAP overriding	40
Total	400

E. Fitness function and Crossover

We define two fitness functions, Hamming Distance (HD) and Levenshtein Distance (LD).

The Hamming distance between two chromosomes x and y is denoted as HD (x, y), which implies the number of positions where corresponding bits are different. Here, x is part of GA population, where y belongs to Test set. The higher the HD for x, the better the fitness level.

Levenshtein Distance (LD) between two chromosomes x and y is denoted as LD (x, y), which implies the minimum number of editing including insertion, deletion and substitutions needed to change x to y. The higher the LD for x, the better the fitness level.

The chromosomes are crossed over based on fitness level. We apply one point cross over based on second segment position. This generates various types of new attack signatures.

For example, if we consider the following two chromosomes x and y. Here, x represents an SQL injection attack in a string type parameter value (based on Fig. 11), where y represent an integer value having XML bomb in the SOAP request body.

Before cross over (x, y)				
000	11010	0	000	00
010	00000	0	000	01

After applying cross over on the second segment, the resultant chromosomes will look like below, where x now represents a new type of SQL injection attack combined with XML bomb.

After cross over (x, y)				
000	11010	0	000	01

010	00000	0	000	00
-----	-------	---	-----	----

Finally, we choose three mutation rates to randomly mutate bits of chromosomes. As an example, the y chromosome's third segment bit (0) is replaced with 1 (this represents privilege escalation attack, where requested method name needs to be replaced).

Applying mutation on y on third segment

010	00000	1	000	00
-----	-------	---	-----	----

V. EVALUATION

We divide attack dataset into two parts: training dataset (40%) and testing dataset (60%). For each of the training dataset, we convert SOAP requests into chromosome representations by editing and implementing a prototype tool in Java. Fig. 8 shows attack detection accuracy for various population sizes while using HD as the fitness function and keeping the mutation rate at 0.3. We can observe that the higher the selection rate for a chromosome to cross over, the better accuracy for attack detection capability we achieve.

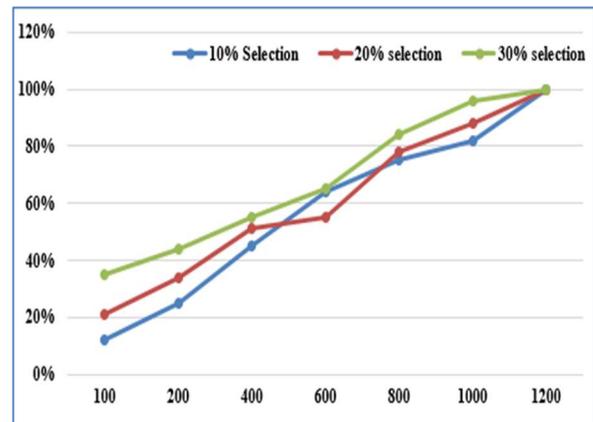


Figure 8: Attack detection accuracy vs. population size for different selection rates (HD, mutation rate=0.3)

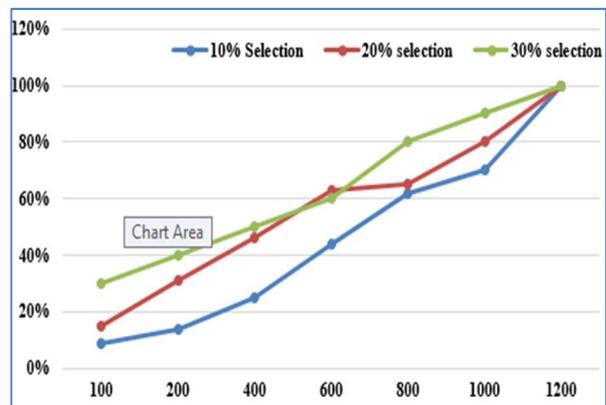


Figure 9: Attack detection accuracy vs. population size for different selection rates (LD, mutation rate=0.3)

Fig. 9 shows the attack detection accuracy for various population sizes while using LD as the fitness function and keeping the mutation rate at 0.3. Each chromosome had varying selection rates, which makes it is easy to see that attacks are detected with more accuracy as the population size increases and the selection rate increases.

Fig. 10 shows attack detection accuracy for various mutation rates. Here, the selection rate is set at 20% and HD is used as the fitness function. The higher the mutation rate is, the increased number of attacks are detected.

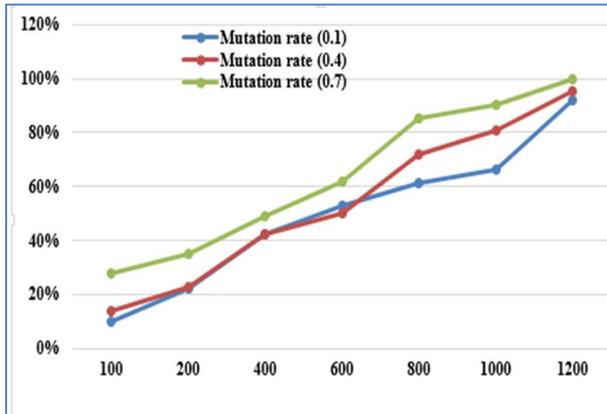


Figure 10: Attack detection accuracy vs. mutation rate (HD, selection rate=20%)

Fig. 11 illustrates that as mutation rate increases, so does the attack detection accuracy using LD as fitness function and the selection rate as 20%. A higher mutation rate shows that the attacks can be detected with more accuracy and in a smaller population.

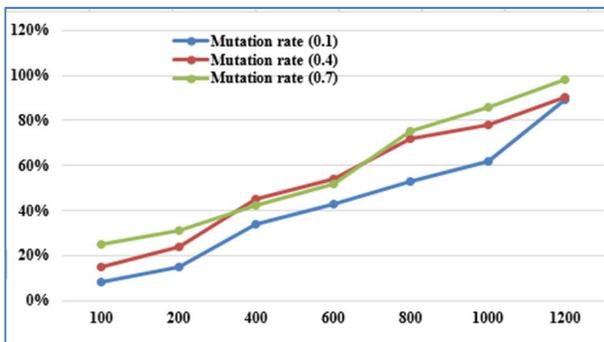


Figure 11. Attack detection accuracy vs. mutation rate (LD, selection rate=20%)

Our initial evaluation shows, that HD performs better than LD. Further, higher selection and mutation rates for crossing over chromosomes, the better the signatures are.

## VI. CONCLUSIONS AND FUTURE WORK

There are many types of attacks that attackers/criminal can use against a computer system. This paper develops a signature-based IDS for web services by addressing the limited attack signatures to test implemented web services. Our approach relies on GA to generate new attack signatures from a set of initial attack signatures. The approach currently addresses the generation of both simple and complex forms of attacks including SQL injection, XPath injection, and SOAP injection. Our future work includes applying multi point cross over, evaluated on a larger scale datasets. We plan to develop an anomaly based IDS in the future using metrics.

## REFERENCES

- [1] M. Vieira, N. Antunes, and H. Madeira, "Using web security scanners to detect vulnerabilities in web Services," Proc. of IEEE/IFIP International Conference on Dependable Systems Networks, DSN '09, pp. 566-571, June 2009.
- [2] S. Tiwari and P. Singh, "Survey of potential attacks on web services and web service compositions," Proc. of 3rd International Conference on Electronics Computer Technology (ICECT), 2011, Volume 2, pp. 47-51, April 2011.
- [3] C. Bartolini, A. Bertolino, E. Marchetti, and A. Polini, "Wstaxi: A wsdl-based testing tool for web services," Proc. of IEEE ICST, pp. 326-335, 2009.
- [4] S. Jan, C. Nguyen, L. Briand, "Automated and effective testing of web services for XML injection attacks," Proc. of Proceedings of the 25<sup>th</sup> International Symposium on Software Testing and Analysis (ISSTA) 2016, pp. 12-23
- [5] NIST, Guide to Secure Web Services, NIST Special Publication 800-95, August 2007, Accessed from <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-95.pdf>
- [6] T. Rosa, A. Santin, and A. Malucelli, "Mitigating xml injection 0-day attacks through strategy-based Detection systems," *IEEE Security Privacy*, Vol. 11(4), pp. 46-53, July 2013.
- [7] V. Patel, R. Mohandas, and A. Pais, "Attacks on web services and mitigation schemes," *Proc. of the 2010 International Conference on Security and Cryptography (SECRYPT)*, pp. 1-6, July 2010.
- [8] J. Chen, Q. Li, C. Mao, D. Towey, Y. Zhan, and H. Wang, "A web services vulnerability testing approach based on combinatorial mutation and soap message mutation," *Service Oriented Computing and Applications*, 8:1-13, 2014.
- [9] N. Havrikov, M. Hoschele, J. P. Galeotti, and A. Zeller, "XMLmate: Evolutionary XML test generation," *Proceedings of the 22<sup>nd</sup> ACM SIGSOFT International Symposium on Foundations of Software Engineering (FSE)*, pp. 719-722, NY, USA, 2014.
- [10] Wapiti, <http://wapiti.sourceforge.net/>
- [11] W3af, <https://github.com/andresriancho/w3af/>
- [12] S. Karumanchi, and A. Squicciarini, "In the Wild: a Large Scale Study of Web Services Vulnerabilities," *Proc. of ACM Symposium of Applied Computing*, Geyoungju, South Korea, 2014, pp. 1239-1246.
- [13] WebScarab, <https://github.com/OWASP/OWASP-WebScarab>
- [14] Wfuzz, <http://code.google.com/p/wfuzz/>
- [15] XML Security Cheatsheet, Accessed from [https://www.owasp.org/index.php/XML\\_Security\\_Cheat\\_Sheet](https://www.owasp.org/index.php/XML_Security_Cheat_Sheet)

- [16] Web Service Security Cheatsheet, Accessed from [https://www.owasp.org/index.php/Web\\_Service\\_Security\\_Cheat\\_Sheet](https://www.owasp.org/index.php/Web_Service_Security_Cheat_Sheet)
- [17] RESTFul Security Cheatsheet, Accessed from [https://www.owasp.org/index.php/REST\\_Security\\_Cheat\\_Sheet](https://www.owasp.org/index.php/REST_Security_Cheat_Sheet)
- [18] N. Antunes, N. Laranjeiro, M. Vieira, and H. Madeira, "Effective Detection of SQL/XPath Injection Vulnerabilities in Web Services," *Proc. of Services Computing*, 2009.
- [19] M. Najjar and M. Azgomi, "A distributed multi-approach intrusion detection system for web services," *Proceedings of the 3<sup>rd</sup> International Conference on Security of Information and Networks (SIN)*, pp. 238-244, 2010.
- [20] The structure of SOAP message, Accessed from [https://www.ibm.com/support/knowledgecenter/en/SSMKHH\\_9.0.0/com.ibm.etools.mft.doc/ac55780.htm](https://www.ibm.com/support/knowledgecenter/en/SSMKHH_9.0.0/com.ibm.etools.mft.doc/ac55780.htm)
- [21] SoapUI, (2016). Accessed from <https://www.soapui.org/security-testing/security-scans/xml-bomb.html>

# Mobile Business Performance Metrics: Framework and Case Study

Ahyoung Kim, Junwoo Lee  
Smart Content Research Section  
Electronics and Telecommunications Research Institute  
Daejeon, Korea  
{ kimay, leeju } @etri.re.kr

*Abstract*—The increases in mobile-phone users changed paradigms in mobile business and extended business area and its applications. Therefore, various attempts have been made to succeed in mobile business. In this study, performance indicators for measuring business performance are proposed, also using proposed performance indicators framework for analyzing usage-focused mobile business is suggested. Based on the characteristics of mobile business, performance indicators are classified to Customer Retention and Product Engagement. The customer's usage data automatically collected from mobile application is analyzed by 2 performance indicators. In this paper, we will suggest the mobile business analysis framework for performance indicators, and introduce the result of analysis through case-study. The suggestion and mobile business analysis framework from this study support decision making in mobile business area such as mobile marketing and mobile commerce.

*Keywords*—performance measure; performance metrics; usage analysis; mobile business

## I. INTRODUCTION

As the usage of mobile device and mobile data service increased, the paradigm of the IT industry has been changing rapidly in hardware-centered service to mobile-based software and application service. Since 2009 when iPhone was introduced and smartphones spread rapidly, the structure of IT industry is changed rapidly to mobile focused industrial structure [1-3].

Accelerating the spread of mobile devices and generalizing the usage of mobile internet, various new mobile services are appeared, and our lifestyle also is changed to mobile focused lifestyle. Through mobile we can connect to the web anytime, anywhere. Therefore, we can provide diverse services immediately at any time, in any location through various application of smartphone device [4-5]. Furthermore, based on the enormous speed and impact on other industries, the importance is increasingly growing.

Based on previous experience, infrastructure evolution regardless of the type and form has had a significant impact on overall economy industries. Therefore, the changes of IT industrial structure and lifestyle are reasonable.

Especially, companies that create value from customer's needs are greatly affected by the changes of business

environment caused by mobile-big bang. The companies need this changes and at the same time it can be new opportunities. However, companies that don't adapt correctly to this flow cloud fall behind in the market. It is very important for effective response. Applying former web-based e-business performance indicator such as PV (page view) or UV (unique visitor) to the changed market environment is difficult. Therefore, in this study, Customer Retention and Product Engagement are selected as mobile business performance indicators. Selected performance indicators should be aimed to optimize in mobile business, and the performances are measured with customer's usages. For this purpose, usage focused mobile business analysis framework that can analyze performance indicator is suggested, and using proposed framework the result of the experimental data is analyzed.

In chapter 2, based on literature review we will analyze performance indicator for E-business and the feature of the mobile business, and derive the mobile data analysis and the related implications of previous research. Based on this theoretical research, in Chapter 3, we will derive the performance indicators for mobile business. Considering derived performance indicators, in Chapter 4, we will design usage-focused mobile business analysis framework. In Chapter 5, we will analyze the result of experimental data. Finally, in Chapter 6, we will provide some concluding remarks regarding our proposal.

## II. RELATED WORK

### A. Mobile Business Performance Metrics

Mobile business is exchanging products, services and information using mobile technology [6]. Namely, mobile business includes business process that makes it possible to trade products, services, and information through mobile device [7]. In other word, mobile business it the administrative tool that adjusts and manages the communication of organization using mobile technology including wireless internet access [8]. Moreover, it could be understood as expended web-based electronic commerce, also E-commerce and E-business that provides service through mobile and wireless network are considered as part of business. [9].

---

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) [No. R0118-16-1005, Digital Content In-House R&D].

TABLE I. CHARACTERISTICS OF MOBILE BUSINESS

Characteristics		Description
<i>Primary Characteristics (Mobility)</i>	<i>Reachability</i>	Attributes which can be contacted anywhere anytime
	<i>Ubiquity</i>	Attributes which can fulfil the need both for real-time information and for communication anywhere
	<i>Convenience</i>	Attributes which are always at hand and are easy to use
	<i>Security</i>	Attributes which are provided authentication of the owner and enables a higher level security
<i>Secondary Characteristics (Localization)</i>	<i>Localization</i>	Attributes which can know where the users is physically located at any particular moment
	<i>Instant Connectivity</i>	Attributes which can be easier and faster to access information
	<i>Personalization</i>	Attributes which can be personalized to represent information or provide services in ways appropriate to a specific users

TABLE II. EXAMPLE OF PERFORMANCE METRICS FOR E-BUSINESS

Financial Metrics	User Metrics	Internal Process Metrics	Learning and Growth Metrics
Online revenue per user	Level of service delivery	Availability of systems	Staff Productivity and Morale
Cost per online user	Satisfaction of existing users	Volume of transactions processed	# of staff trained in new services
Cost-efficiency of Business process	# of new users reached	# of errors	Value delivery per employee

Therefore, mobile business can be defined as all activities to exchange products, services, information and etc. using mobile technology including wireless network with internal and external customers in various management activities. Mobile business has the feature of mobility and localization [10]. Table 1 presents the characteristics of Mobile Business.

E-business performance can be measured in several ways, ranging from Business contribution, user orientation, operational excellence, and ultimately, to future orientation [11-13]. Table 2 presents the example of performance metrics for E-Business.

### B. User Data Analysis Platform

The Flurry is the most well-known product and supports more features and platforms than the other products in spite of the free software. The Flurry analytics make deals with the diverse mobile OS including the BlackBerry and have the capability of collecting data from the mobile device by using the predefined by the Flurry. Additionally, the Flurry service is operated by the big data file system to take care of increasing the amount of data suddenly. The difference between the Flurry and our approach is that our approach can support the real time analysis and is custom-built with a secured data.

The MixPanel is commercial software and supports the complex query creation by using GUI editor instead of the CLI (Command Line Interface) [14-15]. Additionally, this system makes the people to satisfy the analysis result through adequate UI according with a characteristic of the user data. KSuite have a user acquisition analysis, custom events tagging & timeline analysis in general. For the mobile analysis, this system has a special capability like view by device, OS, and revenue tracking. The Country have similar analysis features

based on the statistics like the products mentioned before and have an excellent usability through many types of report creation [16]. The difference from our approach is that these systems are focused on the UI to enhance the functions already made before, but our approach is trying to make a new novel algorithm based on the big data system, cloud environment. Finally, both the Apple and the Google have supported the good analysis tool for the digital content app too. However, these tools have a handicap that these tools are only for the smart content app bought or downloaded in their system.

### III. MOBILE BUSINESS PERFORMANCE METRICS

Performance metrics is important to keep the successful E-Business. But the mobile business is still using the classic metrics such as UV (Unique Visitor) or PV (Page View) to measure the business performance in existing web market. The classic metrics are used to set up KPI (Key Performance Indicators) and to measure ROI (Return On Investment) based on the statistical data. Thus we need the performance metrics of new concept for the mobile environment. In the work, we set up the performance metrics for mobile business. We collect the raw data using mobile applications in order to measure performance metrics of mobile business. The collected raw data is categorized into three types of the static data, the dynamic data and the customer define data according to the data attribute. The Static Data includes data which can be collected only once such as UUID, application name, OS, OS version, device model and resolution. The Dynamic Data includes data that collect every customer activity such as start time, end time, start latitude, end longitude, and network. And the User Define Data includes additional data which is specific for the mobile application besides the Static Data and the Dynamic Data.

Table 3 presents the raw data which can be collected for performance metrics. Using the collected raw data, we calculate user-centric performance metrics for the mobile business. In this work, the performance metrics is categorized into two types of the customer retention and the product engagement according to the business characteristic.

TABLE III. PERFORMANCE MEASUREMENT SET

Data Type	Collected Raw Data
Static Data	UUID, App Name, OS, Model, Resolution
Dynamic Data	Start/End Time, Start/End Latitude, Start/End Longitude, Network
User Define Data	User Define Data

IV. MOBILE BUSINESS ANALYSIS FRAMEWORK

Figure 1 is the system architecture for mobile business analysis framework. The mobile business analysis framework consists of a client module for collecting customer usage data using SDK and a server module for processing and analyzing the collected customer usage data based on cloud.

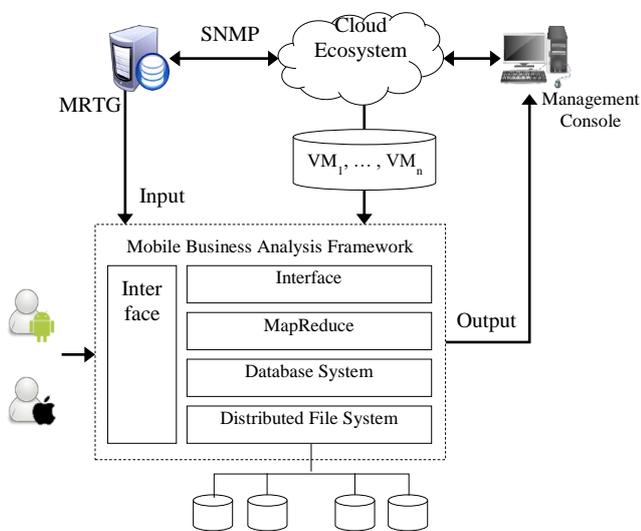


Figure 1. Usage Pattern Analysis for Message Application

The Mobile Business Analysis Framework is made based on a distributed file system with Hadoop to dynamically manage the resources in a cloud computing environment using a Simple Network Management Protocol such as Multi Router Traffic Grapher (MRTG). For monitoring a certain resource, we generally have to use not only a cloud management console but also a network management system like MRTG and the system manager has a difficult time checking and easily observing the digested information about the resources in a cloud computing environment. To investigate this problematic situation in a cloud computing environment, we designed and deployed the cloud service infrastructure based on CloudStack, which is open-source software from Citrix. The model presented in this paper regularly stores the usage data for

computing resources in a cloud service infrastructure with a distributed file system based on Hadoop and HBase using the SNMP among Virtual Machines (VMs).

The mobile business analysis framework has a flexible architecture to satisfy the dynamic requests of users for an analyzed service. First, the mobile business analysis framework collects the usage frequency for the mobile product out of mobile devices that have installed the library for the analysis service. At this time, the data used for the identification of the user is ignored because most people are currently worried that Google and Apple have collected data related to them. Second, the usage frequency is analyzed based on many factors such as the location and device. The frequency of mobile apps based on time is calculated again according to the hour of the day, day of the week, and month of the year because the framework has to generate corresponding candidates for diverse user contexts. The frequency of mobile apps based on the location and device is calculated again, such as the time according to the latitude/longitude, the framework name, and the manufacturer. Additionally, this framework focuses on how long the mobile app is used, along with its frequency of use, as YouTube and Video Player applications are usually used for a long time after being executed. Finally, this framework considers the recentness of the mobile app as an important factor to determine what mobile apps are novel for users because the lifecycle of a mobile app is very short, and users consistently seem to want to use up-to-date applications.

$R$  denote the list of candidates considered for the reliability of the mobile app based on factors such as the time, location, and device.

$$R = \{ R_{time}, R_{location}, R_{device} \}. \tag{1}$$

where  $R_{time}$  is an ordered list of mobile app based on time,  $R_{location}$  is a list of mobile apps based on the location, and  $R_{device}$  is a list of mobile apps based on the device.

$R_{time}$  is an ordered list of mobile apps analyzed by the model using a collaborative filtering algorithm based on the hour of the day instead of the user. This means that  $R_{time}$  consists of the mobile apps selected and ordered by the algorithm, which analyses the usage frequency, duration, and recentness of the mobile app considering the relative importance based on the hour of the day. The others,  $R_{location}$  and  $R_{device}$ , are the means of filtering the candidate applications as an intermediate result from an ordered list selected according with the  $R_{time}$ . In other words, the model can find which mobile apps are good for the users out of the intermediate results by using the location and device. These two parameters are configured based on the users' decisions and are acceptable for the model to determine which mobile app is the best candidate. The model determines which factors the users want to use for the analysis service based on the user context, and sets a value for the variables  $\alpha$ ,  $\beta$  and  $\gamma$ . For example, if the users want to use the location data to find the mobile applications, the model collects the mobile applications based on the current location with a list of candidates already made based on the  $R_{time}$  and  $R_{device}$ .

Let  $U$  denote that the user set consists of the usage frequency per hour, usage duration per hour, usage time per hour, day of the week, month of the year, the operating system version, and the location.

$$U = \{UFT, UDT, URT, DOW, MOY, OSV, L\}. \quad (2)$$

TABLE IV. PERFORMANCE MEASUREMENT SET

Type	Description
UFT	Set for Usage Frequency for Mobile App
UDF	Set for Usage Duration for Mobile App
URT	Set for Usage Recentness for Mobile App
DOW	Set for Day of Week for Execution of Mobile App
MOY	Set for Month of Year for Execution of Mobile App
OSV	Set for Version of OS for Installation of Mobile App
L	Set for Location based on Latitude/Longitude for Mobile App

### V. CASE STUDY

In this paper, we collected the data to evaluate the model explained in the previous chapter over a three-month period using the mobile devices of our team members. In particular, this evaluation is for Google Android OS because the development of the data-gathering algorithm with a client library is relatively easier than for Apple iOS.

The method used to analyze the usage frequency for the mobile apps based on the data collected out of the mobile devices was introduced in the previous chapter. As mentioned before, we determined the standards to measure how many times mobile apps are used by the hour instead of by the day or month because the general usage patterns for the mobile apps fluctuate severely. We collected 120 mobile apps, excluding some applications such as middleware applications from Google and the manufacturer because these applications are useless for collecting user context data.

Generally, the importance of the mobile apps will be determined by the how many people use a certain of the mobile app within a period of the time. In this paper, we add some features to normalize the difference between the usage patterns by an hour in a day in addition to this general characteristic. First of all, we have measured the relationship between an hour in a day based on the usage frequency and the usage duration to determine what time is really important for all mobile apps. And we have analyzed and calculated the objective similarity value of all mobile apps based on the measurement for the relationship between hours in a day.

This normalization process helps us to avoid misunderstand the importance of the mobile apps based on the total frequency number and the total amount of the time are collected. The reason is that the big number of the frequency of the mobile app in a certain of an hour in a day does not mean the importance of the mobile app in comparison with the

number of the frequency of the other hour in a day. For example, a certain of the mobile app was invoked one thousand times at an hour in a day having ten thousand times. The other mobile app was collected twenty times at an hour in a day with forty times execution.

#### A. Performance Measure of Message Types

Figure 2 shows the usage pattern for many apps related with message that are the most famous mobile app in South Korea. As you can see, the graph of the real data for the usage frequency fluctuates heavily according with the time dimension. So the graph for the average value does. However, the graph for the model explained in this paper shows that there is at a high level in spite of the distribution of the real data because the analysis model normalizes the difference between the hours in a day and recalculates the similarity by using the relative weight for the mobile app. The graph based on the usage duration shows that the analysis model in this paper suggests which time is important for this mobile app.

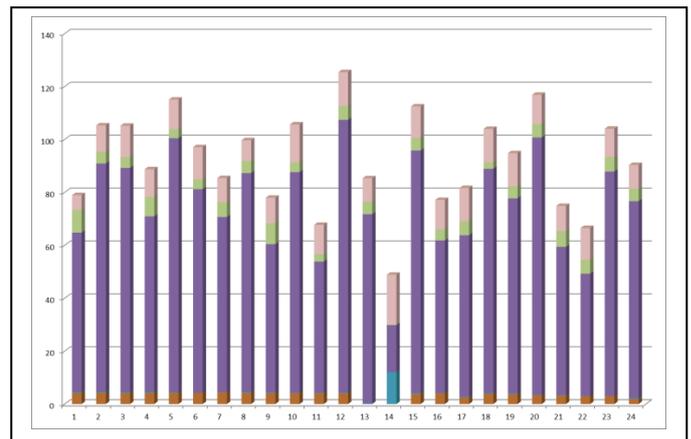


Figure 2. Usage Pattern Analysis for Message Application

#### B. Performance Measure of SNS Types

SNS applications are the most popular applications because mobile devices are consistently connected to other people based on the network environment. Figure 3 shows the usage pattern for SNS applications. SNS applications are popular in that many people want to check and see what kinds of activities other people are currently doing. In addition, the usage frequency and duration of SNS applications are relatively higher than other kinds of mobile applications.

#### C. Relationship between the hours in a day

Figure 4 shows that the relationship between the hours in a day based on the usage frequency of the mobile apps for this experiment. In the commuting time, the similarity of the times in the commuting time is higher than the other times in a day such as the night and the dawn. Similarly, the similarity of the times in a working time is higher than the other times in a day. These phenomena show that we have to shorten the usage data gap between the hours in a day and normalize the real data by using this similarity result. After the normalization process, we

have evaluated the usage pattern based on the frequency and the duration for the usage of the mobile app all candidates within our experiment environment. We have analyzed the equation introduced in this paper by comparing with the two kinds of measure that is a real data and an average value at an hour in a day.

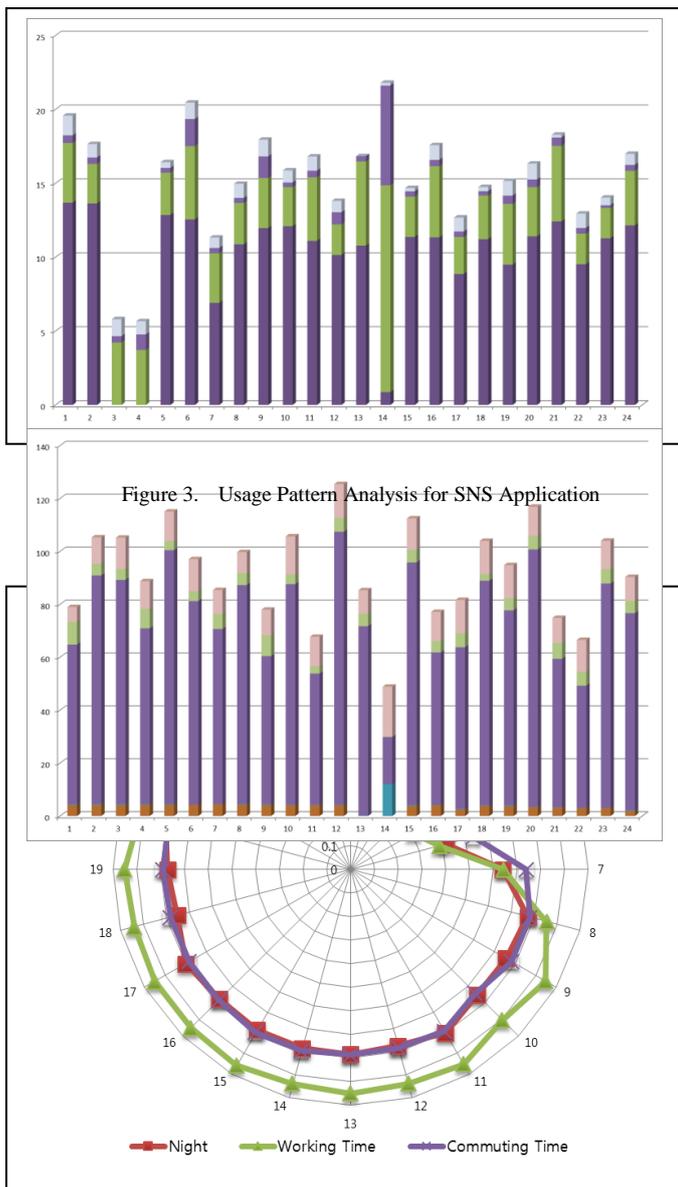


Figure 4. Relationship between the hours in a day

## VI. CONCLUSION

In this paper, we proposed usage-centric mobile business analysis framework via performance metrics and showed a case study with test data. The user data analysis platform consists of a client library and server part based on a large data system. The client library has the main responsibility for collecting user data from the mobile devices and sending these data to the server part of the analysis platform. The server part

analyzes the received data using a predefined data pattern and stores the data in a large data system.

Finally, we conducted an evaluation of the mobile business analysis framework using some data generated by collecting the log data for mobile applications from mobile devices. This evaluation allowed us to make an educated guess of the user ratings instead of using the user profiles or user ratings directly. Additionally, the results of this evaluation show that the popularity of mobile applications can be checked by the hour, and we have to consider a lot of information such as the location, operating system version, day of the week, and time of day, in addition to the usage frequency data for mobile applications. The suggestion and mobile business analysis framework from this study support decision making in mobile business area such as mobile marketing commerce.

## ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) [No. R0118-16-1005, Digital Content In-House R&D].

## REFERENCES

- [1] Bremser, W.G., and Chung, Q.B. (2005). A framework for performance measurement in the e-business environment, The 6th International Conference on Electronic Commerce, 4(4), (pp. 395-412). Electronic Commerce Research and Applications.
- [2] Girardello, A., and Michahelles, F. (2010). AppAware: Which mobile applications are hot?. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (pp. 431-434). ACM.
- [3] Davidsson, C., and Moritz, S. (2011, February). Utilizing implicit feedback and context to recommend mobile applications from first use. In Proceedings of the 2011 Workshop on Context-awareness in Retrieval and Recommendation (pp. 19-22). ACM.
- [4] Böhmer, M., Ganev, L., and Krüger, A. (2013, March). Appfunnel: A framework for usage-centric evaluation of recommender systems that suggest mobile applications. In Proceedings of the 2013 international conference on intelligent user interfaces (pp. 267-276). ACM.
- [5] Lee, K. H., Lee, Y. J., Choi, H., Chung, Y. D., & Moon, B. (2012). Parallel data processing with MapReduce: a survey. ACM SIGMOD Record, 40(4), 11-20.
- [6] Paavilainen, J. (2002). Mobile Business Strategies: Understand the Technologies and Opportunities, Addison-Wesley Professional, London.
- [7] Kalakota, R. and Robinson, M. (2002). M-Business: The Race to Mobility, McGraw-Hill, New York.
- [8] Scornavacca, E., and Barnes, S. J. (2004, May). M-banking services in Japan: A Strategic Perspective, International Journal of Mobile Communications, 2(1), (pp. 51-66). ACM.
- [9] Mylonopoulos, N. A. and Doukidis, G. I. (2003). Introduction to the Special Issue: Mobile Business: Technological Pluralism, Social Assimilation, and Growth, International Journal of Electronic Commerce, Vol. 8, (pp. 5-22). ACM
- [10] Müller-Veerse, F. (1999, November). Mobile Commerce Report, Dulacher Corp., London.
- [11] Van Grembergen, W., and Amelincx, I. (2002), Measuring and managing e-business projects through the balanced scorecard, System Sciences, 2002. HICSS. Proceedings of the 35th Hawaii International Conference on System Sciences. IEEE
- [12] Van Grembergen, W., and Saull, R. (2001), Information Technology Governance through the Balanced Scorecard. In Proceedings of the 34th

Hawaii International Conference on System Sciences (HICSS), CD-ROM. Maui.

- [13] Rosemann, M. (2001), "Evaluating the Management of Enterprise Systems with the Balanced Scorecard", in Van Grembergen, W., Information Technology Evaluation Methods and Management, Hershey (PA): Idea Group Publishing, (pp.171-184).
- [14] Woerndl, W., Schueller, C., and Wojtech, R. (2007). A hybrid recommender system for context-aware recommendations of mobile applications. In Data Engineering Workshop, 2007 IEEE 23rd International Conference on (pp. 871-878). IEEE.
- [15] Karatzoglou, A., Baltrunas, L., Church, K., and Böhmer, M. (2012, October). Climbing the app wall: enabling mobile app discovery through context-aware recommendations. In Proceedings of the 21st ACM international conference on Information and knowledge management (pp. 2527-2530). ACM.
- [16] Böhmer, M., Ganev, L., and Krüger, A. (2013, March). Appfunnel: A framework for usage-centric evaluation of recommender systems that suggest mobile applications. In Proceedings of the 2013 international conference on intelligent user interfaces (pp. 267-276). ACM.
- [17] Adabala, S., Chadha, V., Chawla, P., Figueiredo, R., Fortes, J., Krsul, I., & Zhu, X. (2005). From virtualized resources to virtual computing grids: the In-VIGO system. Future Generation Computer Systems, 21(6), 896-909.
- [18] Wang, L., Von Laszewski, G., Younge, A., He, X., Kunze, M., Tao, J., & Fu, C. (2010). Cloud computing: a perspective study. New Generation Computing, 28(2), 137-146.

## **Session 5: Internet Application and Technology**

Title: Evaluate action primitives for Human Activity Recognition using unsupervised learning approach  
(Authors: Luis F. Mejia-Ricart, Paul Helling, Aspen Olmsted)

Title: Classification of Music by Composer Using Fuzzy Min-Max Neural Networks  
(Authors: Pasha Sadeghian, Casey Wilson, Stephen Goeddel, Aspen Olmsted)

Title: Assessment of Fuzzy Min-Max Neural Networks for Classification Tasks  
(Authors: Pasha Sadeghian, Aspen Olmsted)

Title: Low-cost Detection of Backdoor Malware  
(Authors: Huicong Loi, Aspen Olmsted)

Title: Docker vs. KVM: Apache Spark application performance and ease of use  
(Authors: Walter Blair, Aspen Olmsted, Paul Anderson)

# Evaluate Action Primitives for Human Activity Recognition using Unsupervised Learning Approach

Luis F. Mejia-Ricart, Paul Helling, Aspen Olmsted

Computer Science Department

College of Charleston

Charleston, SC, United States

mejia-ricartlf@g.cofc.edu, hellingp@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—Smartphones and wearable devices are in the frontlines when it comes to the field of Human Activity Recognition (HAR). There have been numerous attempts to use motion sensors in smartphones and wearables to recognize human activity. Most of these studies apply supervised learning techniques, which requires them to use labeled datasets. In this work, we take a sample of these labels, or action primitives (sit, stand, run, walk, jump, lie down), and evaluate them against the resulting labels of several clustering algorithms. We built two datasets (labeled and unlabeled) using accelerometer, gyroscope, and pedometer readings from two fixed-position devices, a smartphone in the side pocket, and a smartwatch strapped onto the left-hand wrist. Ultimately, we want to determine whether these action primitives commonly used in HAR are optimal, and suggest a better set of primitives if not.

*Keywords*—component; sensors; smartphones human activity recognition;

## I. INTRODUCTION

Utilizing smart devices' sensor data to draw better-informed conclusions regarding the user is an increasingly common practice among mobile app developers. Developers use this data to deliver tailored suggestions, automate tasks, and make automated decisions in and on behalf of the user. The field of Human Activity Recognition (HAR) can take advantage of all this data. Albeit most work on smartphone or wearable-based HAR relies only on input data from accelerometers, we consider other variables such as gyroscope and pedometer sensor data.

In HAR, most works use supervised learning because the prediction they're making is general enough to label. In this experiment, we want to analyze data hoping to find more specific traces of activity. The purpose of this unsupervised approach is to reveal patterns in data that could be used as activity primitives for further, higher level activity recognition.

The organization of this paper is as follows. Section II describes the related work. Section III discusses the methodology we used for the current research. Sections IV documents our findings, and Section V gives our conclusions and discussed future work.

## II. RELATED WORK

When making predictions, it is necessary to analyze historical data. This is also true when predicting smartphone user activity recognition. However, if the data is too old, it might already be out of context. This means we must determine a sample data window size. Erdaş Berke et al . [1] used frames of 1 and 4 seconds to determine user activity with 50% overlap at a sampling rate of 0.86 Hz. Abayomi Moradeyo Otebolaku and Maria Teresa Andrade [2] also used a 50% overlap at a 50 Hz sampling rate. Yuan Yuan et al . [1] used a 50% overlap at a 100 Hz sampling rate. Yongjin Kwon et al . [3] used a 50% overlap at a 50 Hz sampling rate in sliding windows of size 0.25 – 7 seconds, and reported their best results were using 0.25 to 0.5-second windows. Using a 50% overlap is seemingly standard, but there is no consensus on the sampling rates and sliding window size.

## III. METHODOLOGY

### A. Data gathering

We collected the sensor data using a smartphone in the left side pocket and the smartwatch strapped onto the left-hand wrist. Both devices were configured to collect accelerometer, gyroscope, and pedometer readings. The sampling rate for accelerometer readings was 20 Hz, amounting to one sample every 50ms. For the gyroscope, readings were collected every 100 ms (10 Hz). and Our sensor data collection sampling rate and sliding window size are still to be determined. Overlapping frames at 50%, allowing each sample to be evaluated twice, like previous work on the field [4, 2, 3]. Sensor positioning will be wrist-fixed for the smartwatch and fixed positioning in the left side pocket for the smartphone.

### B. Feature extraction

One sensor sample determines nothing; what matters is the values of readings over time. As accustomed in related works in the field of HAR, we aggregated the data to gain a more insightful set of data. Previous works used a window length of 1-4 seconds [3, 4]. Álvarez de la Concepción et al. [5] determined the optimal length of each window is 5 seconds. Here, we use a window length of 2 seconds (~40 samples at 20 Hz). Each sliding window has a 50% overlap with the previous window and a 50% overlap with the next window. Also common practice in data collection for HAR, assuring that

every reading will be evaluated twice (with previous readings; with later readings).

We extracted a total of 114 features from the time domain. For every sensor data sliding window, we aggregate the data from the samples to derive the features from the list below. Each feature appears twice in the final feature vector, once the smartphone readings and once for the smartwatch readings.

- Mean for X, Y, and Z values (Accelerometer, Gyroscope)
- Variance of X, Y, and Z values (Accelerometer, Gyroscope)
- Standard deviation of X, Y, and Z values (Accelerometer, Gyroscope)
- Median for X, Y, and Z values (Accelerometer, Gyroscope)
- Index of min/max readings in sliding window for X, Y, and Z values (Accelerometer, Gyroscope)
- Range (max-min) for X, Y, and Z values (Accelerometer, Gyroscope)
- Mean Magnitude (Accelerometer, Gyroscope)
- Total steps detected in window (Pedometer)

### C. Classification

We used 6 different clustering algorithms:

- **K-means**, where  $k = 6$
- **Spectral Clustering**, where  $k = 6$
- **Hierarchical clustering**, once using Ward's method and another using average linkage.
- **DBSCAN** and **Mean Shift** as a shot in the dark to find otherwise not considered clusters

## IV. RESULTS

When compared to our labeled dataset, the results obtained from hierarchical clustering using average linkage, DBSCAN, and Mean Shift did not match to any of our action primitives. DBSCAN was too strict, disregarding many sample windows as noise and failing to come up with a meaningful cluster. Average Linking and Mean Shift failed to come up with clusters similar enough to be considered.

Fig. 1 shows a 3-dimensional scatter plot of the results obtained from Spectral Clustering, K-means, hierarchical clustering using Ward's method, and our labeled dataset. For graphing purposes, we selected 3 arbitrary variables to compare. The resulting clusters from these algorithms exhibit a lot of similarities to our labeled dataset. More specifically, Spectral Clustering and Ward's method were both particularly decisive, while K-means showed noisier results and clusters interweaving.

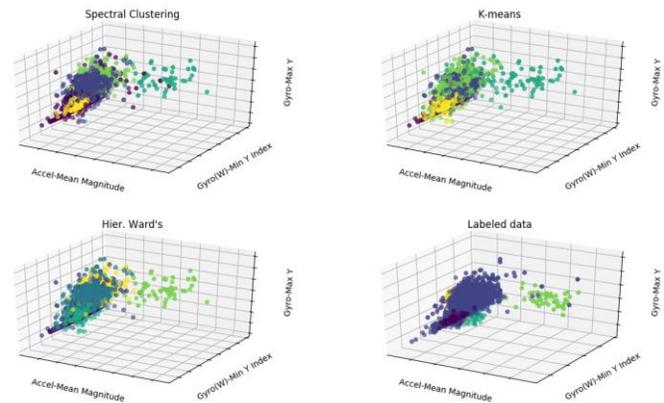


Figure 1. Visualization of data with color-coded clusters for all algorithms. In order: Spectral Clustering, K-means, Ward's, and our labeled dataset

### Cluster sizes

The first 3 labels of our labeled dataset correspond to sitting, walking, and standing idly. The last 3 correspond to running or trotting, lying down, and jumping, respectively.

[4365, 2219, 1793, 88, 77, 49] (Labeled data)

The resulting cluster sizes are as follows

[6452, 1234, 565, 259, 52, 29] (Spectral C.)

[6128, 1032, 947, 222, 206, 56] (Ward's)

[4462, 1672, 976, 975, 341, 165] (K-means)

## V. CONCLUSIONS

We have disregarded the results from our average linking and DBSCAN algorithms as they appear meaningless for now but may include a tweaked version of these algorithms in future work.

When compared to our labeled dataset, the results of our clustering algorithms Spectral Clustering, K-means, and hierarchical clustering using Ward's method appear to reinforce the use of our selected action primitives. The resulting cluster sizes from these three algorithms show that Spectral C. and Ward's method may have grouped together samples from the lying down, standing idly, and sitting down due because they seemed idly. To determine what all the clusters mean, we must further examine cluster data, considering each feature variable.

## REFERENCES

- [1] K. A., H. N., M. S. and P. T., "Optimising sampling rates for accelerometer-based human activity recognition," *Pattern Recognition Letters*, 2016.
- [2] S. González, J. Sedano, J. R. Villar, E. Corchado, A. Herrero and B. Baroque, "Features and models for human activity recognition," *Neurocomputing*, 2015.
- [3] Y. Liu, L. Nie, L. Liu and D. S. Rosenblum, "From action to activity: Sensor-based activity recognition," *Neurocomputing*, 2015.

- [4] L. Liu, Y. Peng, M. Liud and Z. Huange, "Sensor-based human activity recognition system with a multilayered model using time series shapelets," *Knowledge-Based Systems*, 2015.
- [5] A. d. l. Concepción, S. Morillo, Gonzalez-Abril and O. Ramírez, "Discrete techniques applied to low-energy mobile human activity recognition. A new approach," *Expert Systems with Applications*, 2014.
- [6] Y. Kwon, K. Kang and C. Bae, "Unsupervised learning for human activity recognition using smartphone sensors," *Expert Systems with Applications*, 2014.
- [7] A. M. Otebolaku and M. T. Andrade, "User context recognition using smartphone sensors and classification models," *Journal of Network and Computer Applications*, 2016.
- [8] C. Erdaş, I. Atasoy, K. Açıcı and H. Oğul, "Integrating features for accelerometer-based activity recognition," in *The 3rd International Symposium on Emerging Information, Communication and Networks (EICN 2016)*, 2016.
- [9] Y. Yuan, C. Wang, J. Zhang, J. Xu and M. Li, "An Ensemble Approach for Activity Recognition with Accelerometer in Mobile-phone," in *2014 IEEE 17th International Conference on Computational Science and Engineering*, 2014.

# Classification of Music by Composer Using Fuzzy Min-Max Neural Networks

Pasha Sadeghian, Casey Wilson, Stephen Goeddel, Aspen Olmsted

Department of Computer Science  
College of Charleston  
Charleston, USA

pashasadeghian@gmail.com, cawilson1@g.cofc.edu, goeddelsm@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—This work utilizes high-level musical features extracted from a large music database of Sonata pieces composed by Beethoven, Corelli, and Mozart, and assesses the accuracy of Fuzzy Min-Max (FMM) Neural Network and Enhanced Fuzzy Min-Max (EFMM) Neural Network classifiers in classifying the classical pieces by composer. Results of the assessment are provided and show different accuracies depending on the parameters used in the FMM and EFMM models. This study presents a novel approach to the classification of music by composer by presenting two classifiers, namely FMM and EFMM Neural Networks, capable of classifying classical music by composer.

**Keywords**—Fuzzy Min-Max Neural Network; Enhanced Fuzzy Min-Max Neural Network; Classification of Music by Composer

## I. INTRODUCTION

Until recently the task of classifying classical music by composer had been reserved for human music experts. With the advancement of artificial intelligence and machine learning, this task is no longer reserved for human experts.

The goal of this paper is to assess the accuracy of the Fuzzy Min-Max (FMM) Neural Network and Enhanced Fuzzy Min-Max (EFMM) Neural Network classifiers in classifying classical music by composer. The question tackled in this research is “Can a machine using FMM and EFMM classification models accurately recognize who composed a musical piece?”. This work introduces two classification models that can accurately classify musical pieces in groups of three composers.

Section 2 presents the Fuzzy Min-Max Neural Network (FMM) and Enhanced Fuzzy Min-Max Neural Network classification methods which were used for music classification. In Section 3, the assessment methodology is described: the data used and how the samples were generated. In Section 4, the classification results are detailed.

## II. FUZZY MIN-MAX NEURAL NETWORK

The FMM learning algorithm is a three-step process, specifically, hyperbox expansion, hyperbox overlap test, and hyperbox contraction. Learning in FMM beings by using a data set comprising of input samples and target classes,  $A_h$ ,  $h = 1, \dots, N$ , where  $N$  is the total number of data samples. Based on the data samples, FMM generates a number of hyperboxes. Each

hyperbox is represented by a minimum point ( $v_j$ ) and maximum point ( $w_j$ ) in a  $n$ -dimensional space within a unit hypercube ( $I^n$ ). Fig. 1 illustrates a 3-dimensional hyperbox with its minimum point ( $v_j$ ) and maximum point ( $w_j$ ).

Each hyperbox fuzzy set is defined as [1]:

$$B_j = \{A_h, v_j, w_j, f(A_h, v_j, w_j)\} \quad \forall A_h \in I^n \quad (1)$$

where  $B_j$  is the hyperbox fuzzy set,  $A_h = (a_{h1}, a_{h2}, \dots, a_{hn})$  is the input sample, and  $v_j = (v_{j1}, v_{j2}, \dots, v_{jn})$  is the minimum point for  $B_j$  and  $w_j = (w_{j1}, w_{j2}, \dots, w_{jn})$  is the maximum point for  $B_j$ .

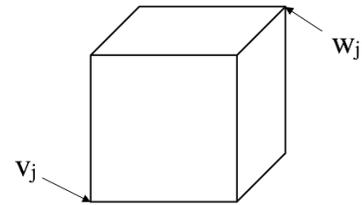


Figure 1. 3-D Hyperbox

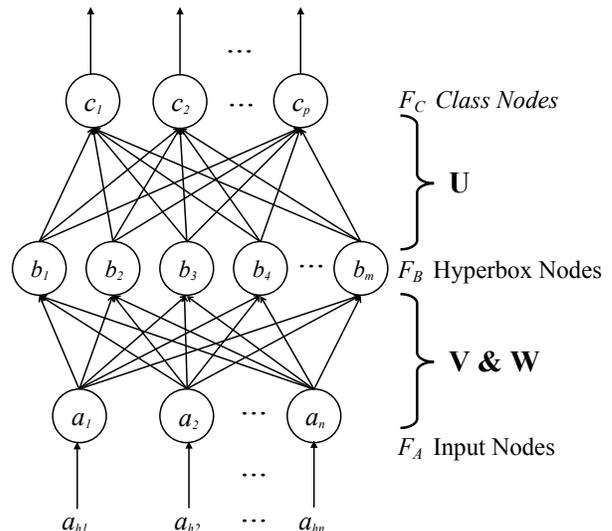


Figure 2. Three-Layer FMM Neural Network

If an input sample is contained in a hyperbox, it is said that the input sample is a full class member of the respective hyperbox. The size of the hyperbox is controlled by a user-defined parameter referred to as the hyperbox expansion coefficient ( $\Theta$ ). The FMM Neural Network has three layers. Fig. 2 illustrates the three-layer structure of the FMM. The input layer  $F_A$  has  $n$  nodes equal to the number of input features. The hyperbox layer  $F_B$  consists of  $m$  sets equal to the number of hyperboxes created during the learning process. The connection between  $F_A$  and  $F_B$  are saved in two matrices ( $V$  and  $W$ ), with the membership function being  $F_B$ 's transfer function [2]. The output layer  $F_C$  has  $c$  number of nodes, each representing a class. The output of each  $F_C$  node represents the degree to which the  $h$ th input pattern  $A_h$  matches an output class. The connections between  $F_B$  and  $F_C$  are binary values and are stored in matrix  $U$ . The connections are defined by

$$u_{jk} = \begin{cases} 1, & \text{if } b_j \text{ is a hyperbox for class } C_k \\ 0, & \text{Otherwise} \end{cases}$$

where  $b_j$  is the  $j$ th  $F_B$  node and  $C_k$  is the  $k$ th  $F_C$  node [1]. For soft decisions, the outputs can be use directly [1]. For hard decisions, the  $F_C$  node with the highest value is chosen as the predicted class [1].

In general, when a new training sample is presented, FMM utilizes the fuzzy membership function, to find the nearest hyperbox matching the sample. The fuzzy membership value of the training sample with respect to the  $b_j$  hyperbox can range from 0 to 1 (inclusive). The sensitivity of the fuzzy membership function is controlled by a user-defined parameter called sensitivity or measure of fuzziness ( $\gamma$ ).

During the hyperbox overlap test step, the FMM determines if any overlap exists between hyperboxes representing different classes. During the hyperbox contraction step, the existing hyperbox overlaps are eliminated by minimally adjusting each of the hyperboxes.

The hyperbox size is controlled differently under FMM and EFMM. Due to the multi-dimensional nature of the input pattern in this work, the effect of the difference in the treatment of the hyperbox size parameter under the two classification models can be significant in terms of average accuracy, standard deviation of average accuracy, as well as the number of hyperboxes. FMM and EFMM also differ in their hyperbox overlap test and their treatment of the overlapping hyperboxes in the hyperbox contraction step. The details of the hyperbox overlap test step and hyperbox contraction step under FMM and EFMM can be found here [2] and [1]. That being said, the description of the FMM in this paper also applies to EFMM.

### III. RELATED WORK

In a study conducted in 2001, Pollastri and Simoncelli employed a dataset of 605 musical pieces composed by Mozart, Beethoven, Dvorak, Stravinsky, and Beatles [3]. In their study, Pollastri and Simoncelli assessed the accuracy of Hidden Markov Models (HMMs) for classifying the musical pieces and compared their results (using HMMs) with the results obtained from music amateurs and music experts [3]. They reported the following accuracies: 42% for HMMs, 24.6% for music amateurs, and 48% for music experts [3].

In a study conducted in 2010, Kaliakatsos-Papakostas, Epitropakis, and Vrahatis focused on musical pieces written by composers Bach, Beethoven, Brahms, Chopin, Handel, Haydn, and Mozart [4]. In their study, Kaliakatsos-Papakostas, Epitropakis, and Vrahatis first used Probabilistic Neural Networks (PNNs) to discover similarities between composers and then used Feedforward Neural Networks (FNNs) to classify the pieces by composer [4]. This study used a dataset consisting of 350 musical pieces in MIDI format with the works transcribed for piano [4].

In two studies conducted in 2015, Herremans, Martens, and Sorensen employed a dataset of 1045 musical pieces, in MIDI format, composed by Bach, Haydn, and Beethoven [5] [6]. The study presented four classifiers, namely, C4.5 Decision Tree, RIPPER ruleset, Logistic regression, Naïve Bayes, and Support vector machines capable of classifying classical music by composer [5] [6]. The reported accuracies ranged from 80% to 86% [5] [6].

### IV. ASSESSMENT METHODOLOGY

In many ways, our work followed the work of Herremans, Martens, and Sorensen [5] [6]. We focused on Sonata piano pieces and selected three composers with a large number of Sonata pieces for inclusion in our train and test datasets. With our emphasis on assessing the accuracy of the FMM and EFMM models in classifying music by composer, we found it appropriate to select musical pieces from the same genre (i.e. Sonata). This way the classifiers would theoretically recognize the composers based on differences in musical style and not the differences in genre.

#### A. Dataset

417 Sonata pieces were selected from the composers Beethoven, Corelli, and Mozart. The pieces were selected from the KernScores database [7]. Table I presents a breakdown of our dataset by the composer.

TABLE I. DATASET

Composer	Number of Pieces
Beethoven Pieces	102
Corelli Pieces	244
Mozart Pieces	71

The data collected consisted of files in MIDI format containing detailed note information. Symbolic music representations, such as MIDI, comprise high-level structured information about the music, e.g., which note is played by which instrument [6].

The software used to extract features was jSymbolic [8]. jSymbolic is a Java-based Open Source software, the software allows for easy extraction of high-level features from MIDI files [6] [8].

Following the works of Herremans, Martens, and Sorensen, the following 12 features were selected for classification tasks [5] [6]:

TABLE II. ANALYZED FEATURES

Feature	Feature Description
Chromatic Motion	The fraction of melodic intervals corresponding to a semi-tone.
Melodic Fifths	The fraction of melodic intervals that are perfect fifths.
Melodic Octaves	The fraction of melodic intervals that are octaves.
Melodic Thirds	The fraction of melodic intervals that are major or minor thirds.
Most Common Melodic Interval Prevalence	Fraction of melodic intervals that belong to the most common interval.
Most Common Pitch Class Prevalence	Fraction of Note Ons corresponding to the most common pitchclass.
Most Common Pitch Prevalence	Fraction of Note Ons corresponding to the most common pitch.
Relative Strength of Most Common Intervals	Fraction of melodic intervals that belong to the second most common interval divided by the fraction of melodic intervals belonging to the most common interval.
Relative Strength of Top Pitch Classes	The frequency of the 2nd most common pitch class divided by the frequency of the most common pitch class.
Relative Strength of Top Pitches	The frequency of the 2nd most common pitch divided by the frequency of the most common pitch.
Repeated Notes	The fraction of notes that are repeated melodically.
Stepwise Motion	Fraction of melodic intervals that corresponded to a minor or major second.

Pitch refers to an absolute pitch, e.g., A in the 6th octave.  
Pitch class refers to a note without the octave, e.g., A.

Samples were randomly selected from the dataset of MIDI files. The samples were selected using Python and were handled by NumPy on Python [9] [10] [11].

*B. Sensitivity or Measure of Fuzziness ( $\gamma$ ) and Hyperbox Expansion Coefficient ( $\theta$ )*

In our tests, we used the following measures of fuzziness  $\gamma = \{0.1, 0.25, 0.5, 1, 2, 5, 10, 15, 20\}$  and the following hyperbox expansion coefficients  $\theta = \{0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 1, 5, 10, 15, 20\}$ . With our focus on the accuracy of the FMM and EFMM classifiers in classifying the musical pieces by composer, we decided to omit the results of hyperbox expansion coefficients larger than 0.25 (due to poor performance in terms of accuracy and high standard deviation of accuracies), we also decided to omit the results of hyperbox expansion coefficients smaller than 0.025 (due to lack of improvement in performance). Larger hyperbox expansion coefficients, keeping other variables constant, result in fewer number of hyperboxes, and fewer number of the hyperboxes result in time saved in training and testing.

*C. Accuracy*

We divided our dataset into train sets and test sets. For each composer, 70% of the pieces were randomly selected and allocated to the train set and the remaining 30% were allocated to the test set. This process was repeated for each test conducted. Table III illustrates the size of our train and test samples.

TABLE III. TEST AND TRAIN SAMPLE SIZE

Test Sample Size	Train Sample Size
130	287

Ten tests were conducted for each combination of hyperbox expansion coefficient (theta) and sensitivity or measure of fuzziness (gamma). Separate tests were conducted for the FMM and EFMM classifiers. For each test, we measured the percentage of correct classifications and reported the average accuracies along with other useful statistics. All tests were run using Python and NumPy [9] [10] [11].

V. RESULTS

*A. FMM Results*

TABLE IV. FMM RESULTS

Theta	Gamma	Average of Accuracy	StdDev of Accuracy	Number of Hyperboxes
0.025	0.1	68.77%	3.44%	242
0.025	0.25	69.15%	1.79%	242
0.025	0.5	69.38%	1.61%	242
0.025	1	67.77%	2.42%	241
0.025	2	68.77%	2.03%	243
0.025	5	69.69%	3.44%	242
0.025	10	70.46%	2.44%	242
0.025	15	67.38%	3.27%	239
0.025	20	68.46%	2.88%	240
0.05	0.1	61.85%	4.04%	108
0.05	0.25	63.62%	4.63%	112
0.05	0.5	64.46%	4.68%	111
0.05	1	63.62%	4.83%	109
0.05	2	63.85%	3.87%	108
0.05	5	63.54%	4.31%	109
0.05	10	62.62%	6.89%	109
0.05	15	62.00%	6.27%	110
0.05	20	62.69%	5.39%	108
0.1	0.1	49.46%	10.99%	28
0.1	0.25	55.15%	10.48%	28
0.1	0.5	52.38%	8.83%	26
0.1	1	57.38%	8.62%	28
0.1	2	48.15%	9.46%	26
0.1	5	55.69%	12.79%	27
0.1	10	54.77%	12.27%	27
0.1	15	50.46%	8.29%	27
0.1	20	53.00%	8.54%	27
0.25	0.1	39.23%	11.88%	3
0.25	0.25	37.62%	19.46%	3
0.25	0.5	43.46%	16.00%	3
0.25	1	35.23%	14.88%	3
0.25	2	47.77%	15.92%	3
0.25	5	47.23%	15.36%	3
0.25	10	37.46%	11.82%	3
0.25	15	42.08%	14.24%	3
0.25	20	33.85%	14.64%	3

*B. EFMM Results*

TABLE V. EFMM RESULTS

Theta	Gamma	Average of Accuracy	StdDev of Accuracy	Number of Hyperboxes
0.025	0.1	69.85%	2.68%	287
0.025	0.25	69.31%	3.39%	287
0.025	0.5	68.31%	2.37%	287
0.025	1	69.38%	2.26%	287
0.025	2	68.92%	3.67%	287
0.025	5	68.85%	2.60%	287
0.025	10	69.92%	2.39%	287
0.025	15	70.62%	2.68%	287

0.025	20	69.54%	3.58%	287
0.05	0.1	70.00%	3.34%	276
0.05	0.25	69.15%	2.03%	275
0.05	0.5	69.08%	2.85%	276
0.05	1	68.46%	3.77%	274
0.05	2	67.31%	2.74%	274
0.05	5	70.38%	3.88%	276
0.05	10	67.38%	2.67%	275
0.05	15	68.77%	2.46%	274
0.05	20	67.85%	2.53%	276
0.1	0.1	63.69%	2.87%	171
0.1	0.25	64.69%	3.91%	170
0.1	0.5	63.85%	4.20%	170
0.1	1	64.23%	3.12%	172
0.1	2	64.00%	3.64%	167
0.1	5	63.69%	4.71%	170
0.1	10	62.00%	2.74%	170
0.1	15	63.54%	2.33%	172
0.1	20	65.92%	4.73%	172
0.25	0.1	58.46%	5.14%	47
0.25	0.25	53.92%	7.63%	47
0.25	0.5	57.23%	4.63%	46
0.25	1	61.77%	6.65%	48
0.25	2	59.23%	7.12%	45
0.25	5	63.92%	6.79%	47
0.25	10	62.77%	5.65%	49
0.25	15	58.08%	6.73%	47
0.25	20	58.77%	5.89%	46

## VI. CONCLUSION AND FUTURE WORK

In this paper, we presented an assessment of the accuracy of FMM and EFMM for classification of musical pieces by the composer. Our work focused on Sonata pieces composed by Beethoven, Corelli, and Mozart. Our highest average accuracy measures were between 68% to 70% accuracy. Based on the results obtained, FMM and EFMM can be recommended for the classification of music by composer. The performance of the EFMM (in terms of predictive power) was consistently superior to the FMM.

Future studies should focus on the following:

- Analyze the optimal hyperbox expansion coefficient and sensitivity or measure of fuzziness with respect to accuracy, time efficiency, and space efficiency.
- Analyze the performance of the FMM and EFMM in classifying musical pieces containing multiple instruments.

- Research the selection of musical features for classification purposes.
- Research the classification of melodic sounds based on similarity to classical pieces.
- Asses the performance of FMM and EFMM in classifying music genre.

## REFERENCES

- [1] M. F. Mohammed and C. P. Lim, "An enhanced fuzzy min-max neural network for pattern classification," vol. 26, no. 3, March 2015.
- [2] P. K. Simpson, "Fuzzy min-max neural networks-Part 1: Classification," vol. 3, September 1992.
- [3] E. Pollastri and G. Simoncelli, "Classification of melodies by composer with hidden markov models," *Proceedings First International Conference on WEB Delivering of Music. WEDELMUSIC*, pp. 88-95, 2001.
- [4] M. A. Kaliakatos-Papakostas, M. G. Epitropakis and M. N. Vrahatis, "Musical composer identification through probabilistic and feedforward neural networks," *Di Chio C. et al. (eds) Applications of Evolutionary Computation. EvoApplications 2010. Lecture Notes in Computer Science*, vol. 6025, 2010.
- [5] D. Herremans, D. Martens and K. Sorensen, "Composer classification models for music-theory building," *Computational Music Analysis*, pp. 369-392, October 2015.
- [6] D. Herremans, K. Sørensen and D. Martens, "Classification and generation of composer-specific music using global feature models and variable neighborhood search," *Computer Music Journal*, vol. 39, no. 3, pp. 71-91, September 2015.
- [7] "KernScores," Center for Computer Assisted Research in the Humanities (CCARH), [Online]. Available: <http://kern.ccarh.org/>. [Accessed June 2017].
- [8] C. McKay and I. Fujinaga, "jSymbolic: A feature extractor for midi files," in *Proceedings of the International Computer Music Conference*, 2006.
- [9] Python Software Foundation, "Python," [Online]. Available: <https://www.python.org/>. [Accessed June 2017].
- [10] E. Jones, T. Oliphant, P. Peterson and et al, "SciPy: Open source scientific tools for Python," 2001-. [Online]. Available: <http://www.scipy.org/>. [Accessed June 2017].
- [11] S. van der Walt, S. C. Colbert and G. Varoquaux, "The NumPy Array: A structure for efficient numerical computation," vol. 13, no. 2, pp. 22-30, 7 March 2011.

# Assessment of Fuzzy Min-Max Neural Networks for Classification Tasks

Pasha Sadeghian, Aspen Olmsted

Department of Computer Science

College of Charleston, Charleston, SC USA

pashasadeghian@gmail.com, olmsteda@cofc.edu

**Abstract**—Statistical methods have been used in order to classify data from random samples. Generally, if we know the statistical distribution of the data, we can utilize specific classifiers designed for that distribution and anticipate good results. This work assesses the accuracy of Fuzzy Min-Max Neural Network (FMM) and Enhanced Fuzzy Min-Max Neural Network (EFMM) classifiers in classification tasks using data from five different statistical distributions: Negative Binomial, Logistic, Log-Normal, Gamma, and Weibull. Results of the assessment are provided and show different accuracies based on the statistical distribution of the data. This study presents a novel approach to the classification of statistical distributions by presenting two classifiers, namely FMM and EFMM Neural Networks, capable of classifying the above statistical distributions.

**Keywords**—Fuzzy Min-Max Neural Network, FMM, Enhanced Fuzzy Min-Max Neural Network, EFMM

## I. INTRODUCTION

Statistical methods have been commonly used in order to classify data from random samples [1]. Typically, if we know the statistical distribution of data, we can use particular classifiers designed for that distribution, and we can anticipate good results from that use [2].

Under several scenarios, it is not feasible to determine whether the sample data were measured with accuracy. Under these scenarios, the imprecision or fuzziness of data should be incorporated in the classification method [3]. Currently, a feasible approach for this modeling is using fuzzy sets proposed by Lofti A. Zadeh [4]. Many classification methods centered on fuzzy sets can be found in the literature and some of which are centered on measures of the probability of fuzzy events [5]. Among them, the Fuzzy Min-Max Neural Network (FMM) classifier was introduced by Simpson [6]. Several variations of the algorithm have been introduced and can be found in the literature, the most recent being the Enhanced Fuzzy Min-Max Neural Network (EFMM) classifier [7].

Ferreira et al. assessed the accuracy of a Fuzzy Gaussian Naive Bayes (FGNB) classifier for tasks using data from five different statistical distributions: Negative Binomial, Logistic, Log-Normal, Weibull and Gamma [3]. This paper aims to follow the work of Ferreira et al., and assess the accuracy of Fuzzy Min-Max Neural Network (FMM) and Enhanced Fuzzy Min-Max Neural Network (EFMM) classifiers for statistical distribution classification tasks using data from five different statistical

distributions: Negative Binomial, Logistic, Log-Normal, Gamma, and Weibull.

Section 2 presents the Fuzzy Min-Max Neural Network (FMM) and Enhanced Fuzzy Min-Max Neural Network (EFMM) classification methods which were used for data classification. In Section 3, the assessment methodology is described: the data used and how the samples were generated. In Section 4 details the classification results for the 5 statistical distributions.

## II. FUZZY MIN-MAX NEURAL NETWORK

The FMM learning algorithm is a three-step process, namely, hyperbox expansion, hyperbox overlap test, and hyperbox contraction. Learning in FMM beings by using a dataset comprising of input samples and target classes,  $A_h$ ,  $h = 1, \dots, N$ , where  $N$  is the total number of data samples. Based on the data samples, FMM generates a number of hyperboxes. Each hyperbox is represented by a minimum point ( $v_j$ ) and maximum point ( $w_j$ ) in an  $n$ -dimensional space within a unit hypercube ( $I^n$ ). Fig. 1 illustrates a 3-dimensional hyperbox with its minimum point ( $v_j$ ) and maximum point ( $w_j$ ).

Each hyperbox fuzzy set is defined as [7]:

$$B_j = \{A_h, v_j, w_j, f(A_h, v_j, w_j)\} \quad \forall A_h \in I^n \quad (1)$$

where  $B_j$  is the hyperbox fuzzy set,  $A_h = (a_{h1}, a_{h2}, \dots, a_{hn})$  is the input sample, and  $v_j = (v_{j1}, v_{j2}, \dots, v_{jn})$  is the minimum point for  $B_j$  and  $w_j = (w_{j1}, w_{j2}, \dots, w_{jn})$  is the maximum point for  $B_j$ .

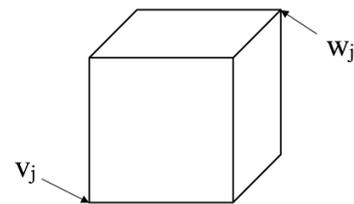


Figure 1. 3-D Hyperbox.

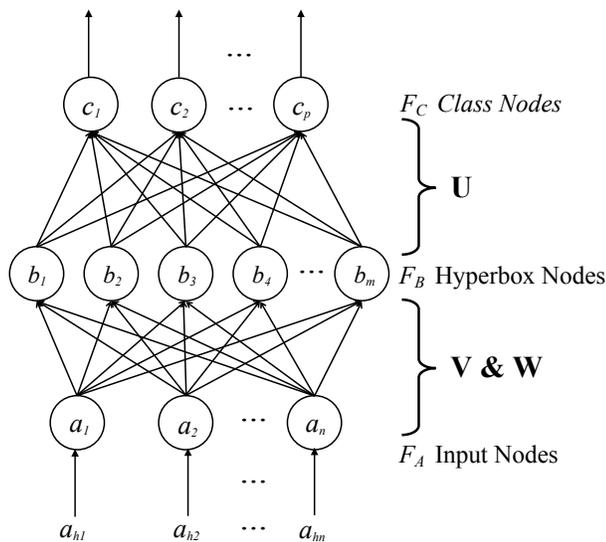


Figure 2. Three-Layer FMM Neural Network.

If an input sample is contained in a hyperbox, it is said that the input sample is a full class member of the respective hyperbox. The size of the hyperbox is controlled by a user-defined parameter referred to as the hyperbox expansion coefficient ( $\Theta$ ). The FMM neural network has three layers. Fig. 2 illustrates the three-layer structure of the FMM. The input layer  $F_A$  has  $n$  nodes equal to the number of input features. The hyperbox layer  $F_B$  consists of  $m$  sets equal to the number of hyperboxes created during the learning process. The connection between  $F_A$  and  $F_B$  are saved in two matrices ( $V$  and  $W$ ), with the membership function being  $F_B$ 's transfer function [6]. The output layer  $F_C$  has  $c$  number of nodes, each representing a class. The output of each  $F_C$  node represents the degree to which the  $h$ th input pattern  $A_h$  matches an output class. The connections between  $F_B$  and  $F_C$  are binary values and are stored in matrix  $U$ . The connections are defined by

$$u_{jk} = \begin{cases} 1, & \text{if } b_j \text{ is a hyperbox for class } C_k \\ 0, & \text{Otherwise} \end{cases}$$

where  $b_j$  is the  $j$ th  $F_B$  node and  $C_k$  is the  $k$ th  $F_C$  node [7]. For soft decisions, the outputs can be used directly [7]. For hard decisions, the  $F_C$  node with the highest value is chosen as the predicted class [7].

Typically, when a new training sample is presented, FMM utilizes the fuzzy membership function, to find the nearest hyperbox matching the sample. The fuzzy membership value of the training sample with respect to the  $b_j$  hyperbox can range from 0 to 1 (inclusive). The sensitivity of the fuzzy membership function is controlled by a user-defined parameter called sensitivity or measure of fuzziness ( $\gamma$ ).

During the hyperbox overlap test step, the FMM determines if any overlap exists between hyperboxes representing different classes. During the hyperbox contraction step, the existing hyperbox overlaps are eliminated by minimally adjusting each of the hyperboxes.

The hyperbox size is controlled differently under FMM and EFMM. However, for the purposes of this study, due to the one-

dimensional nature of the input pattern, the difference in hyperbox size control under FMM and EFMM can be ignored. FMM and EFMM also differ in their hyperbox overlap test and their treatment of the overlapping hyperboxes in the hyperbox contraction step. The details of the hyperbox overlap test step and hyperbox contraction step under FMM and EFMM can be found here [6] and [7]. That being said, the description of the FMM in this paper also applies to EFMM.

### III. RELATED WORK

Many studies have shown that assessment methods or classifiers yield better results when they are applied to data from a particular statistical distribution [3]. Generally, each classifier can yield better results when data follows some specific statistical distributions [3]. In a study in 2014, Ferreira et al. assessed the accuracy of a Fuzzy Gaussian Naive Bayes (FGNB) classifier for classification tasks using six different statistical distributions: Binomial, Continuous and Discrete Uniform, Exponential, Gaussian and Poisson [8] [9]. In a separate study in 2015, Ferreira et al. used Monte Carlo simulations to investigate the accuracy and behavior of an FGNB classifier for classification tasks using five different statistical distributions: Negative Binomial, Logistic, Log-Normal, Gamma and Weibull [3].

In their 2015 study, Ferreira et al. used Monte Carlo simulations and created four dimensions of observations for each statistical distribution [3]. Each dimension is comprising of 4 classes [3]. For each class, 40,000 observations were generated, resulting in 120,000 observations per dimension [3]. In other words, each assessment comprised of 120,000 observations. For each class, 10,000 observations were used for training, and the remaining 30,000 observations were used for testing [3]. The study showed that FGNB is capable of classifying the five statistical distributions [3].

### IV. ASSESSMENT METHODOLOGY

Our work followed the work of Ferreira et al. [3]. We used Monte Carlo simulations and created four dimensions of observations for each statistical distribution. Each dimension is comprising of 4 classes. For each class, 10,000 observations were generated, resulting in 40,000 observations per dimension. For each class, 2,500 observations were used for training, and the remaining 7,500 observations were used for testing.

We ran our initial tests using 400, 4,000, 10,000, and 40,000 observations per class. We noticed a difference in our results when using 400, 4,000, and 10,000 observations. However, we were not able to observe a noticeable difference between the results of 10,000 and 40,000 observations, hence why we chose to use 10,000 observations. By using 10,000 observations, we were able to save a considerable amount of time when running our tests.

#### A. Simulation

Random samples were generated for the 5 statistical distributions. The samples were generated using Numpy on Python [10] [11] [12], using the following parameters for each distribution:

- *Negative Binomial*: For the Binomial distribution denoted by  $X \sim BN(p, k)$ , 2 parameters are necessary for

samples generation. The following parameters were used:

TABLE I. PARAMETERS OF THE NEGATIVE BINOMIAL DISTRIBUTION

NEGATIVE BINOMIAL $X \sim BN(p, k)$	CLASS 1	CLASS 2	CLASS 3	CLASS 4
DIMENSION 1	(0.4,10)	(0.4,30)	(0.2,30)	(0.4,130)
DIMENSION 2	(0.6,10)	(0.3,30)	(0.4,20)	(0.5,140)
DIMENSION 3	(0.4,30)	(0.4,10)	(0.4,130)	(0.3,47)
DIMENSION 4	(0.3,10)	(0.4,80)	(0.5,70)	(0.4,130)

- *Logistic*: For the generation of Logistic distribution samples, the following parameters were used:

TABLE II. PARAMETERS OF THE LOGISTIC DISTRIBUTION

LOGISTIC $X \sim L(\mu, \sigma)$	CLASS 1	CLASS 2	CLASS 3	CLASS 4
DIMENSION 1	(0,2)	(20,2.5)	(43,2)	(60,3)
DIMENSION 2	(13,3)	(60,4)	(35,2)	(90,4)
DIMENSION 3	(20,3)	(40,2)	(108,4)	(72,4)
DIMENSION 4	(79,5)	(6,3)	(110,2)	(40,4)

- *Log-Normal*: Samples were generated from the Log-Normal distribution, using the following parameters:

TABLE III. PARAMETERS OF THE LOG-NORMAL DISTRIBUTION

LOG-NORMAL $(\mu, \sigma)$	CLASS 1	CLASS 2	CLASS 3	CLASS 4
DIMENSION 1	(2,0.2)	(3,0.3)	(3.8,0.3)	(4.5,0.2)
DIMENSION 2	(2.8,0.2)	(2,0.3)	(4.7,0.2)	(3.7,0.3)
DIMENSION 3	(2,0.3)	(2.65,0.2)	(3.7,0.3)	(4.5,0.2)
DIMENSION 4	(4.2,0.2)	(3.5,0.1)	(2,0.4)	(3,0.3)

- *Gamma*: The Gamma distribution has the following notation  $X \sim Gamma(shape, scale)$ . The following parameters were used for generation of the samples:

TABLE IV. PARAMETERS OF THE GAMMA DISTRIBUTION

GAMMA $(shape, scale)$	CLASS 1	CLASS 2	CLASS 3	CLASS 4
DIMENSION 1	(20,0.25)	(40,0.25)	(60,0.25)	(90,0.25)
DIMENSION 2	(12,1.0)	(32,1.0)	(65,1.0)	(110,1.0)
DIMENSION 3	(50,0.33)	(80,0.33)	(120,0.33)	(170,0.33)
DIMENSION 4	(80,0.17)	(130,0.17)	(190,0.17)	(250,0.17)

- *Weibull*: The Weibull distribution consists of two parameters called shape and scale. The following parameters were used for the generation of the samples:

TABLE V. PARAMETERS OF THE WEIBULL DISTRIBUTION

WEIBULL $(shape, scale)$	CLASS 1	CLASS 2	CLASS 3	CLASS 4
DIMENSION 1	(50,5)	(100,10)	(150,20)	(200,20)
DIMENSION 2	(50,5)	(100,10)	(150,20)	(200,20)
DIMENSION 3	(50,5)	(15,20)	(100,10)	(200,20)
DIMENSION 4	(200,20)	(50,5)	(150,20)	(100,10)

### B. Sensitivity or Measure of Fuzziness ( $\gamma$ ) and Hyperbox Expansion Coefficient ( $\theta$ )

In our preliminary tests, we used the following measures of fuzziness  $\gamma = \{0.05, 0.1, 0.2, 0.5, 1, 5, 10, 20\}$  and the following hyperbox expansion coefficients  $\theta = \{0.0001, 0.001, 0.005, 0.05, 0.1, 0.2, 0.5, 1, 5, 10, 20\}$ . With the focus on the accuracy of the FMM and EFMM classifiers in classifying the five statistical distributions, the following parameters were selected,  $\gamma = 1$  and  $\theta = 0.05$ . When using hyperbox expansion coefficients smaller than 0.05, we were not able to observe a substantial improvement in accuracy, however we were able to notice a substantial increase in training and testing time. Furthermore, hyperbox coefficients larger than 1 seemed to substantially increase the variance of our accuracy measurements, thus reducing the robustness of the accuracy of our classifiers. When using hyperbox expansion coefficients larger than 0.05, we were able to observe a substantial decrease in training and testing time, while the tradeoff was a reduction in accuracy. Keeping all other conditions constant, larger hyperbox expansion coefficient result in fewer number of hyperboxes, and fewer number of hyperboxes result in time saved in training and testing.

### C. Accuracy

For each dimension, ten sets of independent observations were generated, and ten independent tests were conducted. The observations for the FMM and EFMM were generated independently of one another. For each test, we measured the percentage of correct classifications and reported the average accuracies along with other useful statistics.

## V. RESULTS

### A. Negative Binomial Distribution

TABLE VI. FMM RESULTS FOR NEGATIVE BINOMIAL

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	75.20%	0.86%	379	30000	10000
2	78.78%	1.54%	271	30000	10000
3	85.07%	0.81%	348	30000	10000
4	79.70%	1.78%	364	30000	10000

TABLE VII. EFMM RESULTS FOR NEGATIVE BINOMIAL

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	75.22%	1.18%	377	30000	10000
2	78.32%	1.86%	273	30000	10000
3	84.92%	1.47%	351	30000	10000
4	79.60%	1.58%	364	30000	10000

### B. Logistic Distribution

TABLE VIII. FMM RESULTS FOR LOGISTIC

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	95.99%	0.30%	1841	30000	10000
2	96.26%	0.18%	2311	30000	10000
3	97.32%	0.18%	2313	30000	10000

4	97.60%	0.14%	2366	30000	10000
---	--------	-------	------	-------	-------

3	84.71%	1.21%	364	30000	10000
4	75.34%	0.27%	166	30000	10000

TABLE IX. EFMM RESULTS FOR LOGISTIC

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	95.73%	0.46%	1840	30000	10000
2	95.79%	0.11%	2304	30000	10000
3	97.19%	0.16%	2309	30000	10000
4	97.33%	0.10%	2364	30000	10000

C. Log-Normal Distribution

TABLE X. FMM RESULTS FOR LOG-NORMAL

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	85.52%	0.21%	2772	30000	10000
2	90.91%	0.26%	2726	30000	10000
3	86.46%	0.51%	2590	30000	10000
4	85.02%	0.33%	2291	30000	10000

TABLE XI. EFMM RESULTS FOR LOG-NORMAL

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	84.86%	0.18%	2766	30000	10000
2	90.69%	0.22%	2729	30000	10000
3	85.64%	0.32%	2575	30000	10000
4	84.82%	0.38%	2285	30000	10000

D. Gamma Distribution

TABLE XII. FMM RESULTS FOR GAMMA

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	86.18%	0.72%	1077	30000	10000
2	97.81%	0.12%	2446	30000	10000
3	94.51%	0.23%	1634	30000	10000
4	97.18%	0.20%	1233	30000	10000

TABLE XIII. EFMM RESULTS FOR GAMMA

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	85.21%	0.65%	1074	30000	10000
2	97.65%	0.14%	2464	30000	10000
3	94.12%	0.18%	1630	30000	10000
4	96.96%	0.11%	1233	30000	10000

E. Weibull Distribution

TABLE XIV. FMM RESULTS FOR WEIBULL

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	75.21%	0.37%	166	30000	10000
2	75.16%	0.41%	169	30000	10000
3	86.26%	0.88%	366	30000	10000
4	75.22%	0.40%	167	30000	10000

TABLE XV. EFMM RESULTS FOR WEIBULL

Dimension	Average of Accuracy	StdDev of Accuracy	Average of Number of Hyperboxes	Test Sample Size	Train Sample Size
1	75.36%	0.37%	164	30000	10000
2	75.42%	0.22%	166	30000	10000

VI. CONCLUSION AND FUTURE WORK

In this paper, we assessed of the accuracy of FMM and EFMM for classification tasks using data generated from different statistical distributions. We simulated five different statistical distributions: Negative Binomial, Logistic, Log-Normal, Weibull and Gamma. For each statistical distribution, four different dimensions were analyzed according to the proportion of correct classifications and the standard deviation of the proportions. Based on the results obtained, FMM and EFMM can be recommended for the classification of data from the Logistic distribution. For the Negative Binomial and Weibull distributions, the accuracies exceeded 75%. For the Gamma and Log-Normal distributions, accuracies exceeded 85%. For the Logistic distributions, accuracies exceeded 95%.

Future studies should analyze the performance of FMM and EFMM for very large sample sizes. Furthermore, future studies should analyze the optimal hyperbox expansion coefficient and sensitivity or measure of fuzziness with respect to accuracy, time efficiency, and space efficiency.

REFERENCES

- [1] A. R. Webb and K. D. Copsey, Statistical pattern recognition, 3rd ed., Wiley, 2011.
- [2] D. G. Strock, P. E. Hart and R. O. Duda, Pattern classification, 2nd ed., Wiley, 2000.
- [3] J. A. Ferreira, E. A. M. G. Soares, L. S. Machado and R. M. Moraes, "Assessment of fuzzy gaussian naive bayes for classification tasks," 2015.
- [4] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338-353, 1965.
- [5] L. A. Zadeh, "Probability measures of fuzzy events," *Journal of Mathematical Analysis and Applications*, vol. 23, no. 2, pp. 421-427, August 1968.
- [6] P. K. Simpson, "Fuzzy min-max neural networks—Part 1: Classification," vol. 3, September 1992.
- [7] M. F. Mohammed and C. P. Lim, "An enhanced fuzzy min-max neural network for pattern classification," vol. 26, no. 3, March 2015.
- [8] R. M. Moraes, "Performance analysis of evolving fuzzy neural networks for pattern recognition," *Mathware & Soft Computing*, no. 20, pp. 63-69, 2013.
- [9] J. A. Ferreira, E. A. M. G. Soares, L. S. Machado and R. M. Moraes, "Assessment of fuzzy gaussian naive bayes classifier using data with different statistical distributions," *III Congresso Brasileiro de Sistemas Fuzzy (CBSF2014)*, August 2014.
- [10] Python Software Foundation, "Python," [Online]. Available: <https://www.python.org/>. [Accessed June 2017].
- [11] E. Jones, T. Oliphant, P. Peterson and et al, "SciPy: Open source scientific tools for Python," 2001-. [Online]. Available: <http://www.scipy.org/>. [Accessed June 2017].
- [12] S. van der Walt, S. C. Colbert and G. Varoquaux, "The NumPy Array: A structure for efficient numerical computation," vol. 13, no. 2, pp. 22-30, 7 March 2011.

# Low-cost Detection of Backdoor Malware

Huicong Loi  
Tandon School of Engineering  
New York University  
Brooklyn, NY 11220  
hl2691@nyu.edu

Aspen Olmsted  
Department of Computer Science  
College of Charleston  
Charleston, SC 29401  
olmsteda@cofc.edu

**Abstract**—Backdoor malware are programs that enable hackers to access unauthorized computer systems by introducing a backdoor. These hackers will use this access to steal company information for personal gain. This malware uses a variety of techniques to hide their presence, and computer security researchers use a growing number of exotic techniques to detect them. However, it is not necessary to expend valuable IT resources on expensive security solutions as most of these backdoors can be detected by simple checks. We tested a wide array of in-the-wild malware to verify the effectiveness of these checks.

**Keywords**—component; backdoor; malware; Trojan; virus; detection; cheap; efficient; low-cost; simple; Windows

## I. INTRODUCTION

Cybercrime is becoming increasingly sophisticated, graduating from the “smash-and-grab” attacks of yesteryear. One of the most dangerous attacks detailed by Ping et al. involves hackers installing backdoor malware on computers in targeted companies [1]. These backdoors use a variety of techniques to avoid detection [2] and allow the hackers free access to the victim’s computers. Due to their stealthy nature, the backdoors allow the hackers to exfiltrate valuable business secrets while remaining undetected for a long time. To combat malware proliferation, computer security researchers have developed a diverse set of techniques to aid detection efforts [3].

In this paper, we develop an efficient method to detect stealthy malware, but without any reliance on exotic technologies or tools. This is accomplished primarily by targeting the command-and-control feature of backdoor malware. We then prove the technique by testing it out on a variety of malware samples captured from the wild.

The rest of the paper is organized as follows: Section II reviews related work. Section III describes the motivation for our work. Section IV describes the data collection mechanism and analysis. Finally, Section V will provide a conclusion and future work.

## II. RELATED WORKS

Idika et al. [3] performed a review of commonly-used heuristic methods used to detect malware activity. In their review, they noted that each heuristic had inadequacies, of which a high false-positive ratio and a large startup cost were the most common issues. They also discussed methods that could be used to defeat the various heuristics. Some heuristics that were

reviewed were: suspicious API call sequences, the frequency of common machine code sequences used in malware as well as program control-flow graphs.

Firdausi et al. [4] analyzed the effectiveness of using machine-learning to classify and detect malware activity. Their process consisted of sending both benign software as well as malware samples to a binary-analysis service which would derive various features and attributes. This information would be weighted and then used with various classifier techniques to predict if a particular program was a malware or not. The authors report a 95% success rate with the J48 classifier technique. However, it has a high running cost as each new item to be analyzed needs to be subjected to the binary-analysis service.

Christodorescu et al. [5] reviewed the effectiveness of some commercial antivirus products such as Sophos, Symantec, and McAfee. They found that commercial products were heavily dependent on signatures and simple obfuscation methods such as renaming variables and inserting garbage sequences would allow a previously flagged malware program to evade detection.

## III. MOTIVATING EXAMPLE

Cybercrime is a very real problem in today’s environment, but private-sector cybersecurity teams often operate with inadequate resources as security is often seen as a cost rather than a requirement [6]. Such teams might not be able to afford the newest cybersecurity solutions, which are becoming increasingly necessary due to the proliferation of malware obfuscation techniques [2] and the ineffectiveness of relying on antivirus software alone [5]. Having an efficient but low-cost method of detecting malware threats is of critical importance to these teams.

## IV. IMPLEMENTATION

We rely on a different approach from the related works, as the solution is intended for teams with limited resources. Due to this constraint, our solution is not intended to be able to detect all types of malware nor to achieve 100% intrusion prevention. The solution is developed for the Windows platform as this is the most common desktop environment.

For most data theft scenarios, a common requirement is that some malware that facilitates an internet connection must be present on the victim’s computer. This connection is used by the hackers to control these victim computers remotely and to run other hacking tools to perform other infiltration tasks [1]. The

malware that creates the internet connection is commonly called a backdoor. By targeting these backdoors, we sever the hackers' access to the victim computers and ensure safety. It is not essential to be able to detect the other types of hacking tools as the hackers would not have any means of controlling them without the backdoors.

The easiest way of finding out if a backdoor is running on your system is to see if some sort of internet connection has been established. All backdoors, except the most advanced kernel rootkits, reveal themselves in this manner. As even expensive cybersecurity solutions have difficulty dealing with kernel rootkits [7], so we do not see this as a limitation of our solution. However, merely creating an internet connection is not sufficient to flag a process as malicious. Internet browsers and various updater systems have perfectly legitimate reasons for creating connections. We need to derive additional characteristics that can differentiate between legitimate and malicious processes.

One question to ask is, "Why not filter by the internet connection's destination"? It is not practical to create a list of all malware sites in the world as there are simply too many and the exact number is very fluid. Some malware camouflages their traffic by using legitimate internet services such as Twitter or Google Drive.

To derive these distinguishing characteristics, we obtained malware from 'theZoo' malware archive and randomly selected 20 for testing. We developed a minimal list that satisfied detection for all of the 20 malware that was tested.

1. Process must make an internet connection
2. Svchost.exe makes an internet connection, but its parent is not services.exe.
3. The process makes internet connection but either does not have an existing parent process, or its ancestor is not explorer.exe.
4. The process makes internet connection but is a common Windows process that should not have an internet connection (explorer.exe, dllhost.exe, regsvr32.exe).
5. Subject all other processes to VirusTotal checks.

We used free software tools to implement our solution. TCPLogView was used to log internet connections as it had a very low CPU load and could be kept on without affecting user experience. A WMIC command was used to check the parent of flagged processes. Finally, VirusTotal has Python API examples which can be used to automate the task of submitting files for evaluation. One area of inefficiency could be the requirement of a large number of files needing to be submitted to VirusTotal for checks. One mitigation could be to develop a database of hashes of files that have been submitted to VirusTotal. Each file would only need to be submitted once, and subsequent checks could be referred to from the database.

We tested the implemented solution against 10 randomly selected malware from 'theZoo' malware archive.

TABLE I. TEST RESULTS

Malware	Detected	Rule
BlackEnergy2.1	no	-
Poweliks	yes	3
Rustock	yes	5
Trojan.Asprox	yes	2
Trojan.Bladabindi	yes	3
Trojan.Kovter	yes	3
Trojan.Loadmoney	yes	5
Win32.Sality	yes	4
Win32.Zurgop	yes	2
Win32.Lepthic	yes	4

We achieved a 90% success rate. The 'BlackEnergy2.1' malware circumvented our checks as it installed and ran as a legitimate Windows service.

## V. CONCLUSION AND FUTURE WORK

In this paper, we proved that our solution is effective against a variety of backdoor malware, even though we only used simple tools and techniques. However, there are still improvements to be made to the system.

- Dealing with malware that is installed as services
- Dealing with malware that injects themselves into the memory of internet-capable processes

## REFERENCES

- [1] P. Chen, L. Desmet and C. Huygens, "A Study on Advanced Persistent Threats," in *IFIP International Conference on Communications and Multimedia Security*, Portugal, 2014.
- [2] I. You and K. Yim, "Malware Obfuscation Techniques: A Brief Survey," in *2010 International Conference on Broadband, Wireless Computing, Communication and Applications (BWCCA)*, Fukuoka, Japan, 2010.
- [3] N. Idika and A. P. Mathur, "A survey of malware detection techniques," Purdue University, West Lafayette, Indiana, 2007.
- [4] I. Firdausi, C. Lim and A. Erwin, "Analysis of Machine learning Techniques Used in Behavior-Based Malware Detection," in *Second International Conference on Advances in Computing, Control, and Telecommunication Technologies*, Washington, DC, 2010.
- [5] M. Christodorescu and S. Jha, "Testing Malware Detectors," in *ACM SIGSOFT International Symposium on Software*, Boston, 2004.
- [6] A. Etzioni, "The Private Sector: A Reluctant Partner in Cybersecurity," in *Privacy in a Cyber Age*, vol. 28, New York, Palgrave Macmillan, 2015, pp. 58-62.
- [7] B. Schneier, "The Failure of Anti-Virus Companies to Catch Military Malware," [Online]. Available: [https://www.schneier.com/blog/archives/2012/06/the\\_failure\\_o\\_f\\_3.html](https://www.schneier.com/blog/archives/2012/06/the_failure_o_f_3.html). [Accessed 21 July 2017].

# Docker vs. KVM: Apache Spark application performance and ease of use

Walter Blair, Aspen Olmsted, Paul Anderson

Department of Computer Science

College of Charleston, Charleston, SC

wmblair@cofc.edu, olmsteda@cofc.edu, andersonpe2@cofc.edu

**Abstract**—*In this paper, we compare the performance and ease of use of an Apache Spark application deployed via Docker and alternatively via KVM. Docker containers have rapidly grown in popularity for dynamic cloud computing and for general-use application deployment and collaboration. Recent work comparing the performance of Docker versus VM deployments has found that Docker’s reduced overhead results in better performance. We compare Docker versus VM deployments of an Apache Spark application in regard to ease of use in a collaborative development environment. We expect that developing the application in a containerized environment will result in fewer post-deployment tests that fail compared to application versions that are developed in non-virtual or VM environments. Future work will closely examine the performance of these two implementations and more generally investigate potential trade-offs between performance and ease of use.*

**Keywords**—Spark, Docker, KVM

## I. INTRODUCTION

Computationally demanding processes and workflows, such as those involved in the fields of genomics and transcriptomics, rely decreasingly on dedicated HPC clusters and increasingly on dynamic cloud clusters. Driven largely by cost benefits, dynamic cloud infrastructures that spin up and pull down VMs or containers as needed promise to be the future of big data crunching in the life sciences [1]. A great deal of research is ongoing to compare cloud infrastructures based on VMs to those based on containers like Docker, and recent publications have used Docker to achieve dynamic cloud performance that is extremely fast and efficient [2, 3].

Enabled by hard and now dynamic clusters, the software tools involved in these computationally demanding workflows have been moving toward the parallelized MapReduce paradigm popularized by Apache Hadoop and more recently updated by Apache Spark [4, 5]. In particular, the ADAM software stack developed by the Big Data Genomics research group has built a processing framework on Apache Spark that has allowed developers to parallelize and standardize much of the computational pipeline and also streamline its development and deployment via Docker containers and Kubernetes container management with a system called Dockstore [6, 7].

Dynamic cloud infrastructure, parallel computing, and computational genomics are three cutting-edge areas of

research, and combining all of them is no easy task. A research lab seeking entry into the genomic and bioinformatic software development must share a Spark development framework and build, maintain, and deploy to a shared and reproducible cloud infrastructure. If a physical cluster is available to the development team, it may be difficult to determine best practices in setting up development and deployment environments that will afford the greatest achievable ease of use during collaborative development while also ensuring the best possible deployment performance. In the current research, we begin investigating such best practices by comparing developers’ ease of use between a Spark application deployed via KVM and via Docker for testing purposes.

The organization of this paper is as follows. Section II describes the related work on Docker/KVM performance for related applications. Section III states the guiding hypothesis for the current research. Sections IV documents the methods that have so far been implemented in pursuit of our hypothesis, and Section V presents our preliminary findings and current aims for the immediate future.

## II. RELATED WORK

Felter *et al.* (2014) compared KVM versus Docker deployments with a number of different applications and found that Docker outperformed KVM in most performance metrics [8]. The authors used KVM specifically but argued that it should be considered a fair representation of other hypervisors like Xen and VMware. Similarly, they used Docker to represent the expected performance of other container technologies. The authors hypothesized that containers would perform better because of their reduced OS overhead while acknowledging that containers have increased security concerns due to the minimal separation of layers in each container. Indeed, they found that Docker containers performed better overall and scaled very efficiently [8].

Paten *et al.* (2015) describe the progress and future aims of the NIH BD2K Center that has produced the ADAM software stack as well as standardized API’s that are all aimed at standardizing and thereby connecting the myriad genomics processing tools and workflows that will drive the future of the field [6]. Of particular interest is the group’s intention to deploy containerized tools and workflows that can be organized with the relatively new orchestration technology Kubernetes. This center has used Docker containers to streamline development as described in O’Connor *et al.* 2017 below, and they also view

Docker containers as important components of application deployment [6].

O'Connor *et al.* (2017) published the launch of the online community for genomics tools called Dockstore [7]. For the BD2K group, Docker has so far been employed in the service of collaboration between developers rather than dynamic cluster performance. Using BD2K's standard API's and schemas allows a developer access to the Quay web service that automatically containerizes a tool, registers it in Dockstore, and makes it available to be incorporated into a workflow. Manual Docker builds are also allowed, but it is clear that the intention of Dockstore is to standardize containerization to the degree that the hassle and waste of incompatible genomics software are minimized [7].

### III. HYPOTHESIS

The performance advantages of Docker deployments over KVM deployments are well documented and make sense given the stripped-down structure of shared-OS containers versus full-OS VMs. The security trade-offs between containers and VMs are also fairly well understood. Our research interest lies in the ease of use of, and particularly the ease of entry into, genomics software development. We hypothesize that establishing a shared Spark development environment within Docker will lead to a reduction in the number of tests that fail after initial deployment compared to deploying from a KVM development environment to a Docker container.

### IV. METHODS

In previous research, we developed a rudimentary sequence read trimming tool called Scalatrim in order to demonstrate that parallelizing the trimming process could improve batch performance in the transcriptome processing workflow [9]. Our initial Docker deployment of this application and subsequent deployment tests will be used as a baseline for the number of errors that might be expected when deploying locally developed software with Docker. With the Spark application of interest running locally, the next step in this research is to write a full development test suite that will be used to verify whether the application meets all requirements in the development environment. In the first version of the application, the development environment was non-VM. When the application passes all of the development tests, the next step will be to deploy the application to a Docker environment, retest, and record the number of tests that fail after deployment.

The same application will then be moved to a KVM development environment, and the software will be patched as necessary to meet the development tests that fail in this new environment. This KVM-tuned application will then be deployed with Docker, and the number of failed tests will be recorded. Finally, both versions of the application will be moved to a Dockerized development environment, updated as necessary, and then deployed with Docker. The number of failed tests will be compared across all application versions and interpreted as an indication of how much effort would be required to deploy an application to Docker from various development environments.

## V. EXPECTED RESULTS

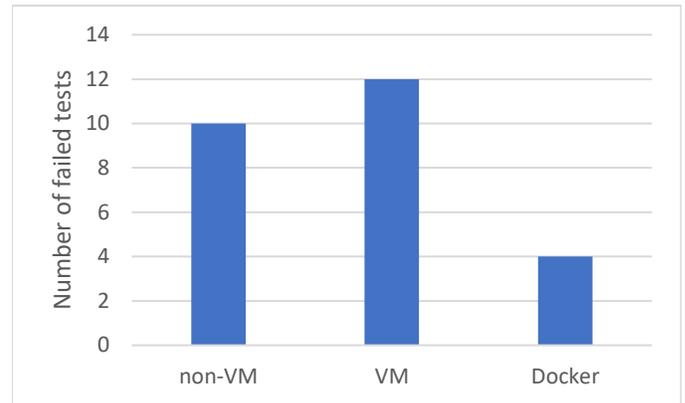


Figure 1 Expected number of failed tests post-deployment

Figure 1 illustrates the results we expect to find after testing all versions of the application after deployment. We expect that the versions developed in non-VM and VM environments will perform similarly to one another and worse than the version developed in a Docker environment. Our expected results support our hypothesis that developing Spark applications in a Docker environment will improve the ease of use of developers by reducing the development and debugging overhead of the deployment process.

We hope in subsequent research to join the ease of use results discussed in the present paper to performance comparisons between the various development-deployment strategies. While other researchers have clearly demonstrated that Docker deployments perform better than VM deployments, an open question is whether applications developed from scratch in a Docker environment perform better than applications developed in a non-containerized environment

## VI. REFERENCES

- [1] L. Stein, B. Knoppers, P. Campbell, G. Getz and J. Korbel, "Data analysis: Create a cloud commons," *Nature*, vol. 523, pp. 149-151, 2015.
- [2] D. Bernstein, "Containers and Cloud: From LXC to Docker to Kubernetes," *IEEE Cloud Computing*, vol. 1, no. 3, pp. 81-84, 2014.
- [3] C. Kan, "DoCloud: An elastic cloud platform for web applications based on Docker," *Advanced Communication Technology (ICACT)*, 2016.
- [4] M. Wiewiorka, A. Messina, A. Pacholewska, S. Maffioletti, P. Gawrysiak and M. Okoniewski, "SparkSeq: Fast, scalable and cloud-ready tool for the interactive genomic data analysis with nucleotide precision," *Bioinformatics*, vol. 30, no. 18, pp. 2652-2653, 2014.
- [5] S. Salloum, R. Dautov, X. Chen, P. Peng and J. Huang, "Big data analytics on Apache Spark," *Int. J. Data Sci. Anal.*, vol. 1, no. 3, pp. 145-164, 2016.
- [6] B. Paten, M. Diekhans, B. Druker, S. Friend, J. Guinney, N.

- Gassner, M. Guttman, W. Kent, P. Mantey, A. Margolin, M. Massie, A. Novak, F. Nothaft, L. Pachter, D. Patterson, M. Smuga-Otto, J. Stuart, L. Veer, B. Wold and D. Haussler, "The NIH BD2K center for big data in translational genomics," *J. Am. Med. Inform. Assoc.*, vol. 22, pp. 1143-1147, 2015.
- [7] B. O'Connor, D. Yuen, V. Chung, A. Duncan, X. Liu, J. Patricia, B. Paten, L. Stein and V. Ferretti, "The Dockstore: Enabling modular, community-focused sharing of Docker-based genomics tools and workflows," *F1000Research*, 2017.
- [8] W. Felter, A. Ferreira, R. Rajamony and J. Rubio, "An updated performance comparison of virtual machines and linux containers," in Performance Analysis of Systems and Software (ISPASS), 2015 IEEE International Symposium on, Philadelphia, PA, USA, 2014.
- [9] W. Blair, A. Olmsted and P. Anderson, "Spark framework for transcriptomic trimming algorithm reduces cost of reading multiple input files," *Submitted*, 2017.

## **Session 6: Internet Applications and Technology**

Title: Meaningful Sandbox Data  
(Authors: Ryan Lile, Aspen Olmsted)

Title: Classifying Influenza Outbreaks by Analyzing and Filtering Twitter Data  
(Authors: Elizabeth Healy, Husna Siddiqui, Aspen Olmsted)

Title: Handling an Organization's Communication Needs with a Single Web Service  
(Authors: Casey Wilson, Aspen Olmsted)

Title: Mobile Multi-Factor Authentication  
(Authors: Andrew Bissada, Aspen Olmsted)

Title: Three Factor Authentication  
(Authors: William Kennedy, Aspen Olmsted)

# Meaningful Sandbox Data

Ryan Lile, Aspen Olmsted  
Department of Computer Science  
College of Charleston  
Charleston, South Carolina  
lilerw@g.cofc.edu, olmsteda@cofc.edu

**Abstract—** In Salesforce sandboxes, there is a need for meaningful data to test new functionalities. In the production environment, there is already good data, but that data does not move over automatically to a new sandbox. Salesforce will move some data over for you if you are willing to pay. There are other very time-consuming alternatives. We propose a new method of data migration where running one program will move over all over any data you may need.

## I. INTRODUCTION

Salesforce has become an industry leading CRM in recent years. It provides the ability to develop full-stack applications on the cloud in Apex and Visualforce. It has great built-in security options and settings to allow both fine a course grained security by user, group, or license type.

One of many requirements for development on the Salesforce platform is the use of sandboxes and testing before the deployment of any Apex code to the production environment. No Apex can be implemented directly in a production environment. Instead, it is required to be developed in a sandbox related to that production environment. There is a required 75% code coverage for all apex. Once the code coverage is met, the code can be pushed to the production environment where it is tested then tested against the existing test classes in production and deployed.

A problem with this system is the availability of meaningful data. When a sandbox is created, no data is moved from production to the sandbox. There are options for Salesforce to move data for you, but it can be expensive and require additional licenses. You could go through the time of creating all new data, but this could be very time consuming depending on the type of data. Using the Salesforce data loading tool, you could pull download records from production then upload them to the sandbox. There is a need for a quick way to load meaningful data from production to a given sandbox.

The following section will describe related work and their outcomes from patents and journals about this topic. We will provide a motivating example for our project in section III. In section IV, we will describe the implementation of a system that streamlines the data migration process. In section V, we will discuss possible future work ideas.

## II. RELATED WORK

In other paradigm's there has been work on matching related records in new databases. David Baker has worked on the concept that the process that pulls and pushes the data assigns globally unique ID's to match records on in any setting, that way

no matter where the data is moved it will be able to match on those IDs.

## III. MOTIVATING EXAMPLE

In many situations, it is useful to have real world data instead of conveniently fabricated data to test new processes, systems, and applications. While implementing a ticketing system for the Gettysburg Foundation, development began in the sandbox. A few dummy records were created to test basic functionality, but to enable a user to test more complex processes, there was a need for real, or meaningful, data. There was a need to test tickets of many different times and many different activities. There was a need for a fast creation of meaningful data. Having the means to move over several real records from production would have made testing significantly easier.

Salesforce offers a sandbox that has a copy of some of the production data however most organizations have very few or one, and purchasing more can quickly become expensive. One option is to use the Salesforce data loader and download a lot of data then upload it to the sandbox but the matching of related objects can become very difficult, and this process requires a lot of download/upload time. Then there are local copies of the production copy left behind. We used the data loader to move the data but spent a lot of time downloading, matching, cleaning, and uploading the data.

## IV. IMPLEMENTATION

For this implementation, a few things need to be set up before the data migration. First, a connected app needs to be created in the production environment. Once this is created a consumer key, and consumer secret is provided for authentication. The user doing the data loading will also need to determine their security token for authentication. A restful API is created in the production environment. A get method needs to be created for each object type to be migrated.

In Salesforce, HTTP endpoints must be registered, as a security precaution, before a callout can be made. Therefore, two endpoints must be registered in the sandbox. The first is "https://login.salesforce.com/services/oauth2/token". This is the OAuth authentication endpoint where the session is can be obtained. There are three methods to do this: Web Server OAuth, User-Agent OAuth, and Username-Password OAuth. For this example, Username-Password is used because it is assumed that a user from the Production environment is migrating data to their own sandbox. Therefore their username and password for production is the simplest way to authenticate the session. The second endpoint to register is "https://\*yourinstance\*/services/apexrest". The URL is the endpoint where callouts to the production web services are made. The specific instance needed is the instance of the

environment the data will be migrated from, in this case, the production instance of the organization.

In the environment containing the data to be migrated, a restful web service must be set up. The methods can be as simple or complex as needed. For simple models, the method can just retrieve any number of records and return them. If more specific records are needed, the user could specify in the query a certain subset of the data or match based on a given field. Only fields contained in the query will be returned. This way you can decide what data you do or don't need in the sandbox. As an example, for a ticket object, there may be no need for the exact price if you are testing functions based on purchase time or ticket type. By limiting the fields to return you can save space in the limited sandbox.

In an API method, Salesforce automatically converts the data to be returned to JSON. However, if more complex logic is needed, the user could also construct a JSON string from scratch. A benefit to constructing a JSON string from scratch is simplifying matching of lookup relationships. Upon insertion into the Salesforce, all objects will be assigned a new ID. The data stored in a lookup field is just a string of the ID of the related object. The moved data will still have the old ID. When you construct your own JSON string, you could move record by record and insert the necessary lookup objects and use the new ID's to match.

On the sandbox side or destination environment for other use cases, an Apex class needs to be set up to accomplish three basic tasks. The first is authentication, the second is acquiring the data, and the third is inserting the data into the sandbox. This class can be a copy of a production class so that a new data insertion class doesn't have to be created for every new sandbox.

To get the necessary authentication, a callout to the token method must be made. This web service is a post method provided by Salesforce. You must supply the client id, client secret, username and password of the user, and the grant type which will be set to "password" for Username-Password OAuth. The response will provide two important values. The first is the instance URL. This is the specific instance of Salesforce organization that is the production version of your organization. This is the beginning of the URL to use for making callouts to production web services. The second is the access token. This is the session id that will allow you to access the production web services. Any call out to the web services of your organization must have a header authentication set to OAuth and the acquired session id.

Once the authentication is taken care of, the data needs to be gathered. Call the restful API created in the production environment to get the data. This step could be very simple, gathering only on the type of object, or very complicated, gathering records of many different types including many lookup or master-detail relationships. This is an area that can be conquered in many ways. For complicated models, matching can happen on either the API end or the local class end. On the API end, a custom JSON string can be created where related objects are values stored in the applicable records. Alternatively, on the sandbox end, multiple callouts can happen, and matching can happen between different responses.

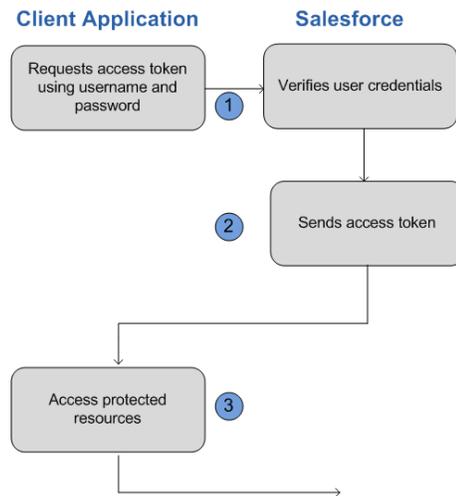


Figure 1. Username-Password OAuth Flow

Finally, once the data is gathered and matching has been accomplished, it is time for insertion. Each model will require different orders of insertion based on relationships. For each record, create a new object of that type, assigning values from the gathered data to the instance variables of the object. Then insert the object.

## V. CONCLUSIONS AND FUTURE WORK

As more and more cloud-based business move to the Salesforce platform, there will be a growing need to migrate data quickly and focus on building the new features of the system. This method of using a restful API to make callouts from the sandbox to get data from the production environment turns a long process to a quick program execution.

For future work, we propose work on efficient record matching for lookup and master-detail fields. This will make more complicated models easier to migrate.

## VI. REFERENCES

- [1] B. Korel, "Automated software test data generation", IEEE Transactions on Software Engineering, vol. 16, no. 8, pp. 870-879, 2002.
- [2] P. Weinmeister, Practical Salesforce.com development without code, 1st ed. [Berkeley, Calif.]: Apress, 2015, pp. 265-274.
- [3] D. Baker, M. Hipara and M. Gonzales, "Intra-platform data migration", 20170046373, 2017.

# Classifying Influenza Outbreaks by Analyzing and Filtering Twitter Data

Elizabeth Healy, Husna Siddiqui, Aspen Olmsted

Department of Computer Science

Graduate School of the University of Charleston

Charleston, South Carolina, USA

healye@g.cofc.edu, siddiquih@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—This paper uses twitter streaming and filtering techniques to determine which cities the flu is most prevalent in real time. The Twitter streaming API was used to collect data and filter using keywords and location. Our results show that more heavily populated cities have more cases of the flu.

**Keywords**—social media; data mining; classification; twitter; influenza;

## I. INTRODUCTION

The Center for Disease Control and Prevention (CDC) defines influenza as a contagious respiratory illness that can show symptoms of fever, chills, cough, sore throat, body aches, headaches, and fatigue [1]. The CDC relies on viral and clinical surveillance reports for influenza detection in the United States.

A report is sent to the CDC each week from public health, outpatient, and clinical laboratories with the number of positive influenza tests and patients with influenza-like illness (ILI). ILI is defined as a fever with a cough and/or sore throat without a known cause other than influenza [1]. Laboratory-confirmed influenza, hospitalizations, and deaths are also reported. These reports come from more than 350 laboratories, 2,800 outpatient health care providers, the National Center for Health Statistics, research and healthcare personnel at the FluSurv-NET sites, influenza surveillance coordinators, and state epidemiologists [1].

This data is collected from Sunday to Saturday of each week. The following week, the CDC receives a summarization on Tuesday and posts it to the public by Friday [1]. The full two weeks it takes to gather and report this information is a major flaw for influenza identification in a timely manner.

Twitter is a free social media platform that lets users share their ideas instantly as a 140-character long tweet. Currently, there are 313 million monthly active users. [2] These tweets can be examined and classified using Natural Language Processing. Twitter data has been proven useful in many different cases including political sentiment to predict election results [3] and text classification to improve natural disaster response [4].

Using data gathered from Twitter, it is possible to see the prevalence of influenza or ILI outbreaks in real time. Many tweets can be misleading because of influenza awareness versus influenza infection. Awareness can imply a user tweeting about the flu but not actually being infected. The filtering method

presented in this research effectively filters out tweets that are not reporting an influenza infection. Identifying influenza outbreaks by analyzing Twitter data can improve preparedness for a future epidemic.

## II. RELATED WORK

The two weeks it takes to gather and report this information is a major flaw. Using data gathered from Twitter, it is possible to see influenza-like illness in real time.

Ginsberg, Mohebbi, Patel, and Brammer used Google search queries to correlate words with the influenza virus. They displayed high accuracy with a mean correlation of 0.97 with the CDC-observed ILI percentages from a large dataset of more than 1 billion queries over a 5-year span [5]. Using data from a Google web search query is only available to Google as a service provider. This is a problem because the gathered data is not available to everyone. Twitter data is a widely available alternative used in this research.

Culotta used linear regression to detect key words in twitter data that estimate national influenza rates. His dataset only consisted of 500k tweets which led to a 78% correlation [6]. Later, Culotta improved on his previous work by using 570 million tweets which gave him a correlation of 95% but lacked filtering techniques [7]. His research didn't account for misspelled words or false tweets. An example of a false tweet would be influenza awareness by talking about the flu epidemic as opposed to having influenza-like symptoms.

Lamb, Paul, and Dredze used sophisticated filtering methods to differentiate between sickness and awareness. They used twitter data to determine influenza trends and demonstrate improvements on influenza surveillance [8]. They researched these trends and methods without comparing them to national statistics which can be beneficial for measuring general disease awareness but not for early detection and prevention.

Paul, Dredze, and Broniatowski improved on their previous work by comparing their data to national statistics but only focused on national prevalence as opposed to city and local levels [9]. City and local knowledge is advantageous for influenza prevention because it can better prepare these areas for an epidemic.

### III. MOTIVATING EXAMPLE

Influenza spreads around the world in seasonal epidemics. Over the years, these pandemics have killed tens of millions of people with each pandemic resulting in a new strain of the flu virus. The most recent of these pandemics was in 2009 when H1N1, commonly referred to as the Swine Flu, caused 503,563 confirmed cases and an estimated 151,700-575,400 deaths [10].

In 2009, Szomszor, Kostkova, and de Quincey did a study in London that used simple filtering and normalization techniques to analyze Twitter data. Their findings showed that the 2009 H1N1 outbreak could have been identified on twitter one week before it emerged from official records [11].

Disaster response and preparedness is a main factor in containing an outbreak. With twitter data and filtering techniques, epidemic preparedness can be drastically improved.

### IV. DATA COLLECTION AND RESULTS

We collected 250,224 Tweets over a period of 4 days from April 7, 2017, to April 11, 2017. The data was collected in 4 segments with each segment pertaining to new keywords. This was done to be able to filter and remove any false tweets.

#### A. General Health Filter

For the first filter, we used the general health terms { 'flu', 'sick', 'ill', 'unwell', 'miserable', 'illness', 'sickness', 'doctor', 'hospital', 'virus', 'disease', 'medicine', 'influenza'}. This was to start with as many tweets as possible that have the possibility of the user having the flu. This filter produced the 250,224 tweets which were then put through the next filter.

#### B. Flu Symptom Filter

The next filter used more specific terms pertaining to flu symptoms and antiviral medications including { 'flu,' 'cough,' 'sore throat,' 'fever,' 'chills,' 'headache,' 'congestion,' 'tamiflu,' 'rapivab,' 'relenza'}. This filter produced 10,933 tweets that showed more prevalence of users having the flu but still didn't filter out false tweets.

#### C. False Tweet Filter

The last keyword filter used keywords {'I have,' 'I have the flu,' 'I've', 'I'm'} to imply the user is talking about themselves and not someone else. This helped remove majority of false tweets such as a user talking about the flu instead of having the flu. This produced 1,626 tweets all pertaining to users who have the flu or influenza-like illness.

#### D. Location Filtering

Twitter users can enable their location which can be collected through the Twitter Streaming API. Since Twitter does not require the user to use their real location, many users put in a fake location that had to be filtered out. Of the 1,626 Tweets collected, 579 users had their location enabled with a real location. These tweets were used to determine where the flu is most prevalent at that given time.

TABLE 1. NUMBER OF FLU CASES BASED ON LOCATION

City	Number of Flu Cases vs Location
	Number of Users
Rio de Janeiro	78
Los Angeles	49
New York City	49
Houston	32
Abuja	27
Washington DC	27
Florida	24
Philadelphia	24
Toronto	24
Chicago	23

• Number of Flu cases in a specific location using the Twitter Streaming API

#### E. Results

As expected, densely populated cities had the most cases of the flu. The top 10 cities were recorded in Table I.

### V. CONCLUSIONS

In the future, we would like to use a larger dataset to show more differentiation between cities. This would help during an epidemic because if one city had a spiked increase of flu cases, then city officials would be able to prepare faster than if they had to wait 2 weeks for results from the CDC.

Other future work could use the same methods of filtering to determine statistics for disaster relief, politics, lice cases in schools, etc.

### REFERENCES

- [1] "The Center for Disease Control and Prevention," [Online]. Available: <http://www.cdc.gov>.
- [2] "Twitter," [Online]. Available: <http://www.twitter.com>.
- [3] A. Smeaton and A. Birmingham, "On using Twitter to monitor political sentiment and predict election results," *Sentiment Analysis where AI meets Psychology (SAAIP) Workshop at the International Joint Conference for Natural Language Processing (IJCNLP)*, 2011.
- [4] Z. Ashktorab, C. Brown, M. Nandi and A. Culotta, "Tweedr: Mining twitter to inform disaster response," *Proc. of ISCRAM*, 2014.
- [5] J. Ginsberg, M. Mohebbi, R. Patel and L. Brammer, "Detecting influenza epidemics using search engine query data," *Nature*, vol. 457, no. 19, 2009.
- [6] A. Culotta, "Detecting influenza outbreaks by analyzing Twitter messages," *arXiv:1007.4748 [cs.IR]*, 2010.
- [7] A. Culotta, "Towards detecting influenza epidemics by analyzing Twitter messages," *SOMA*, pp. 115-122, 2010.

- [8] A. Lamb, M. Paul and M. Dredze, "Separating fact from fear: tracking flu infections on Twitter," *HLT-NAACL*, pp. 789-795, 2013.
- [9] M. Paul, M. Dredze and D. Broniatowski, "Twitter improves influenza forecasting," *PLOS Current Outbreaks*, 2014.
- [10] A. Signorini, A. Segre and P. Polgreen, "The use of Twitter to track levels of disease activity and public concern in the U.S. during the Influenza A H1N1 pandemic," *PLOS One*, vol. 6, no. 5, 2011.
- [11] S. M., K. P. and d. Q. E., "#Swineflu: Twitter Predicts Swine Flu Outbreak in 2009," *Electronic Healthcare*, vol. 69, 2011.
- [12] E. Aramaki, S. Maskawa and M. Morita, "Twitter catches the flu: detecting influenza epidemics using Twitter," in *EMNLP '11 Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2011.
- [13] D. Broniatowski, M. Paul and M. Dredze, "National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic," *PLoS ONE*, vol. 8, no. 12, 2011.

# Handling an Organization's Communication Needs with a Single Web Service

Casey Wilson, Aspen Olmsted  
Department of Computer Science  
College of Charleston, Charleston, SC 29401  
cawilson1@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—Successful communication of information via web services can be a complex and error-prone task, in large part due to network complexity between multiple fine-grained web services. This complexity can decrease ease of maintainability and increase inconsistency between the source and target destinations. A single all-encompassing web service that meets all of an organization's needs can be implemented that performs both generic or highly specific tasks. We have created a REST web service in Salesforce that handles multiple objects and performs specific tasks to test the viability of a single web service to meet all of an organization's needs. We found that one web service can be implemented to handle multiple or specific tasks with a single call. We show that as the number of database actions per single request increases so does the efficiency with which each individual action is processed. A reduction in the number of web services coupled with an increase of the functionality of a single web service provides many benefits when compared to multiple smaller web services.

*Keywords*—web service; coarse-grained; ETS; database.

## I. INTRODUCTION

Often, websites that use large amounts of data store that data in either local or cloud databases. When websites or devices need to communicate with one another, they make use of web services to relay and make use of information. One set of commonly used web services is RESTful web services. REST services can be implemented to handle and manipulate a broad and diverse scope of data (coarse-grained) or a small and specific scope of data (fine-grained). Often websites with data stored in databases communicate with one another through REST services. For each set of connected REST services, both the implementation and their connection must be maintained. Since, by their nature, REST services are distributed, failure can happen in any individual part (either in a service itself or a connection between services), and larger networks of connected services become more difficult to troubleshoot when issues arise than do smaller networks of connected services. If websites had a single web service that could perform all necessary information transfers, the focus of information transfer could be moved away from creating and maintaining larger networks of multiple services and moved toward creating and maintaining larger single coarse-grained services that encompassed all of an organization's needs.

Salesforce, a popular platform as a service, is a common way for organizations to store, transmit, and receive data locally and externally. With its built-in ability to let users write code to create specific REST services and an out-of-the-box database, Salesforce is a good platform to test the feasibility of a single coarse-grained REST web service that provides the ability to perform robust operations on multiple database records and specific operations to alter small pieces of information.

## II. RELATED WORK

One important consideration for approaching this work is the concept of Extract, Transform, and Load (ETL), which in simple terms refers to the extraction of data from a source (Extract), the preparation of the extracted data to be loaded to the intended target (Transform), and loading the prepared data to its appropriate destination. Santhanakrishnan and Olmsted tested the performance of the ETL process on both coarse-grained and fine-grained REST web services and found that coarse-grained services processed faster per record, had more record changes per request, and consumed fewer APIs per request when compared to fine-grain services[1]. A way to ensure the correctness of continuously integrated databases was proposed in [2] called Continuous, Consistent, Extract, Translate, Load (CCETL). By using snapshot isolation of a database, CCETL provided higher guarantees of database isolation and higher throughput than traditional integration techniques. The authors of [3] proposed a system to make the ETL process more extensible and less rigid in the normally tightly couples components of the ETL process. In increasing the flexibility of the process, they were able to insert a new component in between the translate and load procedures that increased performance while solving a specific task, which was not possible through the standard ETL process.

## III. IMPLEMENTATION

To test a REST web service's ability to implement both specific and more general functionality, we created an Apex REST web service class in Salesforce that takes in an XML document, parses the document, and decides what to do with the information in the document. We implemented the class to handle six default Salesforce record types, but the class could easily be extended to implement a greater number of types as well as custom record types.

Depending on what information is contained in the XML document, the REST service class decides which specific record to update or insert, as well as which fields in the record to update. We sent our XML documents to the URL that called our web service using Postman. Each document sent contained either 1, 6, 12, 24, 48, 96, or 144 records to update. 144 was chosen because it is near the Salesforce upsert DML operation governor limit, which is 150. The documents sending a single record represented the most fine-grained request, and the documents sending 144 records represented the most coarse-grained request. Each specific number of records sent was sent five times to average the efficiency between them. The time to process each number is represented in Table I.

#### IV. RESULTS

We were successfully able to implement a single web service that could theoretically handle any number of individual records, were it not for the Salesforce governor limit of 150 upserts. When we averaged the amount of time per record to upsert, we found a general trend that the more records that were upserted in the same call, the less average time it took per record to insert. This is shown in Table I. To ensure that the time difference was a result of the upserts themselves and not a result of the surrounding code determining the actions, we repeated the calls to the REST service, this time removing the upserts from otherwise identical code, and we analyzed the time. We found that the code surrounding the upserts contributed to a fraction of a millisecond of additional time for each record. This suggests that the trend seen is related to how the upserts are handled in Salesforce, and not our specific implementation of the REST service.

TABLE I. TIME FOR RECORD INSERTION

Number of Records per Request	Average Time in ms to insert each record	Range of Times in ms to insert each record
1	58.00	41.00-86.00
6	33.69	32.5 - 35.67
12	25.94	22.25 - 30.00
24	24.63	21.21 - 27.38
48	21.06	19.54 - 24.10
96	18.83	17.57 - 19.42
144	19.15	18.19 - 20.01

#### V. CONCLUSION

When considering our implementation in the context of the ETL process, our implementation, which favors one all-encompassing web service over multiple smaller ones, simplifies the extraction of data because there is because the source only requires one working connection to its target

database, in this case, a correct call to the proper web service. All of the information relayed only needs to find one destination where it can be sent. This increase in the simplicity of data extraction requires an increase in complexity of the transformation process. If any information, regardless of specifics, is sent to the same destination, the destination must be able to handle all of the different possible types of data. For this to happen, the receiving web service must be robust enough to handle a variety of data, and the formatting of the data itself must be specific enough to address all of the particular nuances related to performing a given action.

While the increased complexity in parsing and decision-making of data is a tradeoff for the decreased complexity of the networks of services, the benefits of an all-encompassing service are numerous. Our REST service is able to handle both generic and specific requests, increases efficiency with larger amounts of data, consumes fewer APIs, and is more easily maintainable in terms of a web service network than a more fine-grained service, most tasks, regardless of specifics, can be implemented in a large web service that handles virtually all of an organization's needs more efficiently than multiple smaller services. This would decrease the potential number of errors of multiple individual web pages or web services being unable to communicate.

Furthermore, we have shown that updating data in databases can become more efficient as the size of the data increases. If this increase in performance of larger data sets is coupled with the decrease in complexity of the networks of services, communications between multiple points could become significantly faster and simpler overall, decreasing latency and increasing database correctness.

There are several improvements upon this work that can be investigated. For one example, a direct comparison of the number of inconsistencies in connected databases using either multiple services or a single service could be compared. Additionally, an important aspect that needs to be compared is overall ease of maintainability of both methods by determining the amount of time taken to create and alter the systems set up either to handle multiple web services or only one. Furthermore, a mathematical analysis could be performed to determine the relative change in complexity of many websites using multiple web services, vs. websites using only a single web service.

#### REFERENCES

- [1] G. Santhanakrishnan, A. Olmsted, "Fine-grain vs. coarse-grain web-service for ETL correctness and performance in cloud databases," in *2016 International Conference on Information Society*, Dublin, Ireland, 2016
- [2] A. Olmsted, "Heterogeneous System Integration Data Integration Guarantees," *JCMSE*. Vol. 17, no. S1, pp. S85-S94, 2017
- [3] M. M. I Awad, M. S. Abdullah, A. B. M. Ali, "Extending ETL framework using service-oriented architecture," *Procedia Computer Science*. Vol. 3 pp. 110-114, 2010

# Mobile Multi-Factor Authentication

Andrew Bissada, Aspen Olmsted  
Department of Computer Science  
College of Charleston  
Charleston, SC  
asbissad@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—Security has always been an important topic when it comes to the communication of sensitive data. With hardware advances allowing users the advantage of accessibility used in mobile devices, individuals are now spending more and more time on these devices. Additionally, with the viral popularity of social media applications and single sign-on, users do not always take as many precautions as needed with their information. Multi-factor authentication creates more and varied walls to block out the wrong people from seeing your information. Three-factor authentication can be applied to mobile devices as the first factor is already built in.

**Keywords:** *multi-factor authentication; biometrics; facial recognition; three-factor authentication; MIT App Inventor; Android;*

## I. INTRODUCTION

With the frequency of incidences of data breaches as well as identity theft, in combination with a larger population using technology and in particular mobile devices to save and communicate sensitive information, increased measures for security and improved authentication is more important than ever. There are many levels or factors to authentication. We are used to the first level of authentication, “something you know,” which typically is enforced by a username/password combination. For mobile-device users, the second level, “something you have,” is automatic because they have the device. This is why it is not a large jump to use three-factor authentication. The third factor is “something you are.” This is done with the use of biometrics. Examples of biometrics could be an iris scan, fingerprint scan, facial recognition, etc. From a user experience perspective, it is easier to ask a user to press a button and have their face automatically scanned, then to remember the second set of username/password combo and take the time to type that in as well. From a security perspective, not only is it another wall to discourage hackers, but it does it in an entirely different way. A figure for Multi-Factor Authentication can be seen on the following page. I believe that three-factor authentication can be built into mobile applications in a manner that is worthwhile. [1]

The rest of the paper is organized as follows: Section II reviews related work. Section III describes the motivation for my work. Section IV describes the implementation. Lastly, Section V will provide a conclusion and future work.

## II. RELATED WORK

Software development has used disk encryption for users and companies concerned about their data on their desktops and laptops. This can also be used to protect passwords on mobile devices as well. This is fine, but it doesn’t protect against phishing. Users transition between authenticating “online and on the device much more frequently” [2] and don’t know or care about the difference between the two. “Under these conditions, attackers can much more easily trick users into typing passwords into the wrong web page or application.” [2] This brings about the motivation for better authentication.

Another realm for mobile multi-factor authentication is in health systems, specifically the Telecare Medicine Information Systems. They allow patients to send and monitor their health information with their doctors remotely. Health and lives are at risk here, so security is of utmost importance. Login and authentication start out by having the smart card check if the patient is who they say they are by verifying the ID, PW, and Biometric that the user enters. Then the server and user mutually authenticate each other. Additionally, they get a shared session key that will be used to encrypt/decrypt and authenticate future communications. After this, the patient can log in to the telecare server. This is a secure solution but a little large for the scope of my study. [3]

As individuals are trending towards using their mobile devices at an ever increasing rate, and some of the user cases include storing or exchanging sensitive information, mobile devices need a way to protect users and their data using secure protocols for access and identification. Face and iris recognition are better than fingerprints for this regard. The reason being is, they are reliable and only need the webcam already on the device. Fingerprints, on the other hand, need a dedicated sensor. One proposed way is FIRME (Face and Iris Recognition for Mobile Engagement) which is based on a modular architecture. The architecture “includes separate and replaceable packages.” [4] It first acquires the image. Then different branches “perform detection, segmentation, feature extraction and matching for face and iris separately.” [4] In regards to face, there is a step performed after segmentation for anti-spoofing. All algorithms are optimized not to be demanding and light on computation as there are limited resources on mobile devices.

### III. MOTIVATING EXAMPLE

I wanted to see if it was feasible for three-factor authentication to become the norm for certain categories of mobile applications such as health and banking. I do not think that SMS should play a primary role in authentication. [5] The best way to do this combines both security and ease of use. If something is very secure but very hard to log into, no one will want to use it. This is why people re-use passwords for multiple sites and use simpler logins on mobile devices. On the other hand, if it's very easy to use but not secure, then it is inherently at risk. The best way to fulfill both of these purposes is to make the third factor of authentication, within biometrics, use facial recognition. After entering their username/password, all the user has to do is press one button.

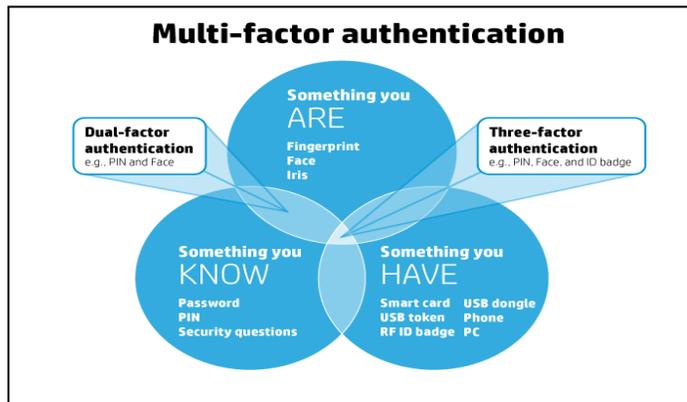


Figure 1. Multi-Factor Authentication

### IV. IMPLEMENTATION

I used the MIT App Inventor tool by Google to create an Android app that authenticates a user. I had to add on the Makeblock computer vision extension which implemented Microsoft Cognitive Services. This allowed me to add in facial recognition. I also used TinyWebDB to store the usernames and passwords. Overall it is a three-factor authentication because you have to have something (the phone), you have to know something (the username/password), and something about you (your face). My code blocks can be seen in the following figures:

Figure 2. App Inventor: Screen 1 Code Blocks

Figure 3. App Inventor: Screen 2 Code Blocks

### V. FUTURE WORK

I still believe that three-factor authentication is beneficial to certain mobile applications. Given more time, I would change a few items in this project. I would first get off of the public TinyWebDB and get my own custom database. I would then change the username/password authentication to be more secure. Being in plaintext is too easy to hack. I would probably use a salted hash. I could even use sign-in from Google/Facebook. Secondly, I would use a different service for facial recognition. Instead of Microsoft Cognitive Services API's, I would look into using OpenCV. It was used in Marsico's article about FIRME, and I have viewed a number of demo videos of it, and it appears to fit better the user experience I have envisioned.

### VI. BIBLIOGRAPHY

- [1] M. Rouse, "Multifactor Authentication," Tech Target, March 2015. [Online]. Available: <http://searchsecurity.techtarget.com/definition/multifactor-authentication-MFA>. [Accessed June 2017].
- [2] D. DeFigueiredo, "The Case For Mobile Two-Factor Authentication," *IEEE*, p. 5, 2011.
- [3] M. Nikooghadam and H. Arshad, "Three-Factor Anonymous Authentication and Key Agreement Scheme for Telecare Medicine Information Systems," *Journal of Medical Systems*, p. 11, 2014.
- [4] M. D. Marsico, C. Galdi, M. Nappi and D. Riccio, "FIRME: Face and Iris Recognition for Mobile Engagement," *Image and Vision Computing*, p. 11, 2014.
- [5] A. Greenberg, "So Hey You Should Stop Using Texts For Two-Factor Authentication," *Wired*, June 2016. [Online]. Available: <https://www.wired.com/2016/06/hey-stop-using-texts-two-factor-authentication/>. [Accessed June 2017].

# Three Factor Authentication

William Kennedy, Aspen Olmsted  
Department of Computer Science  
College of Charleston  
Charleston, SC  
wakenned@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—Authentication, security, and confidentiality are some of the most important topics of cyber security. There have been many solutions presented to users for strengthening the security of login password-based authentication methods. Primarily this has been through the use of two-factor authentication methods. Two-factor authentication is the combination of single factor authentication mechanisms. The growing popularity and acceptance of two-factor methods are driven by the increasing need for privacy and security in this technological age. The success and popularity of adapted security measures are largely dependent on their ease of implementation and convenience to the user. The focus of this research is to address and analyze the implications of using a three-factor authentication model for added security in websites and mobile apps. This paper will present an app we created which could provide a potential method for three-factor authentication that could potentially ensure added authentication assurances without loss of convenience.

**Keywords**- Authentication; Security; Three Factor Authentication; Mobile; Web; Access Control

## I. INTRODUCTION

User authentication is the main building block for any secure cooperative computing system. Security concerns are on the rise in all areas of industry such as banks, healthcare institutions, industry, etc. Due to the proliferation of mobile devices and the heightened interaction between mobile applications and web services, the authentication of users is more frequent for mobile devices than for desktop users [1]. In many instances of multi-factor authentication, both a mobile device and a desktop are necessary and go hand in hand for adequate authentication. One of the drawbacks of multifactor authentication is that user ID's and passwords are abundant, with many users stating that they have more user IDs and passwords than they can remember [2]. This cost of convenience makes the proposed implementation of higher security measures and added authentication factors worrisome to many users and providers.

To better understand the factors in play with authentication, it is first necessary to understand what authentication is. Authentication and the various measures of authentication are used to verify that a specific user or process is who they say they are. It is that simple. There are four standard ways that users are authenticated:

- Something you know – This is the most basic form of authentication with which most users are familiar. This standard is usually presented as a username or password which is known only to the user.

- Something you have – This form of authentication is represented by the user having possession of a physical entity or device. This can be represented as a physical token such as the user's smartphone or other media device generating a temporary and sometimes single use authentication code.
- Something you are – This form of authentication is represented as a biometric signature such as a fingerprint, retina scan, or facial recognition. This is generally seen as one of the strongest forms of authentication when conducted properly [3].
- Someplace you are – This form of authentication corresponds to where a user or process is located, and in response gives or denies access to resources accordingly. This standard can be conducted through the use of a range of IP addresses or geographic location points [3].

For multifactor authentication to be conducted properly, one or multiple of these standards are coupled to heighten security standards. In this work, we will present an app we created which joins three of these standards in order to give a higher level of security and authorization to the user. The organization of this paper is as follows: Section II reviews the related work. Section III describes the motivation for this research. Section IV describes the implementation and results of the created app. Conclusively, Section V will provide a conclusion and discussion of future work.

## II. RELATED WORKS

Li et al. [4] first introduced the concept of using biometrics as a factor in a three-factor authentication system in order to help the grim situation of network security. In this paper, they proposed using passwords, smart cards, and biometrics as a viable solution towards three-factor authentication. It was through this proposal and the analysis of the progression of modern technology and the in-depth look into the drawbacks of fewer security measures which prompted the use of facial recognition as part of the app to be developed for this project. The discussion of the speed/flow of biometrics, as opposed to other authentication factors, was very influential to the progression of this research. The presented scheme could resist many kinds of attacks and protect the hosts multimedia and web resources, which in response, made it very attractive as a potential starting off point.

In Native Autonomous Process Authentication, Olmsted analyzed the promise of multifactor authentication when it came to users while also analyzing the issues that arise when applications try to give credentials to autonomous software processes. Autonomous software processes when allowed credential privileges leave gaps for malicious incursions and attacks [3]. The importance of this paper was paramount as it proposed that in order to have a more secure system, autonomous background processes should not be allowed credentials which in hindsight they should not possess in the first place, as issues can arise.

In Two-Factor Authentication: Cybersecurity for Today's World, HCPro discusses the importance of security in the healthcare industry with "two-factor authentication being the backbone of even a basic cybersecurity program" [2]. The emphasis on weighing convenience and speed over security was part of the main discussion and analysis of this paper. The importance of this paper towards my continued research and app development centered on making the process as seamless as possible with having as few clicks as possible connected to multiple processes.

### III. MOTIVATION

The motivation of this paper is to investigate an orderly approach to the design of a secure three-factor authentication app with the protection of user privacy in the most time efficient and convenient way allowable. The hypothesis is that an app can be developed that performs three-factor authorization without being overly complicated or time-consuming. The app will use three-factor authentication to incorporate the increased advantages of an authentication based on a password, username, and facial recognition through possession of a mobile device [5].

### IV. IMPLEMENTATION AND RESULTS

The design of the app took place in MIT App Inventor 2 beta. The app was designed to consist of two pages in this early stage and to work primarily on Android. The login page functions to take in a username and password. With this app being in the prototype phase the login credentials were hard coded. This means that the user would not be able to create their own account at this time; but would at a later date be able to, when the app can be fleshed out. The user would log in to this page with the username and password, and then the camera would queue up and take their picture, performing a facial recognition check.



Figure 1. Login Screen

The facial recognition functionality was provided through an extension called CamVision1 which works with MIT App Inventor 2. The next step was incorporating the Microsoft Cognitive Service, which allowed for the addition of a computer

vision API which would be necessary for the facial recognition process. Once the facial recognition is completed, the app takes you to the second screen which presents a successful login message.

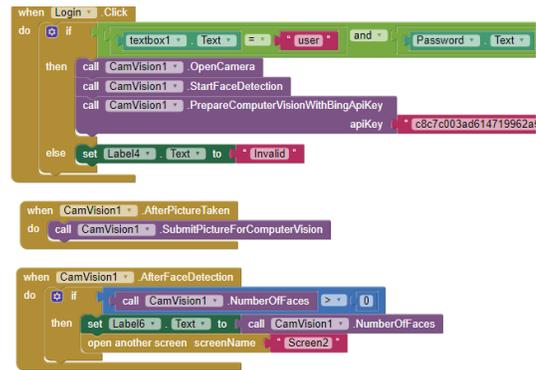


Figure 2. MIT App Inventor 2 Blocks

The app performed as desired with user interaction consisting of the input of a username and password and then having their picture taken all at the expense of a single click of the mouse. The convenience and simplicity of this three-factor authentication app exceeded all expectations for ease of use. The cost of convenience is at its most minimum as presented through the apps one click sign in capabilities.

### V. CONCLUSION AND FUTURE WORK

The implementation of a future market standard of a three-factor authentication method seems all but assured with the use of biometrics and other authentication methods when used in an efficient way. With efficiency comes confidence and with confidence comes reliability. The increased reliability of a more secure platform with three-factor authentication is hard to ignore. With further research, the app can be developed to allow for users to create their own accounts, and for the account development of each user to save credentials and biometric reference tags.

### REFERENCES

- [1] D. DeFigueiredo, "The Case for Mobile Two-Factor Authentication," *IEEE Security & Privacy*, vol. 9, no. 5, pp. 81-85, Sept-Oct 2011.
- [2] HCPro, *Briefings on HIPAA*, vol. 17, no. 1, pp. 1-5, January 2017.
- [3] J. N. M. K. J. L. X. Z. Xiong Li, "Rhubust three-factor remote useer authentication scheme with key agreement for multimedia systems," *Security and Communication Networks*, 2016.
- [4] Y. X. A. C. J. Z. R. D. Xinyi Huang, "A Generic Framework for Three-Factor Authentication: Preserving Security and Privacy in Distributed Systems," *IEEE*, vol. 22, no. 8, pp. 1390-1397, 2010.
- [5] A. Olmsted, "Native Autonomous Process Authentication," in *Proceedings of World Congress on Internet Security 2016 (World-CIS 2016)*, London, UK, 2016.

## **Session 7: Infonomics and e-Technology**

Title: A Secure Enterprise Architecture Focused on Security and Technology-transformation (SEAST)  
(Authors: Md. Tomig Uddin Ahmed, Nazrul Islam Bhuiya, Md. Mahbubur Rahman)

Title: Distributed Query Processing and Data Sharing  
(Authors: Ahana Roy, Aspen Olmsted)

Title: Factors affecting the implementation of information assurance for eGovernment in Indonesia  
(Authors: Rio Guntur Utomo, Robert J. Walters, Gary B. Wills)

Title: The Study of Public Organization's Intention to Use an Open Government Data Assessment  
Application: Testing with an applicable TAM  
(Authors: Chatipot Srimuang, Nagul Cooharajanone, Uthai Tanlamai, Achara Chandrachai,  
Kanokwan Atchariyachanvanich)

# *A Secure Enterprise Architecture Focused on Security and Technology-transformation (SEAST)*

Md. Tomig Uddin Ahmed  
Planning Division,  
Ministry of Planning Dhaka,  
Bangladesh, Email:  
ssa@plandiv.gov.bd

Nazrul Islam Bhuiya  
Bangladesh University  
Of Professional  
Email: nibs1\_nazrul@yahoo.com

Md. Mahbubur Rahman  
Department of Electrical & Computer  
Engineering, MIST, Bangladesh.  
Email:mahbubcse@yahoo.com

**Abstract— Sustainable and flexible development of an organizational business and achievement of current and future goals of an organization depends on business processes, technologies and information systems. In order to address these issues, there have been a number of enterprise architectures proposed in the literature. Unfortunately, many of the existing architectures pay little attention to security and sustainability. To create an accelerated and flexible IT infrastructure, it is needed to align IT and organizational business demands. Here, we proposed an enterprise architecture which is sustainable, secure and convivial to new technology for a developing or developed organizations. It is a robust and easily understandable architecture which includes single source identification system and a knowledge base with sustainable technology-transformation capabilities. This architecture is based on concepts, modeling and processes which can be used as the framework for both business and IT professionals. The proposed architecture includes kernel based security architecture with risk, incident and audit management systems. It is a high performing architecture which includes all the essential features of modern enterprise architecture (EA) enhanced with security and sustainability.**

**Index Terms—Main Architectural Model(MAM); Technology-Transformations; Single Source Identification Model(SSIM); Knowledge Base (KB); Security Architecture(SA); Enterprise Architecture(EA).**

## I. INTRODUCTION

A business process of an organization is regulated by a set of guidelines, policies, principles, technological dimensions, standardization and integration of its operating model to the organization's enterprise architecture (EA). In order to offer better services to business and citizens enterprises are enhancing themselves through aligning IT systems with business process [1-2]. A well-designed IT infrastructure provides a better link among citizens and government [3]. The aim of enterprise architecture is to facilitate an organization to achieve its goals through utilizing available technology in terms of public service delivery [4-5]. As today's living is vastly influenced by information technology, so as it has bigger impact on enterprise architecture (EA) for better

organizational management. Enterprise architecture formulates the proper IT infrastructure depending on the business demands of the enterprise [6]. The vital role of an EA is to identify a framework and to propose the proper methods to implement it. Although, there exists some EA frameworks that depicts implementation mechanism for EA, they are mostly enterprise specific[7]. Enterprise architecture framework (EAF) delineates the framework for modeling enterprise's information systems and its business. Depending on the activities and scope with different perspectives, an EAF may consist of different models. It is a fact that just to model the information system and the business is not sufficient enough for an organization, which has other goals such as to achieve competitive edge and to have strong base to face current and future challenges. Therefore organizations adopt enterprise architecture framework to get their business demands and information system (IS) to be aligned [8-9].

For both government and nongovernment organizations, the scope of business area is increasing tremendously. Demand for better service quality is also increasing dramatically. For the fast delivery of the services, there is a need for business and technology collaboration. While upgrading the technology and infrastructure of an enterprise, it is essential that enterprise is technology-transformation friendly and it is secure enough which is essential for the protection of IS and business. So a model which focuses on technology-transformation and security is very important. In this paper we propose secure enterprise architecture focused on security and technology-transformations (SEAST) which is a comprehensive description of all of the key elements and relationships that make up a secure organizational framework. Security is a critical business concern due to rapidly growing vulnerabilities in the systems. These researches will pave the way to build enterprise architecture models as a framework to design secure architectures.

This proposed architecture is based on single source identification process for monitoring and controlling of the enterprise. For the research and data analysis, knowledge base is collaborated in the Main Architecture Model (MAM). MAM is the heart of the SEAST of this research. This paper describes a secure technology- transformation friendly model which is used to gap analysis and technique to receiving the target. The aim of this architecture is to the organized all logical and physical aspect of EA. For this, we shall discuss the literature review in section II, proposed methodologies in section III,

comparison and discussion in section IV and conclusion in section V of this paper.

## II. LITERATURE REVIEW

The definition of EA is not unique as different literatures define EA differently [9]. According to the systems and software consortium it is the alignment of business and IT systems. Some views it as the combination of models and definitions [10]. Some other view it the process of simplifying complex management problems through IT. Others view it as the combination of hardware, software and network technologies through which an IT project is implemented. Gartner definition describes EA as a process for integration and change of enterprise. Some researchers define from completely different perspective. They view EA as a knowledge base which provides required blueprints [11]. To the decision makers of IT and business, it is the alignment of enterprise architecture and process to the business goals [12]. The major benefit of EA perceived by many is the lowering of the complexity of the enterprise [13-14]. Other researchers elaborated more on the benefits of the enterprise architecture.

A recently proposed EA is a TOGAF framework [17]. Implementing EA with TOGAF ADM provides reflecting the needs of stakeholders, the best method for employees, consideration to the current needs, or an overview for the future needs of an organization. TOGAF ADM provides support in the area of process and data integration and infrastructure and application integration [15]. The framework encompasses stakeholder viewpoints, business process, data and IT systems. At the strategy level, TOGAF ADM helps to balance tradeoffs between benefits and costs by providing methods to measure the value of work packages [16]. In this paper, after reviewing literature we proposed secure enterprise architecture based on security and technology-transformations (SEAST).

## III. PROPOSED ARCHITECTURE

In this proposed architecture, SEAST have been developed by including some components from existing model [17] and we added new component which are deemed necessary for sustainable businesses.

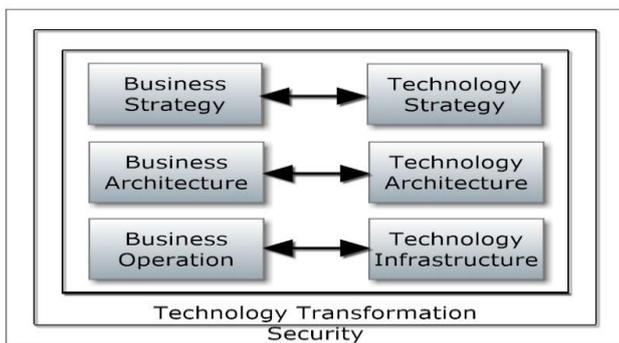


Figure 1. Basic of MAM

Our proposed architecture is scalable to accommodate current and future technologies for both government and private organizations which emphasizes the security and technology transformation.

In Fig. 1 a skeleton of SEAST is depicted which is inter linked with strategy, architecture and operation of business and technology.

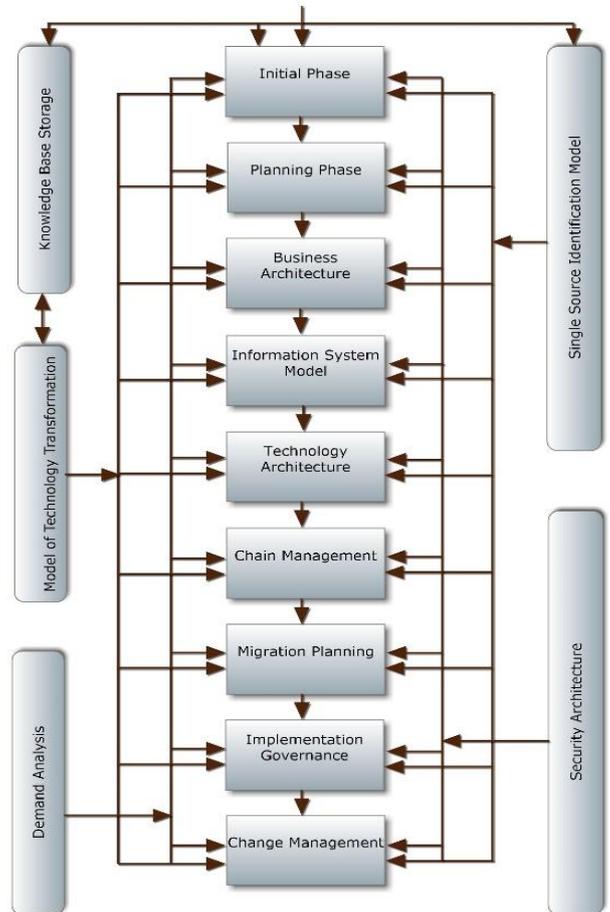


Figure 2. Main Architectural Model (MAM)

In Fig. 2 a main architectural model (MAM) is presented which contains sequential nine phases and all the phases are connected to demand analysis, technology transformation's model, single source identification model, security architecture and knowledge base. MAM is the recipe of the SEAST. It provides interoperability for integration of all aspects and information which will go into one process and the output will become the input of other process. In following paragraphs, we describe the phases, architectures and models of MAM in brief.

### A. Initial Phase

The initial phase of SEAST which is architectural project. It undertakes initial activities and preparation to build an architectural capability. Development tools, definition of architectural principles are defined in the phase.

**B. Demand Analysis**

Demand analysis for an organization is connected to all the phases to determine what type of objectives are needed for achieving the business goal. Every phases of SEAST projects are based on validated business requirements. Need identifications are to be confirmed into the phases with the benchmark of the enterprises.

**C. Planning Phase**

Planning is one of the core stage of MAM which involves to creating of a set of plans to help guide the execution and closure stage of the SEAST project. This phase helps to manage cost, time, quality and success. All process and activities requirement are to be defined into the planning phases.

**D. Business Architecture**

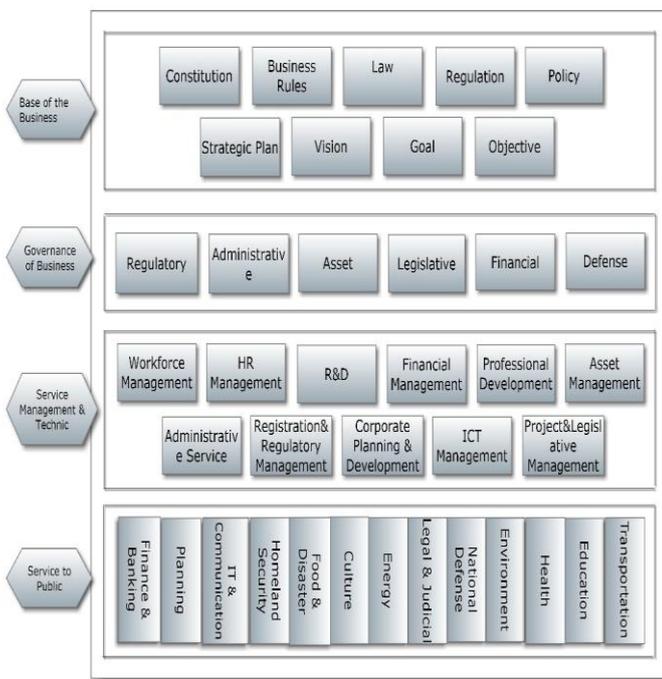


Figure 3. Business Architecture

In Fig. 3 service strategy is described through the business architecture. It also reveals the overall techniques business environment along with the associated information function and processes. There four stages in this sustainable business architecture are the business base, governance of the business, service management technique and service to public. They are strongly related to each other for easy and accelerated services delivery to public. It is important that to achieve business goals business strategies and business drivers are to be chosen accordingly. Also, the architecture to be developed is to be aligned with specific business requirements. The business goals and the drivers are actually defined by business

strategies which are reflected through the business architecture (BA).

**E. Information System Model**

Fig. 4 describes an information system model (ISM) which is an integral part of an organization which function as a basis for IT support. Organizational Information flow is made through the ISM. Naturally the information flow includes collect, filter, process, create, and distribute phases. structure and behavior is.

In this proposed model, application management, middleware, data management, network system, distributed environment and risk management stages are serially connected to make collections, organizations, storages, communication of information and business continuations of an enterprise. Technology-transformation and security stages are connected to all the stages which supports adoption of new technologies, migrations and manages or maintenance of security.

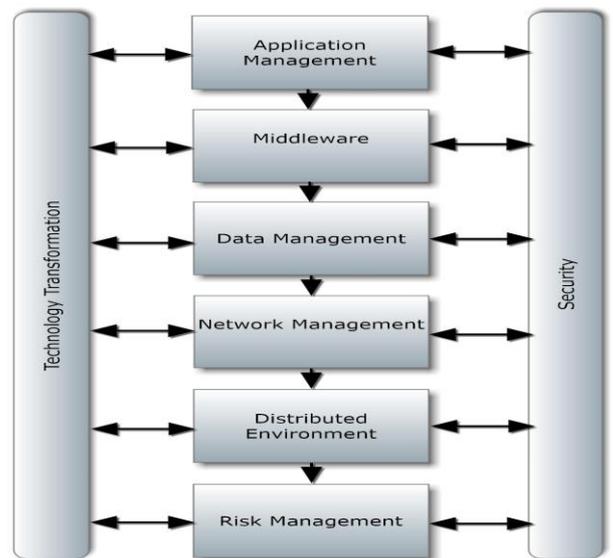


Figure 4. Information System Model

**F. Technology Architecture**

In the MAM one of the architecture is technology architecture (TA) which is the combinations of new technological principles or validated technologies, architectural vision, target data model and application process which construct a technical model (TM). The objective of this step is to map the input to organization’s technological infrastructure to service domain.

**G. Chain Management**

The objective of chain management is to organize and manage organizational resources to ensure the flow of product and services in a MAM. The activities include the

transformation of top to down of the organizational principles, governances, works and services.

#### H. Migration Planning

In the MAM, there should be proper support for change in the enterprise. Requirement for change may be in different aspects. Such as there may be a need to update the roadmap of the architecture. It is important that the resultant value which is expected to add to the business should be properly assessed. The period of transition and transition plan are also to be included in the migration planning. The risk and the capability should be the properly weighted.

#### I. Implementation Governance

In implementation governance phase of MAM provides architectural oversight for the implementation of architecture. In the implementation phases, organizational standard, procedure, architectural approach and operational framework are to be maintained for the organizational-specific development. For the maximum benefit to be obtained, the transformed architecture should be realized through sequence of transformation steps. The other benefit of this approach is that it will minimize the migration and transformation risk. So, the overall transformation is an incremental process and the target architecture is achieved through increments.

#### J. Change management

In this phase, change management is described the development of the enterprise architecture for the fit stages. So far It has been performed strategic direction and top-down architecture of a project generation to achieve corporate capabilities. However, enterprise architecture does not operate in a vacuum. There is usually an existing infrastructure and business which is already providing value. The architecture change management process is to ensure that the architecture achieves, its original target business value. This includes managing changes to the architecture in a cohesive and architected way.

#### K. Single Source Identification Model (SSIM)

We propose here a single source identification model where the module is responsible for performance monitoring, investigation, ensuring high security, acceleration of activity and easy distribution of the information. All phases are connected with SSIM, so we get some answers of what, how, where, who, when and why about the person/project's status. SSIM provides top most monitoring and controlling of the enterprise's people, processes and resources.

#### L. Model of Technology-transformation

One of the purpose of any model that the potential of it should sustain over the next two to three decades with its

competitive advantage. In this model the initial step is to identify the demand of technology. Then gap analysis is performed between current and target technologies. HR management or sustainable HR is the winning soldiers of this stage. Right place, right time, right people and right action will give the competence with combination of skill, knowledge and attitude. Technical mechanism and instrument provides the new or up-to-date infrastructure of the enterprise. All the processes are connected to re-engineering stages for sustainability.

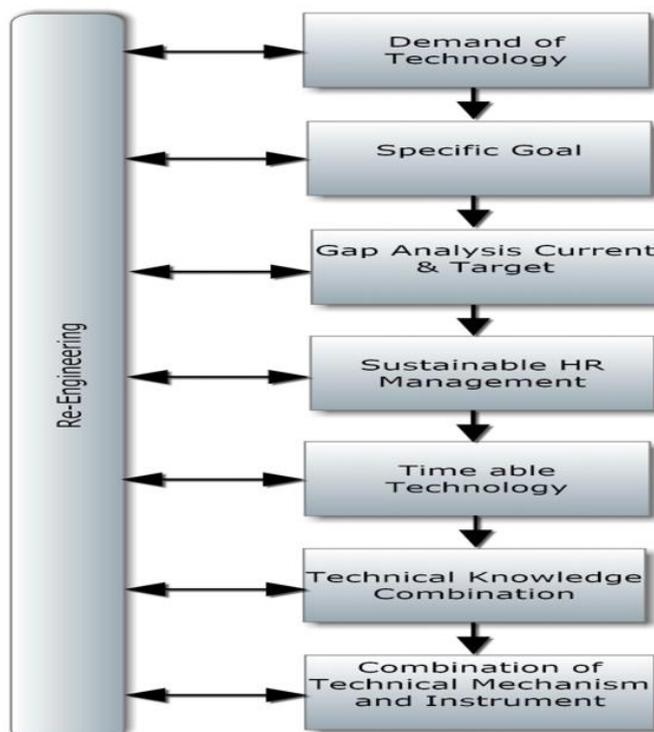


Figure 5. Model of Technology-transformation

#### M. Security Architecture

Data classification stage classifies or levels the enterprise's data which is the more or less secured. Vulnerability management finds out or manage all the weak points of the organizational activities which is to be reported for the action. Threat and attack management provide all aspect of the enterprise's project requirement. Software and hardware security management ensure unauthorized access of the organizational functions of using different type of tools. For the importance of resources and physical security, It is given the secured place with the leveling and preventing policies. Resources and information are supervised with log management and monitoring system which is used updated tools and technologies. Security kernel is the core of the SA which controls everything into the architecture. It is a facilitated computing or controlling system of the kernel that is combination of firmware and software. This software based security kernel is designed with some security techniques and mechanisms.

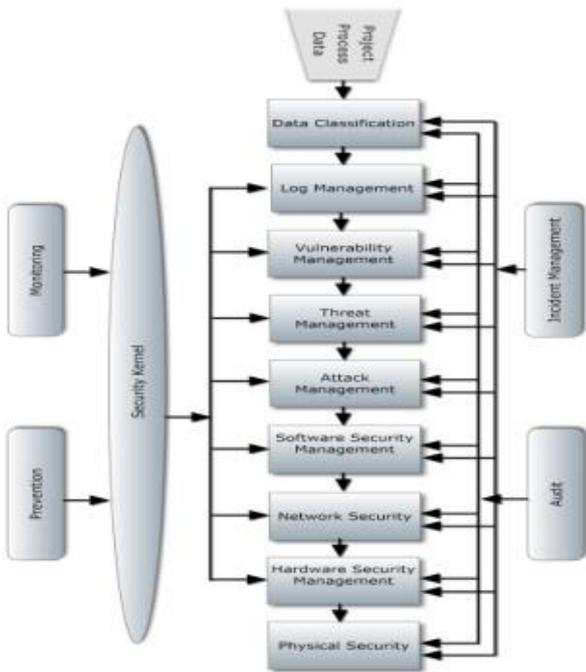


Figure. 6. Security Architecture

N. Knowledge Based (KB)

A knowledge base is required which is important for storage requirement and at the same time an associated intelligent interface is needed. Interface will provide a user friendly and application specific data retrieval and storage facility. KB is connected with technology- transformation, initial stage and SSIM for performing data sharing, data mining, and archiving and big data analysis of the enterprises.

IV. COMPARISON AND DISCUSSION

In this section we provide a subjective comparison of the proposed model with those of well-known Zachman [18], TOGAF [17], FEA[21] and Gatner[19] enterprise architectures.

They all have the similar architectures. But the proposed SEAST is also a combination of processes, reference model and taxonomies. But in the proposed model we used kernel-based security architecture for focusing security and single source identification for monitoring. Knowledge base is required for research and big data analysis. Technology transformation module provides support for adoption of new technology.

We present in Table I comparative rating for different models with ours. SEAST is better than others in some ratings because others lack those features. SEAST has similar features as of TOGAF hence ratings are same.

TABLE I. Comparisons of SEAST, Gatner, FEA, Togaf and Zachman architecture [18-21]

Criteria	Ratings				
	Zach man	TO GA F	FE A	Gart ner	SEA ST
Taxonomy completeness	4	2	2	1	2
Process completeness	1	4	2	3	4
Reference-model guidance	1	3	4	1	4
Practice guidance	1	2	2	4	2
Maturity model	1	1	3	2	1
Business focus	1	2	1	4	2
Governance guidance	1	2	3	3	3
Partitioning guidance	1	2	4	3	2
Prescriptive catalog	1	2	4	2	1
Vendor neutrality	2	4	3	1	2
Information availability	2	4	2	1	4
Time of value	1	3	1	4	3
Security	1	2	2	1	4
Knowledge base	1	1	1	1	4
Technology transformations	1	1	1	1	4
Single Source identification	1	1	1	1	4

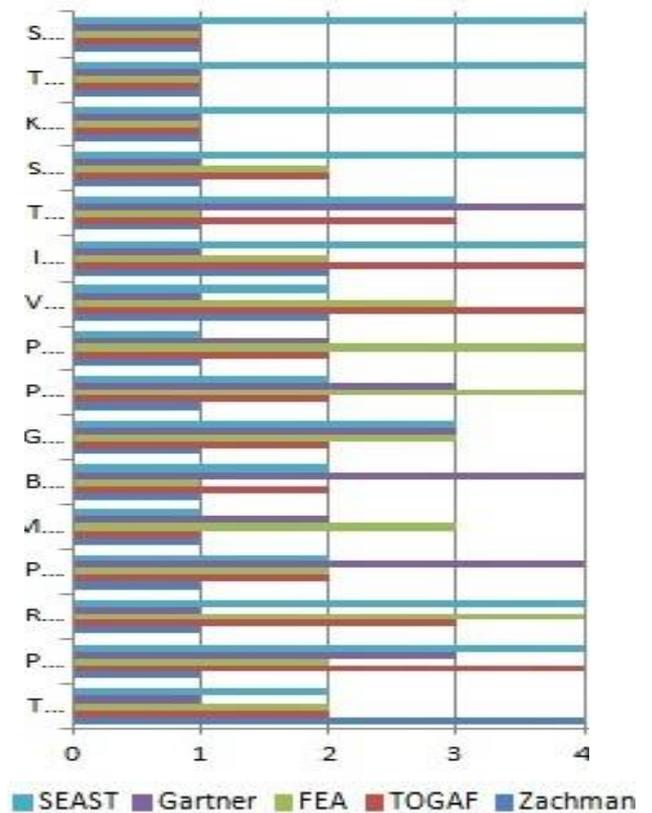


Figure 7. Graphical comparison on EAs

In Fig. 7, (S=Single Source identification, T= Technology transformations, K= Knowledge base, S= Security, T= Time of

value, I= Information availability, V= Vendor neutrality, P= Prescriptive catalog, P= Partitioning guidance, G=Governance guidance, B= Business focus, M= Maturity model, P=Practice guidance, R= Reference-model guidance, R= Process completeness, T=Taxonomy completeness) we show the graphical view of comparative ratings of some well-known EAs and SEAST. It is seen from the figure that SEAST is a more robust architecture than others.

#### V. CONCLUSION

A secure, simplified, sustainable enterprise architecture has been proposed in this paper. The proposed framework is the combination of business strategy, business processes, technology transformation, information system, application, data, security, infrastructure and comprehensive architecture plan. Government or nongovernment organizations or any business organization can use this model to address their current and future business continuity. It is an easily understood sequential and connected architecture. A kernel-based security module ensures security and safety of the enterprise through auditing, preventing and monitoring mechanisms. New technology can easily be adopted for the business through the provision of technology- transformation module. A single source identification process has also been introduced which is responsible for monitoring and controlling all activities based on unique identity.

#### REFERENCES

- [1] S. Al Nasrawi and M. Ibrahim, "An enterprise architecture mapping approach for realizing e-government," in third international Conference on Communications and Information Technology (ICCIT), 2013, pp. 17-21.
- [2] T. Kaddoumi, M. Watfa "A Proposed Agile Enterprise Architecture Framework" The sixth international conference on innovative computer technology (INTECH), 2016
- [3] I. Shaanika and T. Iyamu, "Development of Enterprise Architecture in The Namibian Government: The use of Activity Theory of Examine the Influencing Factors", The Electronic Journal of Information System in Developing Countries. Vol.71, 2015.
- [4] J. J. Korhonen, J. Lapalme, É. de, Doug McDavid, "Adaptive Enterprise Architecture for the Future " 2016 IEEE 18th Conference on Business Informatics, 2016
- [5] V. Weerakkody, M. Janssen, and K. Hjort-M adsen, "Integration and enterprise architecture challenges in e-government: a European perspective," International Journal of Cases on Electronic Commerce (IJCEC), vol.3, pp. 13-35, 2007
- [6] L. Matthias, J. Mendling, and J. Recker. "A comprehensive EA benefit realization model - An exploratory study." System Science (HICSS), 2012 45<sup>th</sup> Hawaii International Conference on. IEEE, 2012.
- [7] J. Hagan, J. Paula "Guide to the (Evolving) Enterprise Architecture Body of Knowledge." MITRE Corporation).
- [8] Zachman, John A. "Concepts of the framework for enterprise architecture." Zachman International, Inc., La Canada, Ca (1997).
- [9] R. BabakDarvish, H. Shirazi, A. FarahmandNezhad, and S. Kharazmi. "Presenting a framework for agile enterprise"
- [10] M. Alaeddini, &S. Salekfarid, "Investigating the role of an EA project in the business-IT alignment in Iran", Information Systems Frontiers, 15(1), pp.67-88, 2013.
- [11] S.Bente, U. Bombosch ,& S. Langade, "Collaborative Enterprise Architecture: Enriching EA with Lean, Agile, and Enterprise 2.0 Practice", 2015
- [12] L. Urbaczewski and S. Mrdalj, "A COMPARISON OF EA FRAMEWORKS", Issues in Information Systems, 7(2), pp.18-23. Available at: <http://www.iacis.org/iis>, 2016.
- [13] A. Lapkin, et al., "Gartner Clarifies the Definition of the Term `EA`", Gartner, (August), p.6. Available at: <http://my.gartner.com/portal>, 2017
- [14] F. Armour, S. Kaisler, & S. Liu, "A big-picture look at enterprise architectures", IT Professional, 1(1), 1999.
- [15] B. Cameron, & E. Mcmillan, "Analyzing the Current Trends in EA Frameworks", Journal of Enterprise Architecture, February (9) pp. 60-71. Available at: <http://ea.ist.psu.edu/documents>, 2017
- [16] R. Winter & E. Sinz, "Enterprise architecture", Information Systems and e-Business Management, 5(4), pp. 357-358, 2007
- [17] The Open Group. TOGAF 9.1, 2011. <https://www.opengroup.org/togaf>
- [18] J. Zachman, "The Zachman Framework", available at: <https://www.zachman.com/resources>, 2017
- [19] The Gartner Framework Analysis", <https://www.gartner.com/doc/1987417/enterprise-architecture-process-framework-cio>, 2017
- [20] C. M. Firmansyah, Y. Bandung, "Designing an Enterprise Architecture Government Organization Based on TOGAF ADM and SONA", International Conference on Information Technology Systems and Innovation (ICITSI), 2016
- [21] R. Sessions "A Comparison of the Top Four Enterprise-Architecture Methodologies" Object Watch, Inc. May 2007.

# Distributed Query Processing and Data Sharing

Ahana Roy, Aspen Olmsted

Department of Computer Science  
College of Charleston, Charleston, SC 29424  
roya@g.cofc.edu, olmsteda@cofc.edu

**Abstract**— Data privacy is becoming an increasing concern in industries where huge data repositories are dealt with. Most of them need to share private data across distinct organizations. In this paper, we propose a modification to an existing framework wherein data sharing across distributed databases is handled without compromising the sensitive nature of information. This paper focuses on replacing an inner join with semi-join reducer technique in one of the distributed sites of an existing query execution model. We will demonstrate how transmitting redundant data can be eliminated deploying a reducer technique which ultimately results in minimization of communication cost.

**Keywords:** Data privacy, Semi-joins, Query optimization, Distributed database, Transmission cost

## I. INTRODUCTION

Distributed database system [1] [2] [3] can be defined as a collection of logically interrelated data distributed over several sites. Storage of user-specific data in a system and sharing of the same with multiple organizations has become very common especially with the rise of Internet-connected systems. For example, in a hospital, certain patient-specific information should not be disclosed to drug providers and others, instead be limited to highly specialized practitioners. At the same time, besides maintaining privacy, we also need to optimize queries on a network of distributed machines, as sharing of data comes into the picture. In a distributed database system, the database query extracts data from several different sites, so in this case, the important factor is to reduce the amount of data transmission to the maximum extent.

We propose a modification to the query execution architecture initially proposed in [4]. There are three essential components in this design, namely data owner (or site 1) where the information is stored, a query asker (or site 2) or a participating organization which runs SQL-like queries against distributed data and a randomly generated blind comparer (or site 3) which stays aloof of the relation contents, performs required joins and returns the results to the query asker. The introduction of semi-joins in the blind comparer to optimize query performance across distributed database system is our aim. There is a need to design query plan in a way such that it minimizes communication cost in transferring relations from one site to another in a distributed database. With semi-join reduction [5], only desired part of a relation is shipped to the relevant site where a natural join is performed. In this process, we need to compare sizes of the relations and consider a semi-join if there is a possibility of elimination of dangling tuples from a particular relation. Similarly, a sequence of semi-joins

can be profitable as it acts as a full reducer for more than two relations in view. We shall denote semi-join of relations  $R(X, Y)$  and  $S(Y, Z)$ , where  $X$ ,  $Y$ , and  $Z$  are sets of attributes, as  $R \bowtie S = R \bowtie_X (projection_Y(S))$ . That is, we project  $S$  onto the common attributes (in this case  $Y$ ), and then take the natural join of that projection with  $R$ .  $Projection(S)$  is a set-projection, so duplicates are eliminated.

The organization of this paper is as follows. Section II describes related work trying to improve query plans using semi-join across distributed machines on a network. In Section III, we give an overview of the proposed query execution model. Section IV describes the performance of using semi-join instead of natural join and the benefits and limitations of using the same. Section V contains a conclusion and possible future work in generating improved query optimization techniques across distributed databases.

## II. RELATED WORK

Researchers have found several optimal and suboptimal reduction to reduce the total amount of data transferred between sites in a distributed database (in contrast to traditional centralized databases). There are many algorithms that employ semi-join operator for the reduction phase of a query optimization strategy [6] [7] [8] [9] [10] [11]. Semi-joins never increase the size of any join. A combination of joins and semi-joins is proposed in [12]. The authors stated that cyclical query graphs could not be fully reduced by semi-joins alone, so a combination of joins and semi-joins is required. More recently, methods based on hashing have been investigated [13]. Bloom filters were first introduced in [14] and had been used in distributed query processing [15] [16] [17] [18]. The hash-semi join [13] is based on Bloom filters and produces good results. PERF joins were introduced in [19]. The PERF join works with the tuple scan order and thus limiting collisions. In addition, PERF joins do ensure that relations being joined are reduced as much as possible before the final join. In [20] a variation of bloom filter called Complete Reducing Filter (CRF) is proposed. Based on a combination of composite semi-join and PERF join, the main purpose of it is to lower the transmission cost by keeping the join location information. Composite semi-joins is also a possible query optimizing plan [21]. Details about 2-way semi-joins are discussed in [22]. Here, the authors establish that such joins should be used when the selectivity of the reducing attribute is either high or low. They recommend using two separate semi-joins when the selectivity is in a middle range. In [23] the authors implemented three different strategies to compute semi-join with simple SPJ queries. The essentially

compare the performance involving joins involving one or more than one joining attribute. Our motivating example [4] proposed a system for storing and querying private data in a distributed manner. In this paper, we try not to compromise the three privacy features: data privacy, query privacy and anonymity of communication at the same time keeping query optimization in main consideration.

### III. ARCHITECTURAL MODEL AND QUERY EXECUTION

#### A. Architectural Model

The query execution framework introduced in [4] splits querying into two phases. Phase I performs a global search for records specific to a person which in turn returns data handles which indicate that the record exists without revealing the exact location in the network. Phase II uses data handles to execute relational algebraic query concealing original data and query from the query asker and data owner component respectively. In this paper, we shall assume query asker has sufficient authorization over the records requested. Fig. 1 shows the procedure for executing a query involving a semi-join across three nodes in a distributed database.

- Step 1: Query asker uses the data handle to route a message to the data owner requesting a join. The authors in [4] propose that each provider selects a random number, hashes it and transmits to either of them. The sum of this hashes is used to select a blind comparer where the join of two relations is performed. Two onion skin routes are used in this process to route information from query asker, and data owner, identities of these are thus hidden from the blind comparer.
- Step 2: Original query is split into two parts. The first part consists of a select statement which gets sent to relevant data owners using data handles as onion skin routes. These data handles also contain instructions to send the results of a query execution to the blind comparer using the specified route. Our architecture proposes that an additional instruction should be present to route a projected part of the query back to the query asker.
- Step 3: Query asker projects the second part of the original query based on the common attribute or the predicate in the where clause of the original query. This is then shipped to blind comparer which performs the natural join of relation returned from data owner and the subset of the remainder query shipped from query asker.
- Step 4: Blind comparer returns the join result to query asker.

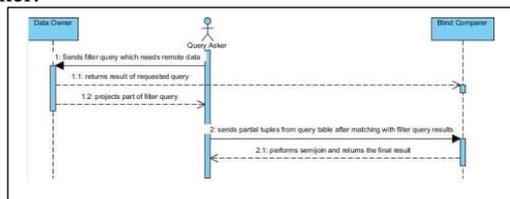


Figure 1. Query execution model

#### B. Query Execution

We shall demonstrate our proposal using the SQL query as an example used in [4]. One important advantage of a distributed system is the ability to replicate data, that is, to make copies of the data at different sites. One motivation is that if a site fails, there may be other sites that can provide the same data that was at the failed site [5]. Suppose, there has been a database crash at the query asker site (pharmacist in this case). A backup of patient information is kept with data owner site. The pharmacist is not authorized to retrieve any data other than the pickup date record of a patient vulnerable to a specific drug. Fig. 2 shows the use case diagram.

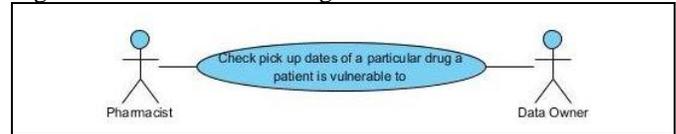


Figure 2. Use case diagram of pharmacist model

The corresponding SQL query may be written as:

```
SELECT d.pickupdate
FROM (SELECT * FROM conflicts c
      CROSS JOIN nonces n
      INNER JOIN remote(drug_history) d
      ON n.nonce = d.nonce
      WHERE c.drug = d.drug
);
```

The conflicts table in the above query is a list of the conflicting drugs available as shown in Table 1. The nonces table is shown in Table 2. It contains a set of data handles for the patient as found in a global search operation in [4]. For simplicity, we assume each of the above two tables consists of two records only.

TABLE 1. conflicts

drug
A
B

TABLE 2: nonces

nonce
34
56

Fig. 3 shows the parse tree which is executed in a bottom-up fashion starting at the leaves for the local tables (conflicts and nonces). After the cross-join operation on local tables, we need to join on a remote table (a table residing on another site, data owner). In this paper, we assume the filter query is specified by the query asker (in this example, the pharmacist) for each data handle. The data owner is also instructed to send the relevant result of the query to a randomly selected blind comparer. A part of the execution result is also asked to return

to query asker in certain situations, for example when the cross-join result of local tables is huge.

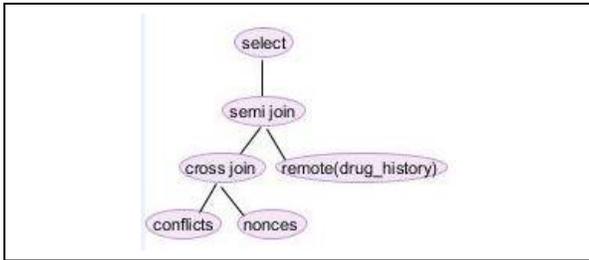


Figure 3. Parse tree of original query

A filter query may be written as:

```

SEND (
    SELECT d.nonce, d.drug, d.pickupdate
    FROM drug_history d
    WHERE d.nonce = 34
);
    
```

From the original query, we can see that the where clause equates over a common attribute drug. The partial or projected filter query result that should be sent back to query asker may be written as:

```

SEND (
    SELECT d.drug
    FROM drug_history d
    WHERE d.nonce=34
);
    
```

As soon as the query asker gets a list of drugs from the data owner for a patient, it equates with the matching column of the query table shown in Table 3TABLE.

TABLE 3. Query\_table

drug	Nonce
A	34
A	56
B	34
B	56

Now, suppose we get from data owner the drug as “A” for nonce 34, then we only send the first row of Table 3TABLE to the blind comparer eliminating the third row.

#### IV. PERFORMANCE REVIEW

We shall analyze joins and semi-joins and consider scenarios where these might outperform each other.

- Data owner returns null record for a patient (or nonce) on executing the filter query. In that case, sending the query table (or the cross-join result containing four rows) from the query asker to blind comparer would be a waste. If the query table consists of huge records, then the existing design will have more transmission

cost which is fruitless. An intermediate step as proposed here shall act as a check post which determines whether all or part or none of the query table needs to be shipped to the joining site.

- The selection of joins in the distributed system depends on the data transmission from one site to another. In our case, if transmitting the whole of the cross join of local tables cost more, we can think of reducer algorithms using semi-join as demonstrated in this paper. The following pseudo-code [24] may help determine the selection of join technique:

- Step 1: Read query table at query asker (Table 3TABLE)
- Step 2: Read the filter query projection from data owner to query asker containing the join attribute (JA)
- Step 3: Compute the cost of semi-join operation denoted as  $Scost$ . We denote resultant table which is the result of semi-join (joining join attribute with joining table or JT followed by selection and projection operation) as R. We can compute  $Scost$  as follows:  
 $Scost = Cost (transmitting JA from data owner to query asker) + Cost (projecting JA over JT) + Cost (R)$
- Step 4: Compute the cost of join operation or  $Jcost$ . This is aligned as per the existing design of query execution model in Fig. 1 where an inner join is performed between the results of filter query from data owner and query table from query asker. The cost of performing such a join is as follows:  
 $Jcost = Cost (transmitting filter query from data owner to blind comparer) + Cost (transmitting query table from query asker to blind comparer) + Cost (R)$
- Step 5: IF ( $Scost < Jcost$ ) Then  
 Execute Operation with Semi-Join  
 Else  
 Execute Operation with Join  
 End IF

#### V. CONCLUSION

We have discussed semi-join reduction method as a query optimization technique in distributed query processing involving data sharing across different machines on a network in which data privacy is also a concern. Semi-joins do implement more operations as compared to joins, but it reduces the number of bytes transferred from one site to another to a large extent. Further, it is to be noted that if transmission cost is of main consideration, then semi joins are beneficial to joins. We will experiment scalability of this approach, and also a combination of join and semi-join using real-time data in future and measure the effectiveness of the same. Also, we intend to evaluate the performance of PERF joins, composite semi-joins, 2-way semi-joins and other algorithms as discussed in Section

II and measure how beneficial or profitable they are compared to various other optimizing plans across distributed databases that have the constraint of revealing minimum information to other sites.

REFERENCES

[1] N. Mukherjee, "Synthesis of Non Replicated Dynamic Fragment Allocation Algorithm in Distributed Database System," in *Proceeding of international conference on advances in Computer Science*, 2010.

[2] T. V. Kumar and V. Singh, "Distributed Query Processing Plans Generation Using GA," *International Journal of Computer Theory and Engineering*, vol. III, 2011.

[3] R. Ghaemi, A. M. Fard and H. tabatabee, "Evolutionary Query Optimization For Hetrogenous Distributed Database System," in *World Academy of Science, Engineering and Technology*, 2008.

[4] M. Siegenthaler and K. Birman, "Sharing Private Information Across Distributed Databases," in *Eighth IEEE International Symposium on Network Computing and Applications*, Cambridge, MA, 2009.

[5] H. Garcia-Molina, J. D. Ullman and J. Widom, *Database Systems: The Complete Book (2nd Edition)*, Pearson.

[6] P. Bernstein, N. Goodman, E. Wong, C. Reeve and J. Rothnie, "Query processing in a system for distributed databases," *ACM Transactions on Database Systems*, pp. 105-128, 1981.

[7] P. Black and W. Luk, "A new heuristic for generating semi-join programs for distributed query processing," in *IEEE COMPSAC*, 1982.

[8] P. Apers, A. Hevner and S. Yao, "Optimization algorithms for distributed queries," *IEEE Transactions on Software Engineering*, pp. 51-60, 1983.

[9] "Improvement algorithms for semi-join query processing programs in distributed database systems," *IEEE Transations on Computers*, pp. 959-967, 1984.

[10] H. Kang and N. Roussopoulos, "Using 2-way semi-joins in distributed query processing," in *3rd Int. Conf. on Data Engineering*, Los Angeles, 1987.

[11] N. Roussopoulos and H. Kang, "A pipeline N-way join algorithm based on the 2-way semi-join program," *IEEE Transaction on Knowledge and Data Engineering*, pp. 486-495, 1991.

[12] M. Chen and P. S. Yu, "Combining join and semi-join operations for distributed query processing," *IEEE Transactions on Knowledge and Data Engineering*, pp. 534-542, 1993.

[13] J. Tseng and A. Chen, "Improving distributed query processing by hash-semijoins," *Journal of Information Science and Engineering*, no. 8, 1992.

[14] B. Bloom, "Space/time tradeoffs in hash coding with allowable errors," *Comm. ACM*, pp. 422-426, 1970.

[15] D. DeWitt, S. Ghandeharizadeh, D. Schneider, A. Bricker, H. Hsiao and R. Rasmussen, "The GAMMA database machine project," *IEEE Transactions on Knowledge and data engineering*, pp. 44-62, 1990.

[16] G. Qadah, "Filter-based join algorithms on uniprocessor and distributed-memory multiprocessor database machines," *Lecture Notes in Computer Science, Springer Verlag*, pp. 388-413, 1988.

[17] G. Qadah and K. Irani, "The join algorithms on a shared-memory multiprocessor database machine," *IEEE Transactions on Software Engineering*, pp. 1168-1683, 1988.

[18] P. Valduriez and G. Gardarin, "Join and semijoin algorithms for a multiprocessor database machine," *ACM Transactions on Database Systems*, pp. 133-161, 1984.

[19] Z. Li and K. Ross, "PERF join: an alternative to two-way semijoin and Bloomjoin," *Proc. Of CIKM*, pp. 137-144, 1995.

[20] Y. Zhang, "Variation of bloom filters applied in distributed query optimization.," *Electronic Theses and Dissertations. Paper 4508*, 2003.

[21] W. Perrizo and C. Chen, "Composite semijoins in distributed query processing," *Information Science*, pp. 197-218, 1990.

[22] N. Subhani and J. Morrissey, "Investigating the 2-Way Semijoin for Distributed," Ontario.

[23] S. M. Mahajan and V. P. Jadhav, "Tri-variate Optimization Strategies of Semi-Join Technique on Distributed Databases".

[24] M. Sharma, D. G. Singh and R. Virk, "Analysis of Joins and Semi Joins in a Distributed Database Query".

# Factors affecting the implementation of information assurance for eGovernment in Indonesia

Rio Guntur Utomo, Robert J Walters, Gary B Wills

School of Electronics and Computer Science

University of Southampton

Southampton, United Kingdom

{rgu1n15, rjw1, gbw}@soton.ac.uk

**Abstract**—Electronic government (eGovernment) services are aimed to improve government services to the public by improving the quality and availability of services that can be accessed regardless of time and place. Consequently, the services must always be available at any time, and any threat to the information and systems should receive attention to ensure business continuity in the event of an incident. Accordingly, in the implementation of eGovernment, information assurance (IA) should be considered. In Indonesia, the implementation of eGovernment is still in the early stage according to the eGovernment roadmap of Indonesia. However, there is no study so far that has focused on IA for eGovernment in Indonesia. Therefore, research on a framework of IA is needed to support the implementation of eGovernment in Indonesia. The aim of this research is to develop an IA framework for eGovernment within the Indonesian context. The development of the framework is divided into four stages, which are identifying the factors from IA international standards for organisations, determining success factors from literature, identifying the challenges, and evaluating and harmonising all the factors categorised into three categories. This research proposes an IA framework, which is expected to assist eGovernment implementation in Indonesia.

**Keywords**—information assurance; egovernment; information security; influential factors; implementation;

## I. INTRODUCTION

In the modern world, technology is an integral part of everyday life and cannot be separated from progress and human development [1] and from everyday life [2]. One of the technologies is the Internet, which continues to grow and evolve and has become an integral part of daily life [3]. Since 2014, Internet users in Indonesia have increased from 34.9% to 51.7% by the end of 2016 and will continue to grow for many years [4]. Internet technology has been used in various fields, such as business, health, and education [5]. Examples of popular Internet usage are for email, online shopping, and social media [6]. Internet technology has also facilitated government initiatives to improve their services, which is often called electronic government or eGovernment [7].

The World Bank [8] defined eGovernment as the use of information technologies (such as wide area network, the Internet, and mobile computing) by government agencies that could transform relations with citizens, business, and other government organisations. The term eGovernment in Indonesia refers to the use of information technology (IT) in the service

procedures organised by government organisations [9]. In Indonesia, the implementation of eGovernment initiatives began with the publication of the Presidential Instruction No. 3 in 2003 [10].

Implementation of eGovernment provides many advantages, such as improved quality of service in which eGovernment systems allow public, business, and government sectors to have access 24 hours a day, seven days a week to government information [11]. Reduced costs and process levels in organisations have streamlined the operational procedures, which are also benefits from the implementation of eGovernment [12]. In addition, the performance of government agencies in providing public services to customers will be more effective and efficient [13]. Moreover, the implementation of eGovernment will also increase the transparency and service to the public, improve service and efficiency, reduce transaction costs, and provide benefits from an economic perspective [14].

Despite the benefits of eGovernment, there are also problems regarding its implementation. The availability of services has become a significant concern [15]. Moreover, according to Basu [16], assurance of the security of the communications and its sources has also become an issue. Users are mainly concerned about the integrity of the communicated information. In addition, with eGovernment reliance on information systems and services, it is more vulnerable to threats and needs to be protected [17]. To overcome this problem, information assurance is needed as a mechanism to protect information systems and services.

The main purpose of IA is to protect the business by reducing risks associated with information and information systems [18]. The activity is driven by risk analysis and cost-effectiveness with a comprehensive and systematic management of security countermeasures [19]. Additionally, IA relies on multiple, related, organisational actions and controls in the form of the defence in depth model [20]. All IA processes are carried out to support corporate governance [21]. With services and business continuity assured, it is expected that the eGovernment services in Indonesia will be implemented successfully; therefore, the purpose of implementation of eGovernment will be achieved, which is to improve the effectiveness, efficiency, and quality of service to the citizens.

To implement eGovernment in Indonesia successfully, the IA of eGovernment in Indonesia requires attention. Therefore, the aim of this study is to develop a framework of IA to support to the implementation eGovernment in Indonesia.

## II. LITERATURE REVIEW

### A. eGovernment in Indonesia

The Indonesian government consider the use of IT advances to provide better public services. Consequently, it is necessary to establish a network of information and transactions of public services that have the quality and scope to reach and satisfy the wider community. Which is also need to be accessible in all parts of Indonesia at any time at a cost that is affordable to citizens in the form of eGovernment [10]

Indonesia officially began implementing eGovernment with the publication of the Republic of Indonesia Presidential Instruction No. 3 of 2003. The development of eGovernment is an effort to implement an electronic-based governance to improve the quality of public services effectively and efficiently. Through the development of eGovernment, management systems and processes carried out in the government environment are restructured by optimising the utilisation of IT. The utilisation of IT includes two related activities [10], among others:

(1) IT advances for public services can be accessed easily and cheaply by citizens throughout the country;

(2) Data processing, information management, systems management, and work processes are done electronically.

Since 2003, there has been a steady increase of local government capability on using and managing IT for their eGovernment initiative. A successive survey by the Ministry of Communication and Informatics in 2009 and 2011 showed a 3.7% increase in the eGovernment rating of provincial governments [22][23]. According to the eGovernment roadmap of Indonesia [24], until 2014, the eGovernment system development was still in the form of silos. Furthermore, from 2015 to 2018, eGovernment systems and national eGovernment infrastructures were being integrated.

The year 2019 will begin the optimisation era, where services like G2G, G2B, G2C, and G2E will begin to be implemented. According to Karokola [25], eGovernment service security is often not considered at the initial stages, and it is argued that it should be considered from the initial stage of information systems, IT, and information and communication technology (ICT) development for eGovernment services. It is expected that the development of a framework that focuses on IA for eGovernment in Indonesia can be helpful in the development of eGovernment services in Indonesia.

### B. IA Standard-based Frameworks

The International Organisation for Standardization (ISO) and International Electrotechnical Commission (IEC) standard [17], Information Assurance for Small and Medium Enterprises (IASME), and Control Objectives for Information and Related Technologies (COBIT) 5 are three best practice standard-based frameworks that have been developed and recognised

internationally. Although ISO/IEC 27001 is intended for ISMS, according to Hibbard [18] ISO/IEC 27001: 2013 is more closely aligned with IA.

The standard ISO/IEC 27001: 2013 is an ISMS standard developed and published by ISO and IEC. This standard provides guidelines for establishing, implementing, maintaining, and continually improving information security management systems. The standard states that the purpose of ISMS is to manage and control information security risk and to protect and maintain confidentiality, integrity, and availability. The standard also identifies the outcomes of an effective ISMS implementation, which are adequate control over information security, good governance in handling and securing information, and having a mechanism for measuring the success or absence of security control. The standard also has guidelines for integrating ISMS with organisational strategies based on ten key requirements [17].

In 2011, the IASME Consortium published the IASME document aimed at providing guidance on SMEs to assess and acknowledge the level of maturity of their business information security. The processes are adopted from international standards and EU guidelines, which are simple, fast, and cost-effective. The advantage of IASME is that this standard can be adjusted with other standards, such as ISO/IEC 27001, Publicly Available Specification (PAS) 555, Communications-Electronics Security Group (CESG) 10 Steps to Cyber Security, and Centre for the Protection of National Infrastructure (CPNI)/SysAdmin, Audit, Network, and Security (SANS) 20 Critical Controls for Cyber Defence.

In addition, IASME works by applying control sets to all business types and adjusting their implementation regarding the business risk profile. Although developed for smaller businesses, the IASME process can now be adapted to any business size. For effective implementation, IASME is based on 12 factors for its guidance [26].

Moreover, COBIT 5 is a framework of principles, practices, analytical tools, and models that are globally accepted and can help in identifying critical business issues related to governance and management of information and technology for enterprises. The COBIT 5 for Assurance focuses on defining assurance objectives that align with enterprise objectives by maximising the value of assurance initiatives [27]. In addition, COBIT 5 provides guidance for establishing and sustaining assurance for enterprises and provides a structured approach on how to provide assurance over enablers.

The scope of COBIT 5 for Assurance consists of two perspectives, namely, the assurance function perspective that describes what is needed in building and the assurance and assessment perspective functions that describe which assurance needs to be provided. These two perspectives are built on the seven common governance and management enablers of the COBIT 5 framework [27].

Although these factors are from international standards, it is still necessary to identify other factors from relevant literature that are expected to complement IA aspects that are not covered by the standards.

### C. IA Critical Success Factors

According to Bullen and Rockart [28], critical success factors (CSFs) are a limited number of areas that must be met and implemented properly for the organisation to achieve its goals and objectives. A good implementation of CSFs will ensure the successful performance of an individual, department, or organisation [28].

There are several studies on the CSFs of IA implementation within organisations. Birchall et al. [29], stated in their study that the business strategy and strategic direction of the organisation affect IA. Moreover, IA is often only considered a technical problem, but in practice, IA should be approached holistically, which is connected to business and strategy.

The Ministry of Information and Communications Technology [30], published a National Information Assurance Policy (NIASP) that can be used in all sectors. The NIASP presents the necessary and relevant foundation for implementing an ISMS in the organisation. According to MICT, key factors in protecting information are awareness and education for users and employees, as they are the users and managers of that information.

As a part of the 'Information Assurance Strategy NHS Lanarkshire', Tannahill [31] described IA as a process to ensure the confidentiality, integrity, and availability of information assets. The strategy of IA outlined the key areas of strategic focus and action, which involve leadership, governance, risk management, policy, operations, and monitoring and compliance of effective IA.

Cherdantseva and Hilton [19], published a reference model of information assurance and security (RMIA). This model conveys a perception of IA as a complex organisational and managerial concern and requires comprehensive and systematic treatment. The model also listed security countermeasures in implementing IA and security, which are focused on organisational, technical, legal, and human-oriented aspects.

In 2014, Lincolnshire Police [32] released an IA strategy, standards, and working practices. The strategy aimed for further development of IA capability. The IA strategy was based on information management values, such as standards, business management, people management, information sharing, and data/information management.

A group within the UK Government Communications Headquarters (GCHQ), CESG, published 'The Information Assurance Maturity Model and Assessment Framework' [33]. The IAMM consists of five levels with three main IA goals. The key process for the IA goals are leadership and governance; training, education, and awareness; information risk management; through-life IA measures; assured information sharing; and compliance. For organisations, these processes facilitate achieving the maturity to accomplish trust in the information systems and processes, both internally and between organisations.

The Combined Communications Electronics Board (CCEB) published 'Information Assurance for Allied Communications and Information Systems' [34], which was intended for the five member nations of Australia, Canada,

New Zealand, the United Kingdom, and the United States. The document listed IA principles as well as components and defines the IA policies and procedures to enable a secure combined information environment.

Chris Cope, the lead auditor of ISO27001 and a CESG certified professional, listed principles for effective IA [35]. The principles are intended to enhance the security of any organisation. The principles have strong emphasis on business alignment, a holistic and risk-driven approach, and good governance within a less policy-constrained environment, as CESG withdrew the mandatory requirement to use Information Assurance Standards 1 and 2 (IAS1 & IAS2).

### D. Challenges

Expensive and uneven infrastructures become an obstacle to the implementation of eGovernment in the local government; it also resulted in limited access to eGovernment in certain places [36]. The Indonesian government does not have an interconnected government network and does not have a secure government network [24]. Moreover, at the local level, the available bandwidth is around 128 kbps to 1 Mbps, and the central level and citizens face a bandwidth of 10 to 200 Mbps and 100 Mbps to 1 Gbps, respectively [24]. The length of optical fibre that is available is around 50,000 km. The Indonesian government owns about 300 data centres that are not yet integrated with each other and are without backup or a disaster recovery plan [24].

Besides infrastructure, in implementing the eGovernment initiatives, cultural issues play a key role as these also influence any organisational changes regarding initiatives [37][38]. Culture is defined as values, beliefs, norms, and behavioural patterns of a group [39] that dictate how people think, solve problems, make decisions, and behave [40]. In Indonesia, some cultural issues influence the process of eGovernment implementation, such as the following no culture of sharing, no culture of documenting, no full support from leaders, resistance toward openness, and resistance to change of mindset [36][41][42].

In addition to infrastructure and cultural factors that hinder the implementation of eGovernment, according to Khalil et al. [43], in developing countries such as Indonesia, the digital divide factor and trust and privacy must also be considered in the implementation of eGovernment. Differences in class, race, ethnicity, and geography in developing countries, such as Indonesia, have resulted in the emergence of a gap in access to technology, especially the Internet [43]. In addition, to achieve the successful implementation of eGovernment, trust must be established between government institutions as well with the citizens. With the enormous amount of user information that must be managed, regarding the issue of privacy of information, the government should consider the responsibility with the intention that the user information is well protected [43].

Other challenges regarding security in Indonesia are organisational structure and coordination. The creation of an organisation like a National Cyber Agency to be in control of handling information security issues is required to oversee the other organisations that already exist in managing government

information security issues including eGovernment [44]. With many government institutions in Indonesia, there is a need for coordination between institutions so that the duties of each institution do not overlap in protecting eGovernment information [44][45].

From identifying these issues, it appears that many factors are challenging in the implementation of eGovernment services as well as IA, such as cultural issues, the digital divide, and trust for the users, along with infrastructure, coordination, and security, which affect the implementation of IA. These challenges are addressed in the proposed framework to ensure continuity of eGovernment services in Indonesia.

### III. THE PROPOSED FRAMEWORK

The development of the framework was divided into four stages. The first stage was identifying the IA factors from the international standards. The second stage was determining IA success factors from the literature, and the third stage was identifying the challenges. In the fourth stage, all 79 factors were evaluated and analysed to determine similar concepts and remove duplication. The remaining 18 factors were harmonised by their concepts and divided into three categories based on the scope and association. The three categories, namely, organisational management, implementation management, and Indonesian context, were based on the scope and association of the factors. Factors that are closely related to managerial and human aspects are incorporated into the organisational management category. Moreover, the more technical factors are categorised as implementation management. Finally, context-focused Indonesian factors are then incorporated into the Indonesian context.

#### A. Organisational Management

The following factors are associated with organisational management in implementing IA for eGovernment within the Indonesian context.

- Leadership and Commitment

Leadership and commitment from the top of the organisation in the implementation of IA are critical for the achievement of IA through the initial planning [31]. Top management must ensure that the policy and objectives of IA are in line with business needs. In addition, the availability of required resources should be ensured [17].

- Policy, Legal, and Compliance

The policy aims to guide the IA to be in line with business needs [26]. The legal department needs to ensure legal certainty for the use of information, intellectual property rights, and the use of software and other products [19]. Furthermore, the compliance of the information systems with policies and standards also need to be ensured [33].

- Management Review and Continual Improvement

Top management should undertake periodic review of the continuing suitability, adequacy, and effectiveness of the IA [17]. From the reviews, continual improvement for the

suitability, adequacy, and effectiveness of the IA is expected [17].

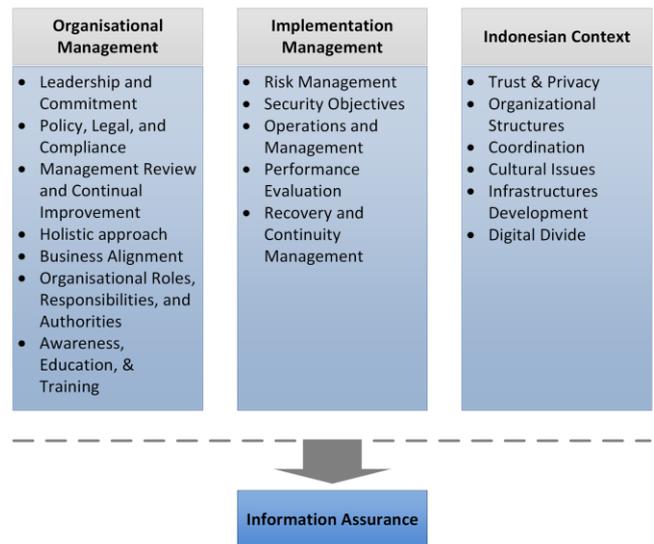


Fig. 1. Proposed Information Assurance Framework

- Business Alignment

In practice, IA do not work independently, but to support business processes [32]. The IA should be able to accommodate business needs (Bunker). Since business runs on risk, by planning IA that focuses on business objectives, business risk can be minimised and information held by an organisation is assured [35].

- Organisational Roles, Responsibilities, and Authorities

Top management must ensure roles in the organisation [30]. Ensuring responsibility and authority by top management is required to confirm IA is in accordance with the standard [17]. Top management should also receive reports related to IA performance [17].

- Awareness, Education, and Training

People who work in the organisation must be aware of the information policy and its contribution to the effectiveness and performance of the IA as well as the implications of not complying with the IA requirements [17]. Moreover, all employees working in the organisation are required to undergo relevant training and education corresponding to their job functions [30]. It is intended that all staff will be competent in their respective fields [27].

#### B. Implementation Management

The following factors are associated with implementation management in implementing IA for eGovernment within the Indonesian context.

- Risk Management

The first step in risk management is asset enumerating and planning, which aims to calculate assets owned by the

organisation and categorise them [17]. In addition, this process also plans the use of assets for the organisation. Assets that have been identified will be assessed to define the total risk of the information asset [26]. From the results of the assessment, the treatments for risks will be determined [17]. Risks cannot always be eliminated but can be minimised.

- Security Objectives

Information security objectives must be determined relevant to the functions and levels and be consistent with the information security policy [17]. Moreover, in determining the security objectives, the needs of information security as well as the results of the risk assessment and risk treatment must be considered.

- Operations and Management

Organisations must ensure the plan, implementation, and control needed to comply with information security requirements [17]. To ensure the optimal security, it is necessary to update the security systems in accordance with the latest updates from the provider [26].

- Performance Evaluation

Performance evaluation includes internal audits, monitoring, measurement, analysis, and evaluation. Internal auditing needs to be done to confirm whether the IA complies with needs of the organisation and the international standards [17]. Information relating to the effectiveness of implementation and maintenance of IA will also be obtained from the results of the audit. Monitoring, measurement, analysis, and evaluation are used to evaluate the performance and effectiveness of the IA. In addition to monitoring activities in the network, the organisation should also monitor the activities related to business activities [26].

- Recovery and Continuity Management

Backup and restore is the capability to maintain the integrity and availability of information systems during an incident or disaster [17]. In the event of major failures of information systems, business continuity must be maintained and work as usual [26].

### C. Indonesian Context

The challenges related to factors in the context of implementing eGovernment and IA in Indonesia as well as other developing countries are presented in the following section.

- Cultural Issues

Cultural issues need to be considered in the implementation of IA [27]. This is related to cultural issues that influence behaviour in organisations in Indonesia that can affect IA performance [36][41][42].

- Infrastructure Development

Implementation of eGovernment requires good infrastructures for the system to able to provide services as intended [27]. In addition, infrastructures are also needed for the information security process to be achieved according to

the security objectives [43]. Indonesia and other developing countries generally have difficulty in developing basic infrastructures [24][36].

- Digital Divide

Differences in class, race, ethnicity, and geography in developing countries, as well as Indonesia, have resulted in the emergence of the gap in access to technology, especially the Internet [43]. This issue should be addressed regarding IA performance, especially as under-developed regions still need to achieve the initial objectives.

- Trust and Privacy

To achieve the successful implementation of eGovernment, trust must be established between government institutions as well with the citizens [43]. With the enormous amount of user information that must be managed, regarding the issue of privacy of information, the government should consider the responsibility with the intention that the user information is well protected [43].

- Organisational Structures

The creation of an organisation like a National Cyber Agency to be in control of handling information security issues is required [46]. This organisation will oversee the other organisations that already exist in managing government information security issues including eGovernment [44].

- Coordination

With many government institutions in Indonesia, there must be coordination between institutions so that the duties of each institution do not overlap in protecting eGovernment information [44][45].

## IV. CONCLUSION

In summary, eGovernment is an initiative that aims to improve the quality of public access and the quality of public services. With the demanded services that must always available, assuring the continuity of eGovernment services becomes one of the problems in the implementation of eGovernment. One way to guarantee eGovernment services is with the implementation of IA. By combining technological, human, and organisational aspects that have a strong emphasis on strategic risk management, IA has a broad connotation that includes reliability, authentication, and nonrepudiation and provides restoration of information systems when an incident occurs, which ensures business continuity. In Indonesia, eGovernment is still in the development stage, and from the literature review, so far there is no IA framework that focuses on assuring eGovernment service in Indonesia. Therefore, this research proposed an IA framework for eGovernment in within the Indonesian context. It is important to identify factors for the effective IA implementation. By extracting factors from international standards ISO/IEC 27001:2013, IASME, and COBIT 5 and combining these with IA success factors as well as security and eGovernment challenges in the Indonesia context, the framework was developed and proposed. The framework consists of 18 factors, which are the result of the synthesis process of a combination of 79 factors. These factors

are categorised into three categories: organisational, implementation, and the Indonesian context. For future work, the framework will be validated using the triangulation method. The first step is the literature review, which has been done, with the result, which is the proposed framework. The next step is expert interviews and a practitioner survey that will be conducted with the aims to review and confirm the framework.

## REFERENCES

- [1] World Bank, "Technology & development," Global Economic Prospects 2008: Technology Diffusion in the Developing World, 2008.
- [2] OECD, "The case for E-Government: Excerpts from the OECD report: The E-Government imperative," OECD Journal on Budgeting, 3(1), 2003, pp. 1987–1996.
- [3] Techterms, "Internet," Techterms, 2015, Available at: <https://techterms.com/definition/internet> [Accessed 18 June 2017].
- [4] APJII, Bulletin APJII Edisi 05 November 2016..
- [5] Helsper, E.J. and Eynon, R., "Measuring types of Internet use," 2016 pp. 1–16.
- [6] Sooper Articles, "10 most common uses of internet," Sooper, 2009, [online] Available at: <http://www.sooperarticles.com/internet-articles/10-most-common-uses-internet-19405.html> [Accessed 28 Jun. 2017].
- [7] Dodd, J., "Delivering on the e-government promise," A government technology industry profile, 2000
- [8] World Bank. "e-Government," World Bank, 2015, [online] Available at: <http://www.worldbank.org/en/topic/ict/brief/e-government> [Accessed 9 Feb. 2017].
- [9] Sipatuhar, I.S. and Sutaryo, "Faktor-faktor penentu implementasi e-government pemerintah daerah di indonesia," Simposium Nasional Akuntansi XIX, 2016, pp. 24–27.
- [10] Presiden Republik Indonesia, "Kebijakan dan strategi nasional pengembangan e-government," Instruksi Presiden Republik Indonesia Nomor 3 Tahun 2003, 2003
- [11] Ndou, V., "E-government for developing countries: opportunities and challenges," The Electronic Journal on Information Systems in Developing Countries, 18(1), 2004, pp. 1–24.
- [12] Seifert, J.W., "A primer on e-government: sectors, stages, opportunities, and challenges of online governance," Report for Congress, 2003, p. 24.
- [13] Wang, H., & Rubin, B. L., "Embedding e-finance in e-government: a new e-government framework," Electronic Government, an International Journal, 1(4), 2004, 362–373.
- [14] Cohen, S. and Eimicke, W., "The future of e-government: a project of potential trends and issues," In: Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS-36), 2002 p. 146b (1-10).
- [15] Jaeger, P.T. and Thompson, K.M., "e-government around the world: lessons, challenges, and future directions," Government Information Quarterly, 20(4), 2003, pp. 389–394.
- [16] Basu, S., "E-government and developing countries: an overview," International Review of Law, Computers & Technology, 18(1), 2004, pp. 109–132.
- [17] The International Organisation for Standardization (ISO) and International Electrotechnical Commission (IEC), "ISO/IEC 27001:2013," September 2014.
- [18] Hibbard, E.A., "Introduction to information assurance," Storage Networking Industry Association, 2009.
- [19] Cherdantseva, Y. and Hilton, J., "A reference model of information assurance & security," 2013 International Conference on Availability, Reliability and Security, September 2013, pp. 546–555.
- [20] May, C. (CERT/CC T. and E.C.), "Defense in depth: foundation for secure and resilient it enterprises," (September 2006), p. 346.
- [21] Rathmell, A., Daman, S., O'Brien, K. and Anhal, A., "Engaging the board corporate governance and information assurance," Information Assurance Advisory Council (IAAC), 2004.
- [22] Kominfo, "Indonesia ict whitepaper," Pusat Data Kementerian Komunikasi dan Informatika, 2010.
- [23] Kominfo, "Pemeringkatan e-government indonesia (indonesian e-government ranking)," Direktorat e-Government. Kementerian Komunikasi dan Informatika, 2012.
- [24] Anggono, B.D., "eGovernment Indonesia update 2015-2019," Regional training workshop in Asia and the Pacific: Sustainable development and disaster risk management using E-Government, 2015.
- [25] Karokola, G.R., "A framework for securing e-government services: the case of tanzania," 2012.
- [26] Information Assurance for Small and Medium Enterprises (IASME), "The standard for information assurance for small and medium enterprises (iasme)," (2.3), 2013.
- [27] Control Objectives for Information and Related Technologies (COBIT) 5, "COBIT 5 for assurance," 2015
- [28] Bullen, C. V. and Rockart, J.F., "A primer on critical success factors," Working papers, (69), 1981, pp. 1–64.
- [29] Birchall, D., Ezingard, J.-N., McFadzean, E., Howlin, N. and Yoxall, D., "Information assurance: strategic alignment and competitive advantage," 2004.
- [30] MICT, "National information assurance policy 2.0.," CNSS Instruction No. 4009, CNSSI No. (4009), 2014.
- [31] Tannahill, C., "Information assurance strategy NHS Lanarkshire (4)," 2013.
- [32] Burge, R., "Lincolnshire police information assurance strategy, standards and working practices," June 2014, p. 59.
- [33] Communications Electronics Security Group (CESG), "The information assurance maturity model and assessment framework. (2.1)," 2015.
- [34] Combined Communications Electronics Board (CCEB), 2015. Information assurance for allied communications and information systems.
- [35] Cope, C., "10 Principles for effective information assurance," 2015.
- [36] Hardjaloka, L., "Studi penerapan e-government di indonesia dan negara lainnya sebagai solusi pemberantasan korupsi di sektor publik. rechtsvinding, 3(3)," 2014.
- [37] Avgerou, C., "Information systems for development planning. International Journal of Information Management, 13(4)," 1993, pp. 260–273.
- [38] Walsham, G., Symons, V. and Waema, T., "Information systems as social systems: Implications for developing countries. Information Technology for Development, 3(3)," 1988, pp. 189–204.
- [39] Leung et al., "Culture and international business: recent advances and their implications for future research," Journal of International Business Studies, 36(4), 2005, pp. 357–378.
- [40] Hall, E.T., "Beyond culture," Contemporary Sociology, 1976.
- [41] Wicaksono, D. A., "E-government in indonesia : the opportunities and challenges. Development," 2003, 1–2.
- [42] Djumadal, J. S., "Implementasi e-government, sebuah harapan penuh tantangan di provinsi daerah istimewa yogyakarta," In Konferensi dan Temu Nasional Teknologi Informasi dan Komunikasi untuk Indonesia, 2008.
- [43] Khalil, M.A., Lanvin, B.D. and Chaudhry, V., "The e-government handbook for developing countries," A project of InfoDev and the Center for Democracy & Technology, November 2002.
- [44] Ardiyanti, H., "Cyber-security dan tantangan pengembangannya di indonesia (cyber-security and the development challenges in indonesia)," Politica, 5 (dinamika masalah politik dalam negeri dan hubungan internasional), 2014, pp. 95–110.
- [45] Operananta, L., "Cyber security: Indonesia's challenges and opportunities to move forward," 2015.
- [46] Setiadi et al., "An overview of the development Indonesia national cyber security," International Journal of Information Technology & Computer Science (IJITCS), 6, 2012, pp. 106–114..

# *The Study of Public Organization's Intention to Use an Open government Data Assessment Application: Testing with an applicable TAM*

Chatipot Srimuang

Technoprenurship and Innovation Management Program  
Graduate School, Chulalongkorn University  
Bangkok, Thailand  
e-mail: joesichon@gmail.com

Uthai Tanlamai

Department of Accountancy, Chulalongkorn Business  
School, Chulalongkorn University  
Bangkok, Thailand  
e-mail: uthai@cbs.chula.ac.th

Kanokwan Atcharyachanvanich

Faculty of Information Technology,  
King Mongkut's Institute of Technology Ladkrabang  
Bangkok, Thailand  
e-mail: kanokwan@it.kmitl.ac.th

Nagul Cooharajanane

Department of Mathematics and Computer Science  
Faculty of Science, Chulalongkorn University  
Bangkok, Thailand  
e-mail: nagul.c@chula.ac.th

Achara Chandrachai

Department of Commerce, Chulalongkorn Business School,  
Chulalongkorn University  
Bangkok, Thailand  
e-mail: achandrachai@gmail.com

*Abstract—Open Government Data (OGD) project can help citizen to better understand government administrations. In the open data initiatives. Thailand is now in the process of providing a draft OGD country roadmap. This paper, we develop OGD assessment application based on our developed assessment model and test the application acceptance. The developed assessment model consists of nine components which include the enterprise architecture as a new the assessment component. The enterprise architecture is now becoming the key activity to transform the digital government. Therefore, to enhance the user experience in assessment, the web-based application is developed. Testing the application acceptance by using Technology Acceptance Model (TAM). The data were collected from 30 public organizations. The result shows that most public organizations intend to use the application proposed model*

*Keywords-component; Open Government Data, Evaluate Open Government Data, Open Data Assessment, Open Data, TAM*

## I. INTRODUCTION

The new paradigm of two-ways communications and social medias such as Twitter, Facebook and Line, stimulated governments all around the world to develop e-Government and digital government for enhancing citizen participation and collaboration in civil affairs, for example, public policy making and budget planning [1]. Both research on e-Government and Digital government can be separated into two parts, the first one focus on policy and social science research including the structure, directions, and citizen behaviors of the modern state in the society. The other, technical research focus on applications and systems to support their government activities

uses to governance. Various government organizations recently produce and collect a wide broad range of different data types and sources in order to conduct the digital government tasks [2][3]. The OGD project can be help citizens to better understand government administrations such as how well it manages budgeting expense. In addition, the OGD implementation fosters the valuable use of citizens in government data. The concept of releasing government data that need to be easy to access, completeness and re-usable to improve public services by all. Government need to seek feedback from the public in order to continuous OGD improvement [4][5]. Another important in measure the OGD progress was developed a measurement tool to assess the yield comparative data between public organizations that were subject to the Thailand's 3 Year Digital Government Master Plan 2016 – 2018 [6] and Official Information Act, B.E. 2540 in Thailand. The components of OGD evaluation in organization form were drawn directly from the concept of the OGD to determine whether ODG goals have been satisfied in each component [7].

However, there are various problems with OGD in Thailand. The lack of digital knowledge and skills lead to a low data quantity and quality, while the work process of most government agencies still involve much paper work with no guidelines and procedures for staff. There are many existing researches on OGD assessment models (OGDAM), but their research does not suit the context of Thailand. In addition, there is no OGD assessment application being used in Thailand. This research aims to develop an assessment model to evaluate OGD by developing a model to OGD assessment

application and testing the application acceptance. This paper is divided into five sections. This introduction to the OGD research forms the first section. In the second section, a review an existing background literature on assessment components of open government data to develop the OGDAM, summary of the Thailand's three years Digital Government Master Plan (2016 – 2018) and Technology Acceptance Model (TAM). The third section describes the methodology to develop and test user acceptance by using TAM. In the fourth section, the results of TAM are presented. Finally, the last section is a discussion and concludes of the developed OGDAM and TAM results with a research summary and proposed future research.

## II. LITERATURE REVIEW

### A. *a review an existing on assessment components of open government data*

1. Policies and Plans describes high level activities of minister and his cabinet. This group will determine government responsibilities in order to identify directions of the open government data [7]. OGD policy making should identify OGD goal both long term and short term, how to then meet open governments data with enable smart procedures and standards [8]. However, many countries tried to set up OGD strategies which emphasized on rapidly release open government datasets. As well as missed public interest, value creation and innovation which is the output of OGD [9][10].

2. Laws and regulations can be both advantage and disadvantage. For instance, legal framework improvement was critical to facilitate government data accessibility and re-use, data sharing and securities [10]. Legal perspective was significant how OGD can be complied with legislation, including freedom of information acts, OGD policies, open government directives and statements [11]. However, certain legal requirements such as outdated laws will make diversification for OGD initiative in order to promote and collaborate the work within government agencies for transition towards digital government.

3. Organization Management builds organizational capabilities in this area need an ability to properly manage open government data programs with other agencies in different areas, for example, training, management activity, and performance assessment. Also, it is important to be able to measure open government maturity and to ensure appropriate internal and external coordination [12]. These are strategic elements which need to be understood and addressed at national level and in organizational strategies and plans [12]. Government agencies need to further increase the participation in the agencies and considering to help change the culture and attitudes of public officials in order to make data available and sharing with peers [12].

4. Personal Capabilities of an entity to appropriately perform organization, training and management activities as planned. All staffs and management should be trained suitable skills in e-government and open data principles [13]. Additional, resources development and measurement were a

strong tendency for agencies to stretch entities to the success of their efforts [14].

5. Technology Infrastructure is an infrastructure which is the basis of government agencies to release their data [15]. OGD platforms would be a central platform for enabling public organizations and citizens to use officially updated data [16]. Any organization with high quality of integrated ICT-systems such as network, cloud and portal have a higher opportunity to OGD achievement. Government shall be provided OGD platform such as C1-RES of the ENGAGE project which help to visualize core data sets in a user driven fashion [16].

6. Open government data principles; the popular guidelines of OGD Principles with eight principles was created by [4]:

6.1 Complete, all public data is made available excluded privacy, security or privilege data.

6.2 Primary, raw data releases at the source, not modified.

6.3 Timely, data releases as soon as possible for the high value data.

6.4 Accessible, data releases with access free and no limitation for anyone.

6.5 Machine Readable, data releases with a structured manner for automated processing.

6.6 Non-Discriminatory, data releases without any registration for all.

6.7 Non-Proprietary, data releases which is not own by a single entity.

6.8 License, data release without any violate such as copyright, patent or trade secret regulations.

7. Innovation is the output of OGD implementation was innovation to produce both commercial and social benefits [17]. There is various type to measure innovation perspective such as data access measurement which indicated access indicators or degree of involvement in the encouragement and assistance to the work of re-user' agents [18]. OGD can help create service and product innovations especially, information technology and communication industries can increasingly use open data to develop novel applications or services and produce both commercial and social benefits [19].

8. Participation related to the increasing number of published dataset would result in more participation and innovation [19]. The main potential of OGD initiative was citizen participation and collaboration for creating value and innovation in government services [20]. Participation refers to the extent to which citizens can participate in the government services of the OGD such as using datasets or access to government data [20].

As the research gap identified that there is no research identifying enterprise architecture as the assessment perspective, which becomes the key activity to transform the previous government work process to a government digitalized process. Driving OGD needs to change the structure of an organization ) enterprise architecture( and to determine the roles that different units play in the OGD program.

9. Enterprise architectures are significantly information technology evolution for expanding new systems that enhance organization goals. This can be executed in both technical and business term. Technical term included web-based application, hardware and work process activities change from the paper work environment to the e-government environment in government agencies. In the business terms such as mission, business functions, and policy environments. Enterprise architecture is a management tool and the commitment of an e-government agency depends on the familiarity of the executive officers with this concept [20].

*B. The Thailand's three years Digital Government Master Plan (2016 – 2018)*

To meet the goal of Thailand's three years Digital Government Master Plan (2016 – 2018) (in next 3 years, the plan contains with four critical elements namely; government integration, smart operations, driven transformation and citizen-centric services. From all elements, the Thai Government will be lifted up to a Digital Government along the plan [6].

1. Government Integration: Thai government agencies will integrate information and operation between different agencies to arrange government share services at a single point.
2. Smart Operations: government will provide digital technology to support digital personnel's work activities such as government cloud system, government mail and electronic transactions. The connection between each agency will be proceeded within three years. The important consideration was big data and analytic tools.
3. Driven Transformation: government agencies need to transforms an organization in aspects such as human resources, work processes, technology and regulations. Central government and the cabinet will fully support with policy, vision and directive the importance of development in utilizing technology.
4. Citizen-centric Services: the service sectors will stimulate public participation so that there is responding the citizen requirements based on an individual's needs.

*C. An applicable of Technology Acceptance Model (TAM)*

TAM designed for modeling user with acceptance of information technology [21] and adjusted a theoretical of the reasoned action (TRA) by [22]. The TAM model consists of behavioural intention to use, which is measured users' attitudes of the model and the perceived usefulness of the model. Perceived ease of use is one of TAM model which effect from attitudes and perceived usefulness, considering a general person's salient acceptance in term of using technology. It is useful to develop performance of each person. "Perceived ease of use is a person's salient belief that using the technology will be free of effort" [23]. Moreover, [24] observed that TAM

excludes factors that might be significant predictors of information technology and information system usage. For example, TPB provides constructs that do not exist in TAM theory. As a result, [24] created perceived resources for TAM model. Perceived resource is to extent that a person trust which both personal and company resources needed to employ an information system like individual ability, computer hardware and web-based application documentation, data, time and man power.

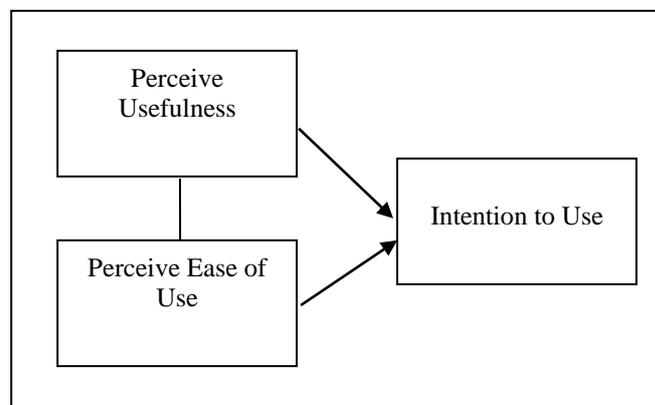


Figure 1. Sutable Technology Acceptance Model

*D. OGD assessment web-based application*

Once the nine components were investigated that they can be used for assessing the progress of OGD in Thailand. Therefore, to enhance the user experience in assessment, the web-based application is developed. The web-based application is implemented based on Visual Studio 2015, SQL Server 2012, jQuery, HTML, Java script, and Chart.js which developed to Web Application and set up on Server Hosting (see Fig. 2).

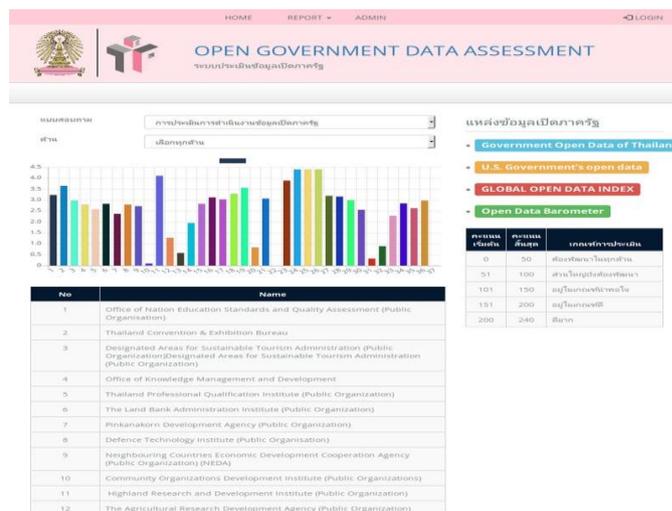


Figure 2. OGD Home page

This web-based application allows all public organization in Thailand to assess their open government data progress. The application contains three processes; input, process and output. (1) For input, each public organization shall register both three assessors especially who work in the fields of policy, information technology and operation, and approver who is the general director. The assessors have to assess their activities all nine components. The approver has an authorize to review and approve the result of the organization. (2) For process; this web-based application feature can calculate the result and shows the result in both total score and each component score. (3) For output, the application can compare the result with other public organization. Moreover, it can be generated an executive summary for the organization.

The web-based application can collect data, calculate and report the assessment results. The process to assess and approve can be show in step below:

Step #1. The user logs in and assess these nine components with other department in the same organization and review information (in Thai language) before submitting to approver (see Fig. 3).

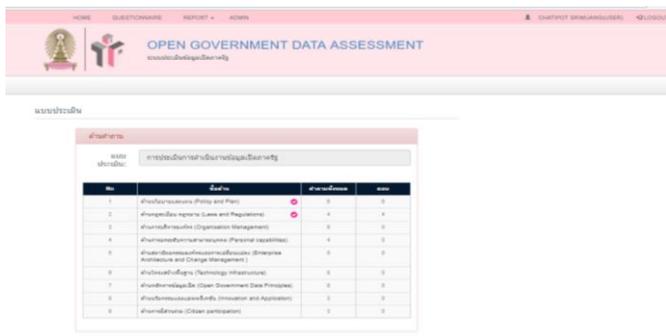


Figure 3. Assessment process in nine components

Step #2. The result will be reviewed before send to approver who is generally the general director of organization to approve the result. However, the general director has authorized to review and send back to user to revise the result (see Fig. 4).



Figure 4. Reviewed process

Step #3. When the result has been sent to approver, the approver will review or send feedback to assessor. After reviewing and accepting the result, an approver can approve by click the green button. The approval processes can be shown in Fig. 5.

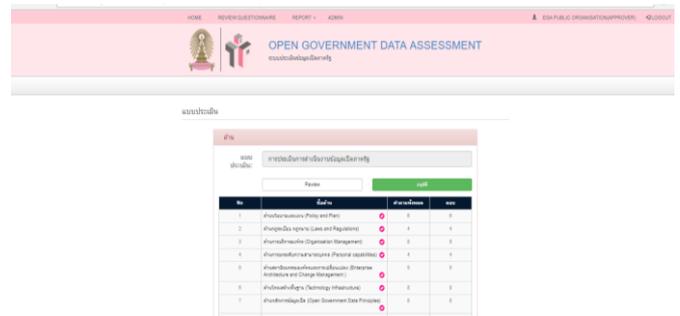


Figure 5. Reviewed process

Step #4. After approval, the result will be show in the first page in the website (see Fig. 2).

Step #5. The result can be show in all thirty-nine-public organization in Thailand and can compare the summary result between public organization (see Fig. 6).

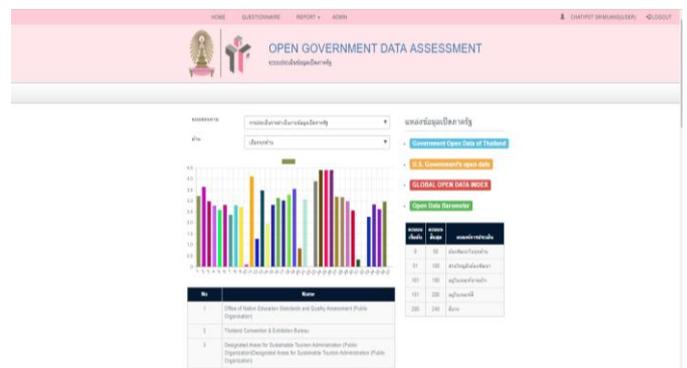


Figure 6. Suitable Technology Acceptance

Step #6. Web-based application can show the progress of organization in each dimension (Fig 7).

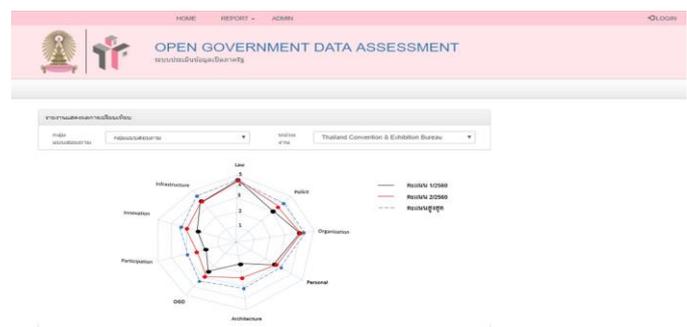


Figure 7. Score Comparison in each Component

## III. METHODOLOGY

Before In order to come up with an assessment model and the assessment application, this research consists of three stages;

*Proposed Open Government Data Assessment Model*

*Stage 1:* Reviews and determine the open government data assessment components will be suitable to the context of Thailand.

*Stage 2:* Develop an OGD assessment application using Visual Studio 2015, SQL Server 2012, jQuery, HTML, Java script and Chart were to be a web application installed in Server Hosting named <http://www.th-ogdam.com>.

*Stage 3:* Survey questionnaires created from reviewing an applicable TAM and data collected from 30 public organizations referred to Public Organizations Act B.E. 2542 in Thailand. The sample focused on IT manager as assessment users who are authorized to use the assessment model. Questionnaire in Thai language is consisted of 4 parts: 1) Demographic data of the respondent; 2) Basic data of the public organization; 3) Five points Likert-scale rating of TAM including perceive usefulness, perceive ease of use and intention use; and 4) Open-ended any suggestions. The data collection occurred during March 1, 2017 to June 1, 2017.

## IV. FINDING

This research used statistical analysis to test the acceptance model of OGD web-based application in the assessment of the OGD implementation in organizations. The research analyzed the multi-item scales, as shown in the table 1. TAM analysis will be demonstrated for further validation. The total number of respondents is 30 persons; 21 people 70.0% are males and 9 people 30.0% are females. The highest proportion in terms of age of the respondents are the age between 26-35 years old same as age 36-45 years old, which is at 50% or 15 persons in both group. Furthermore, most respondents 15 persons 50.0%, have a master's degree. For statistic result can be shown in Table 1.

TABLE I. TAM ANALYSIS

Component	OGD web-based application			
	Questions	n	Mean	Std.
Perceive Usefulness	Useful	30	4.17	0.79
	Online assessment	30	3.67	0.76
	Work quicker	30	3.63	0.64
	Increase accuracy	30	4.03	0.77
	Efficiency system	30	3.80	0.76
	Management easy to review	30	3.97	0.67
Perceive Ease of Use	Easy to learn	30	3.70	0.62
	Easy to use	30	4.03	0.65
	Flexible to assess	30	3.90	0.89

Component	OGD web-based application			
	Questions	n	Mean	Std.
	Clarity	30	3.97	0.67
	Decrease assessment process	30	3.87	0.73
	More understandable	30	3.87	0.63
	Accurate result	30	4.00	0.74
Intention to use	More Quality	30	3.97	0.77
	More confident in result	30	4.03	0.83
	Prefer to use	30	4.50	0.63

The research aims to study TAM of OGD web-based application which assess progress of OGD implementation in Thailand's public organization. The statistic result demonstrated into three groups;

Perceive Usefulness, participants responded the highest in OGD web-based application is Useful with mean equal 4.17 followed by Increase accuracy with mean equal 4.03. However, the lowest is Work quicker with mean equal 3.63.

Perceive Ease of Use, participants responded the highest in OGD web-based application is Easy to use with mean equal 4.03 followed by Clarity with mean equal 3.97. However, the lowest is Easy to learn with mean equal 3.70.

Intention to Use, Participants prefer use OGD web-based application is the highest with mean 4.50 and the lowest is more quality with mean 3.97.

Researchers also asked the last question that would participants like to use OGD web-based application and 28 persons or 93.3% prefer to use this web-based application.

## V. CONCLUSION

The concept of open government data is to enable citizens and allows them to have access to public data for free and without restrictions. To assess the implementation of open government data in Thailand, this research develops a new workable OGD web-based application. The research result shows that the participants intent to use OGD web-based application because accurate and more confident result. Moreover, respondents believe that OGD web-based application will enhance their work such as increase accuracy, efficient system and clarity.

This research can be concluded that the research contributes useful information to government and especially the public organizations in Thailand on develop OGD web-based application that assess the OGD implementation progress.

In particular, it has been concluded that OGD web-based application has key features to implement OGD capabilities for public organizations in Thailand and web-based application allows a government agency to assess its digital government stage from operational level and trustable data.

Future research is required to validate the reliability of OGD web-based application by sending a questionnaire with the five-point Likert scale to more participant especially public organizations in Thailand. This will therefore, provide

guidelines and the best practice for public organizations in Thailand to implement.

REFERENCES

- [1] C. M. Chan, (2013). From open data to open innovation strategies: Creating e-services using open government data. 2014 47th Hawaii International Conference on System Sciences 0 (pp. 1890–1899).
- [2] B. Ubaldi, “Open Government Data TOWARDS EMPIRICAL ANALYSIS OF OPEN GOVERNMENT DATA INITIATIVES. doi: 10.1787/5k46bj4f03s7-en (2013).
- [3] C. Alexopoulos, E. Loukis, S. Mouzakitis, M. Petychakis and Y. Charalabidis., “Analysing the Characteristics of Open Government Data Sources in Greece” ©Springer Science+Business Media New York 2015
- [4] M. Solar, G. Concha and L. Meijueiro, “A Model to Assess Open Government Data in Public Agencies” ©IFIP International Federation for Information Processing: EGOV, 2012 LNCS 7443, pp. 210–221
- [5] R. P. Lourenço and Serra L., “An Online Transparency for Accountability Maturity Model”. IFIP International Federation for Information Processing, EGOV 2014, LNCS 8653, pp. 35–46.
- [6] Electronic Government Agency (Public Organization) [Online]. “Thailand’s three year Digital Government Plan (2016-2018), 2016. [https://www.ega.or.th/upload/download/file\\_430e6925f329591155fb3801cdd08d60.pdf](https://www.ega.or.th/upload/download/file_430e6925f329591155fb3801cdd08d60.pdf) [Access Date: 10 September, 2017]
- [7] R. Almazan and J. Garcia., “Towards an Evaluation Model for Open Government: A Preliminary Proposal” IFIP International Federation for Information Processing: EGOV 2014, LNCS 8653, pp. 47–58, 2014
- [8] H. J. Wang and J. Lo, “Adoption of open government data among government agencies”. *Government Information Quarterly*, 33(1), 80–88. doi: 10.1016/j.giq.2015.11.004
- [9] G. Lee and Y. Kwak, “An Open Government Maturity Model for social media-based public engagement” *Government Information Quarterly*. 29, 492–503 (2012).
- [10] J. Huntgeburth and D. Veit, “Evaluating Open Government Initiatives” © Springer International Publishing Switzerland 2015
- [11] A. Zuidervijk, N. Helbig, J. R. Gil-Garcia, and M. Janssen, “Special Issue on Innovation through Open Data - A Review of the State-of-the-Art and an Emerging Research Agenda” Guest Editors’ Introduction. *Journal of theoretical and applied electronic commerce research*, 9(2), 1-2. doi: 10.4067/s0718-18762014000200001 (2014).
- [12] R. Fuentes-Enriquez, and Y. Rojas-Romero, (2013) . Developing accountability, transparency and government efficiency through mobile apps: The case of Mexico. Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance (pp. 313–316) . New York, NY, USA: ICEGOV '13, ACM. <http://www.dx.doi.org/10.1145/2591888.2591944>. [Access Date: 20 September, 2017]
- [13] E. Loukis, Y. Charalabidis and A. Androutsopoulou, “Evaluating a Passive Social Media Citizensourcing Innovation”. 9248, 305-320. doi: 10.1007/978-3-319-22479-4\_23, 2015.
- [14] N. Veljković, S. Bogdanović-Dinić and L. Stoimenov, “Benchmarking open government: An open data perspective” *Government Information Quarterly* © 2014 Elsevier Inc. All rights reserved
- [15] E. Styrin, N. Dmitrieva and A. Zhulin, “Openness Evaluation Framework for Public Agencies”. ICEGOV'13, October 22–25, 2013, Seoul, South Korea.
- [16] E. Barry and F. Bannister, Barriersto open data release: Aview from top, presented at the European Group for Public Administration, Edinburgh, September, 2013
- [17] P. Parycek, J. Höchtel and M. Ginner, “Open Government Data Implementation Evaluation”. *Journal of theoretical and applied electronic commerce research*, 9(2), 13-14. doi: 10.4067/s0718-18762014000200007, 2014
- [18] E. Kalampokis, E. Tambouris and K. Tarabanis, “Open Government Data: A Stage Model” ©IFIP International Federation for Information Processing: EGOV 2011, LNCS 6846, pp. 235–246, 2011
- [19] J. Bertot, P. McDermott and T. Smith, “Measurement of Open Government: Metrics and Process” IEEE 45th Hawaii International Conference on System Sciences, 2491-2499. doi: 10.1109/hicss.2012.658, 2012
- [20] D. Sayogo, T. Pardo and M. Cook, “A Framework for Benchmarking Open Government Data Efforts” IEEE 47th Hawaii International Conference on System Science 1896-1905. doi: 10.1109/hicss.2014.240
- [21] M. Fishbein, and I. Ajzen, (1975), *Belief, intention and behavior: an introduction to theory and research*. Addison Wesley; 1975.
- [22] F.D. Davis, R.P. Bagozzi, and P.R. Warshaw (1989), “User acceptance of computer technology: a comparison of two theoretical models”, *Management Science*, Vol. 35 No.8, pp. 982-1003.
- [23] D. Robey, (1979). User attitude and Management Information System Use, “Academic of Management Journal. September 1979.
- [24] K. Mathieson, E. Peacock, and W. Chin, (2001), “Extending the Technology Acceptance Model: The Influence of Perceived User Resources,” *Database For Advances in Information Systems*, Vol. 32, Issue 3, pp. 86 – 112.

## **Session 8:        Wireless Networking and Communication**

Title: Propagation modelling and performance assessment of aerial platforms deployed during emergencies

(Authors: Faris A. Almalki, Marios C. Angelides)

Title: A Comparative Analysis of MANET Routing Protocols through Simulation

(Authors: Thomas Nash, Callum Brill, Aspen Olmsted)

Title: Multi-Channel Steganographic Protocol for Secure SMS Mobile Banking

(Authors: Omega Obinna, Eckhard Pfluegel, Charles A. Clarke, Martin J. Tunnicliffe)

# Propagation modelling and performance assessment of aerial platforms deployed during emergencies

Faris A. Almalki and Marios C. Angelides  
 Department of Electronic and Computer Engineering  
 College of Engineering Design and Physical Sciences  
 Brunel University London  
 Uxbridge, United Kingdom  
 {faris.almalki, marios.angelides}@brunel.ac.uk

**Abstract**—This paper proposes a channel model for aerial platforms deployed for disaster relief that aims to optimize the underlying ad hoc network performance. Channel modelling has been motivated by existing empirical and ATG propagation models and their resilience and scalability for aerial platforms. Grade of Service (GoS) is the benchmark used to set the desired performance of a trunked system. Thus, two GoS concepts are being considered: probabilities of blocking (Erlang B), and delay (Erlang C). The Received Signal Strength (RSS) performance of rescue teams is then simulated. Our prediction results show that aerial platforms offer much promise in providing robust communication links during emergency relief operations.

**Keywords**— *Low Altitude Platforms; High Altitude Platforms; Emergency Situations; Propagation Models; QoS*

## I. INTRODUCTION

In man-made or natural disaster scenarios, wireless communication systems have the further merit that they are less vulnerable to physical damage. When disaster strikes, terrestrial communication links are often disrupted, yet for disaster relief workers, such links are essential during rescue operations. Under such conditions, it is essential to have a large-scale and robust communication network for broadcasting instructions to people, seeking rescue aid, as well as providing inter-departmental communications. Disruptions caused by physical damage can be incredibly costly and time consuming to restore, as they require maintenance or sometimes replacement of complex network hardware to re-establish communications. This is especially problematic if major installations such as terrestrial towers or fiber-optic cables are involved [1, 2].

Outer space communication systems, such as satellites, or aerial High Altitude Platforms (HAPs) and Low Altitude Platforms (LAPs) might serve as better solutions. Authors in [1, 2, 5] offer several examples of space communication systems that had successfully established ad-hoc networks for disaster relief. This is mostly based on outer space communication systems covering a wide area, offering deployment flexibility, forecasting disaster evolution, providing last-mile connectivity, re-configurability, or in case of emergency and disaster situations offering unique Line of Sight (LoS) advantage.

However, satellite systems are not without limitations, for instance, propagation delay, handover complexity, high power consumption and cost. All these in addition to environmental issues such as high gas emissions during satellite launches, as well as signals that have no regard for geographical or political boundaries, which might or might not be a desirable feature [1, 4].

Aerial platforms are of flexible size since they are usually helium-filled and solar powered airships and can be used for various applications and services such as telecommunications, broadcasting, surveillance, emergency services, and navigation. Its position in the sky could take advantage of the strengths of terrestrial and satellite communication systems, whilst avoiding some of their weaknesses. Deploying an Ad Hoc network using aerial platforms for disaster relief swiftly bridges communication gaps through a soft infrastructure, a fast start-up time, gradual growing, on-demand capacity assignment, with low capital investment, as well as low ongoing operating costs [2, 4].

The rest of this paper is organized as follows: Section II presents related works on LAP and HAP technologies from an emergency response perspective. Section III presents our channel modelling, including Trunking and QoS. Analysis of the resulting performance is presented in section IV. Finally, section V paper concludes.

## II. RELATED WORK

Aerial platforms are increasingly seen as an innovative solution to the last-mile problem. They could provide many of the satellite advantages, but without the distance penalty. Receivers may experience a better signal quality, as the system offers LoS communications, hence, less propagation delay in relation to satellite systems. Our review of relevant literature reveals that there has been some consideration of the capabilities of aerial platforms for emergency situations.

A trail of Emergency Broadband Access Network (EBAN) is designed in [5] to provide wide area hotspots for emergency relief in Indonesia using a tethered-blimp balloon. The evaluation is based on measured Received Signal Level (RSL) and Signal-to-Noise Ratio (SNR). In this trial, Wireless Fidelity (WiFi) and Worldwide Interoperability for Microwave Access (WiMAX) services are considered. At an altitude of 0.4km above ground, the balloon's coverage area is 47.39km<sup>2</sup> with a constant 54Mb/s downlink throughput, but as coverage increases to 72km<sup>2</sup> throughput fluctuates.

The WiFi gives a satisfactory performance for Internet access and easily achieves a LoS for rural users but with challenging consequences for urban users. WiMAX offers better capacity, less interference, and has better coverage with no LoS (nLoS) where objects block signals. Recommendations for future work includes an investigation of the impact of emergency traffic on the various communication interfaces at different altitudes [6].

Researchers in [7] investigate the performance of 4G Long-Term Evolution (LTE), and WiFi multimode base stations installed on aerial stations to deliver coverage for first responders in emergency situations. Directional antennas are

utilized in the aerial platform to provide either macrocell, microcell or picocell coverage. Their results show that the performance of different link segments, whether LTE or WiFi, at varying aerial-station altitudes between 0.5km and 2km is quite high. However, packet delay increases as the number of parallel services rises. Recommendations for future work includes an investigation of the impact of emergency traffic on the various communication interfaces for different altitudes.

Authors in [8] investigate the challenges of WiMAX, WiFi, LTE, ZigBee, and XBee wireless technologies for enabling aerial drone platform in Alpine environments, in order to support short term winter events and provide a novel but viable solution in emergency and rescue situations in a hostile environment. The candidacy of an omni-directional antenna type is considered in this case and WiMAX has been suggested as a suitable wireless technology for drone communications for several reasons: flexibility, safety, QoS, interference, throughput, installation, and coverage area.

In [9] an optimal LAP location is examined for emergency conditions to minimize path loss and maximize RSS. An ATG propagation model is considered at 1km altitude, 2GHz frequency band and transmission power set at 40 dBm. Their results show that the maximum urban radius achieved is 20km with path loss and RSS of 130dB and 82dBm respectively.

Authors in [10] highlight the behavior of an LTE network in a disaster scenario using a HAP. The propagation model considered is ATG at 17km altitude, transmission power at 30dBm, and multibeam antenna gain at 38.7dBi. The simulation results show that the RSS floats between -70dBm and -110dBm, and the system can restore 92% of throughput. Nevertheless, further HAP configurations and link designs are required to enhance QoS results.

In [11] the authors study the coverage of a LAP system equipped with WiFi access points to provide terrestrial users with wireless communication services for short-term events and/or emergency situations. Their results show that at 0.5km altitude, 2.4GHz frequency band and transmission power set at 35dBm, the maximum urban radius achieved is 6km, with path loss of 120.5dB and RSS of -80dBm.

Authors in [12] present an algorithm that calculates the optimal placement to cover an area using different station types, i.e. portable terrestrial stations (PTSs), and LAPs. Both assume being equipped with LTE technology for random situations. The authors present a case from the disaster caused by Hurricane Katrina. Their results confirm the advantage of deploying LAPs in terms of high bandwidth utilization, wide coverage, and required number of base stations to cover a specified area in relation PTSs. However, increasing the number of LAPs may cause interference with terrestrial stations. Therefore, it is recommended to positioning LAPs on the boundaries of the disaster area, so that interference is reduced whilst leaving no coverage gaps.

Authors in [13] investigate the performance of LTE and Wi-Fi technologies in an urban Australian emergency scenario using a tethered LAP. A ray tracing ATG path loss model and three empirical propagation models are simulated at many LAP altitudes, and four network performance indicators are measured, i.e. path loss, outage probability, delay, and throughput. Their results show that LTE outperforms WiFi under all conditions, while it is inferred that cost, coverage, and deployment time should be

considered for suitable selection of technology for LAPs. One important enhancements offered by LAPs is their ability to increase the footprint area compared to terrestrial networks due to an increased LoS probability. However, this depends on LAP altitude, frequency band, and antenna specifications.

Considering multiple antenna technology improves performance as indicated by Alamouti's scheme of Multiple Input Multiple Output (MIMO) antenna technology. The scheme maximises capacity and improves QoS and coverage extension range. The effect of MIMO antennas on near space solar powered platforms performance and capacity is discussed in [14, 15], where it is argued that the antenna gain is optimized, to prevent users from experiencing weak radio across many miles.

The ITU's International Mobile Telecommunications-Advanced (IMT-Advanced) standard for 4G offers access to various telecommunication services and supports mobile applications for various environments that offer high data rates to users. In aerial platform technology, both WiMAX [9] and LTE [10] perform well and are suggested as good candidates for better coverage, whether in LoS or nLoS, increased capacity and less interference [16].

### III. PROPOSED MODEL

The authors in [5, 7] propose investigating the impact of emergency traffic on aerial platforms. However, deployment of aerial platforms for short-term emergency communications is scarcely reported in the literature and where it is the constituent models do not utilise the full range of link budget parameters let alone taking the view of GoS. Therefore, in proposing a model for emergency communications that uses LAP and HAP technologies the aim is to include as wide a range as possible of link budget parameters and also consider trunking and GoS performance.

A special attention is paid to assessing the performance of rescue teams for disaster relief operations through analyzing the performance of RSS predictions in an urban scenario. The model has sourced calculations from the emergency operations in Florida state in USA when struck by Hurricane Irma in August 2017 when wired and wireless communication links between residents and rescue teams were destroyed. The aerial platform system requires two main segments, a space and a ground segment as pictured on Fig.1 [4].

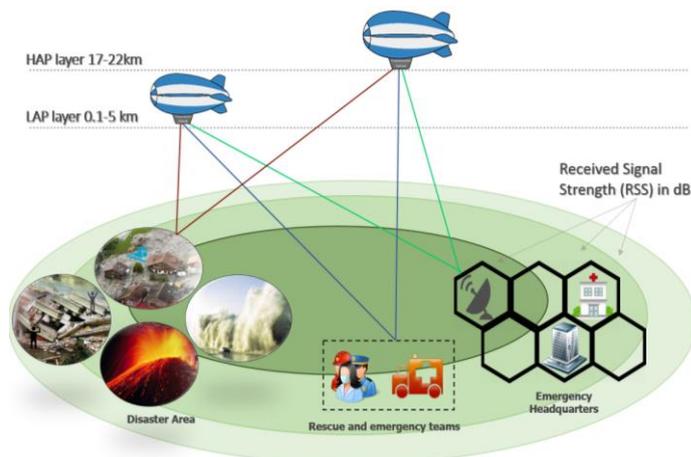


Figure 1: The aerial platform system structure for emergency conditions

The resulting architecture consists of various layers which can provide wireless services for rescue operations and act as an ad hoc node for emergency communications. Each of the layers has different hardware and software capabilities. The ground stations are linked by platforms using both backhaul links and also host gateways to external networks, and intermediate wireless sensor nodes in the ground. RSS is affected by distance and/or shadowing as signals experience reflection from obstacles that interfere with them.

### A. Link budget parameters

The first aspect to consider in the proposed model are the channel characteristics which are modelled using two types of propagation models: An empirical Okumura propagation model for LAP and an ATG propagation model for HAP [17-19]. Okumura is best at predicting path loss in dense areas with tall buildings using a range of frequencies and achieving altitudes up to 1km. Its path loss  $P_L$  calculation is as follows:

$$P_L = L_f + Amn(f, d) - G(h_t) - G(h_r) - G_{area} \quad (1)$$

$$L_f = 32.44 + 20 \log(f) + 20 \log(d) \quad (2)$$

$$G(h_t) = 20 \log(h_t/200), 10m < h_t < 1000m \quad (3)$$

$$G(h_r) = 10 \log(h_r/3), h_r \leq 3m \quad (4)$$

$$G(h_r) = 20 \log(h_r/3), 10 > h_r > 3m \quad (5)$$

where  $L_f$  is Free space path loss (dB),  $Amn(f, d)$  is Median attenuation relative to free space (dB),  $d$  is Transmission Distance (m),  $f$  is Carrier frequency (GHz),  $G(h_t)$  is Transmitter antenna height gain factor (dB),  $G(h_r)$  is Receiver antenna height gain factor (dB),  $h_t$  is Transmitter antenna height (m),  $h_r$  is Receiver antenna height (m),  $G_{area}$  is Gain due to environment (dB). Correction factors relate to type of terrain. The values of  $Amn$  and  $G_{area}$  can be worked out by empirical plots [19].

ATG depends on the elevation angle and uses path loss, and shadowing, whether LoS or nLoS. Shadowing and reflection gives rise to nLoS. Fig. 2 shows selected ITU-R empirical parameters as a, b, c, d, and e for different environments simulated with a LoS probability using a wide range of elevation angles [18].

Environment	a	b	c	d	e
Urban	187.3	0	0	82.10	1.478
Suburban	120	0	0	24.30	1.229
Rural	101.6	0	0	3.25	1.241

Figure 2: Selected ITU-R parameters for different environments

Calculation of the total path loss  $P_L$  is as follows:

$$PL_T = \rho_{LoS} \times PL_{LoS} + \rho_{NLoS} \times PL_{NLoS} \quad (6)$$

The probability of having LoS connections at an elevation angle of  $\theta$ , depending on the type of environment is given by:

$$\rho_{LoS} = a - \frac{a-b}{1 + \left[\frac{\theta-c}{d}\right]^e} \quad (7)$$

$$\rho_{NLoS} = 1 - \rho_{LoS} \quad (8)$$

The path losses for LoS and NLoS are as follows:

$$PL_{LoS} \text{ (dB)} = 20 \log \frac{4\pi(f)(d)}{c} + \eta_{LoS} \quad (9)$$

$$PL_{NLoS} \text{ (dB)} = 20 \log \frac{4\pi(f)(d)}{c} + \eta_{NLoS} \quad (10)$$

Where  $\eta_{LoS}, \eta_{NLoS}$ : average additional loss to free space depending on type of environment.

The rest of the link budget parameters that are considered for inclusion in the proposed model are: RSS, SINR, throughput, and coverage radius. RSS depends on  $P_t, P_L, G(h_t), G(h_r)$  as well as connector and cable loss  $L$ . SNIR depends on RSS, Noise figure  $N$  and Interference  $I$ . As there exists no formula with which to calculate the exact throughput based on  $P_L$  and SNIR, a prediction is made using Shannon's formula. Predicted throughput  $C$  depends on bandwidth  $B$  and SNIR. Calculation of the coverage footprint uses the elevation angle from the user's position a significant departure from current empirical models [23-26]. The RSS, SNIR and throughput are predicted as follows:

$$RSS = P_t + G(h_t) + G(h_r) - P_L - L \quad (11)$$

$$SNIR = \frac{RSS}{N+I} \quad (12)$$

$$C = B \times \log(1 + SNIR) \quad (13)$$

### B. Trunking and Grade of Service

The proposed channel model is evaluated using the mathematical concept of Trunking. A trunked mobile radio system provides access to users on demand from an available number of channels. A small number of channels can accommodate a large random number of users due to limited radio spectrum. GoS aims to measure the ratio of users accessing a trunked system during the busiest hours. The busy hours in our case are during or immediately after natural or man-made disasters. Thus, it is vital to consider the two main concepts of GoS: the probabilities that a call is either blocked "Erlang B", or experiencing a delay greater than a certain queuing time "Erlang C" [20]. The probability of being blocking  $P_B$  is expressed as follows:

$$(P_B) = \frac{A^C/C!}{\sum_{i=0}^C A^i/i!} \quad (14)$$

$$A_u = \lambda_T \times H \quad (15)$$

$$A = U \times A_u \quad (16)$$

$$A_c = A \times R_n \quad (17)$$

$$U_{Ser} = C \times R_n \quad (18)$$

$$U_{Sub} = \frac{A}{A_u} \times R_n \quad (19)$$

Where  $P_B$  is the ratio between the number of lost calls and the total number of calls,  $C$  is the number of channels,  $A$  is total traffic,  $H$  is the call duration, and  $\lambda_T$  is arrived calls rate,  $U$  is number of users, and  $A_u$  is the call rate per user.  $A_c$  is carried traffic,  $R_n$  is number of cells.  $U_{Ser}$  is the maximum number of users who can be served at a specific time, whereas  $U_{Sub}$  is the percentage of users who subscribe to the LAP/HAP systems. The probability of delay  $P_d$  formula is expressed as follows:

$$(P_d > 0) = \frac{A^C}{A^C + C!(1 - \frac{A}{C}) \sum_{i=0}^{C-1} A^i/i!} \quad (20)$$

$$(P_d > t) = (P_d > 0) \times \exp\left(-\frac{C-A}{t}\right) / H \quad (21)$$

$$D = (P_d > 0) \times \frac{H}{C-A} \quad (22)$$

Where  $P_d > 0$  is the likelihood that a call is initially denied access to a channel in the system,  $P_d > t$  is the probability that the delayed call is forced to wait more than  $t$  seconds and  $D$  is the average delay of all calls either delayed or not.

IV. RESULTS AND DISCUSSION

Reporting in the literature on testing WiMAX for LAP or HAP is scarce despite offering the advantage of independent links with minimum interference, especially for short term event and/or in emergency situations. Thus, our simulation will serve predictions of a fuller range of link budget parameters using WiMAX MIMO antenna specifications. Specification of simulation parameters of MIMO antennas has been provided by Netronics mobile WiMAX Telecom Company. The frequency band is 2.5GHz, on the Transmitter side, Power = 37dBm, Antenna Gain = 17dBi, Diversity gain = 5dBi, Rx Sensitivity = -90dBm, and loss = 5dB, on the Receiver side, Power = 27dBm, Antenna Gain = 2dBi, Diversity gain = 2dBi, Rx Sensitivity = -88dBm, loss = 0.5dB. The LAP and HAP altitudes, which denote transmitter antenna height  $h_t$ , are set at 1km, and 20Km, respectively. The receiver antenna height  $h_r$  is set at 1.5m. The total system Bandwidth B is 10MHz, and the full duplex channel bandwidth is 10.94kHz.  $\eta_{LoS}$  is an average of 4dB, whereas  $\eta_{NLoS}$  is 10dB. The minimum elevation angle from an aerial platform perspective is assumed to be 10 degrees.

A simulation of equations (1) to (22) is carried out using MATLAB, where link budget and GoS parameters sourced from the recent disaster caused by Hurricane Irma in Florida. Another simulation tool that has been deployed is a 3D RF and propagation tool “Remcom Wireless InSite software”. Table 1 shows all the predicted numerical results produced by the simulations and Fig. 3 through to Fig. 10 visualise these.

TABLE 1: Numerical predictions for LAP and HAP

Parameter	LAP	HAP
<b>Link budget</b>		
PL (dB)	-142.39	-156.60
RSS (dBm)	-82.39	-96.60
Throughput (Mb/S)	1.28	0.86
Footprint Radius (km)	24	216
<b>GoS of 0.02</b>		
$P_B$	0.091	0.064
$A_c$ (Erlangs)	7,776	69,984
$U_{Sub}$	155,520	1,399,680
$U_{Ser}$	51%	10%
$P_d > 20$	6.1%	6.4%
D (Seconds)	0.81	0.81
<b>RSS predictions of Rescue Teams</b>		
Rescue Team 1	-48.70	-46.89
Rescue Team 2	-45.38	-43.66
Rescue Team 3	-50.49	-47.03
Rescue Team 4	-68.72	-50.91
Rescue Team 5	-70.23	-51.76

Fig.3 and Fig.4 visualize the link budget predicted results of the Okumura, and ATG models and their effect on LAP and HAP, respectively, based on an ad hoc scenario. PL predictions are used for monitoring system performance and coverage to achieve a certain level of reception. PL shows a gradual increase as aerial attitudes increase. Thus, due to the short distance, the PL for LAP is better in comparison to HAP by around 14dB. PL values for LAP and HAP are below the maximum allowable path loss (MAPL) value of 158dB. Keeping transmission power constant at different transmitter

altitudes yields varying levels of RSS. As RSS is linked to PL the Okumura model for LAP achieves better predicted result than the ATG model for HAP, because as distances increase PL increases, and RSS decreases. Throughput decreases with distance as well as higher PL. Thus, the Okumura model for LAP yields the best predicted result with 67% in comparison to HAP’s throughput. Network coverage is affected by transmitter and receiver antenna specifications, geomorphology, and minimum elevation angles. Hence, the HAP radius is by far better than that of LAP due to the big difference in altitude. The peak coverage of 1km altitude of LAP is 24km, whereas for HAP is 216km at 20km altitude. The hexagonal LAP cells are of a radius of 2km and a cluster size of 7 whereas the hexagonal HAP cells are of a radius of 6km and also a cluster size of 7. We assume that each user makes 1 call per hour with an average duration of 3 minutes.

Fig. 5 and Fig. 6 show the probabilities of blocking and delay respectively. Fig.5 shows the probability of blocking as functions of the number of channels and traffic intensity in Erlang in different channels including those dedicated for LAP and HAP.  $P_B$  means that a new call arriving is rejected because all servers (channels) are busy. This measures traffic congestion in the telephone network in cases of lost calls.  $P_B$  in the case of LAP is higher than that of HAP due to the smaller number of cells. The number of LAP and HAP cells are approximately 9360, and 84,240 respectively. LAP cells are distributed evenly over Orlando in Florida over a total area of 1497km<sup>2</sup> and population of around 300,000. HAP cells are distributed over 65% of Florida state over a total area of 121,213km<sup>2</sup> and a population of 13 million as Fig. 7 shows.

The maximum traffic carried,  $A_c$ , is higher with HAPs in comparison to LAPs by 160% due to the huge capacity (number of channels and cells). Furthermore, the total number of users,  $U_{Ser}$ , which can be served at 2% GOS is 155,520 for LAP and 1,399,680 for HAP. The percentage of users who can subscribe to the aerial platform service  $U_{Sub}$  in the covered area are 51%, for LAP and 10% for HAP. Fig.6 shows the probability of a call being delayed, as a function of the number of channels and traffic intensity in Erlang. The probability of delay in both LAP and HAP is quite similar with an average delay of 0.81 seconds.

The 3D RF tool has been used to understand the ad hoc structure for disaster relief, and measure the performance of five rescue teams, as Fig. 8 shows. In this figure, one single aerial platform, and five rescue teams have been considered in an ad hoc scenario of a typical urban city in US. This architecture aims to measure and demonstrate the propagation power and RSS predictions of five rescue teams for both LAP and HAP. In this scenario, a MIMO antenna specification is considered and an elevation angle of 60 degrees to obtain more LoS connectivity. Rescue teams can obtain more resilient communications and effective responses. Due to the high elevation angle and thus less coverage area, the aerial platform is placed at the edge to also act as a relay between inter-departmental, headquarters and rescue teams.

Fig. 9 and Fig. 10 show RSS predictions for LAP and HAP respectively. As PL increases with distance and/or geomorphology due to multipath, RSS decreases. Fig. 9 and Fig. 10 demonstrate the RSS predictions for rescue teams in and ad hoc network over an urban environment. Keeping the transmission power constant at different transmitter altitudes yields varying levels of RSS. LoS connectivity is the main reason why RSS prediction results float within the same levels, which give rescue teams reliable communication links.

Overall, RSS yields a reasonably good average across all five rescue teams. The best RSS float between -45 dBm and -70dBm due to their antenna gains. The ATG model for HAPs achieves better predictions in comparison to the Okumura model for LAPs. The reason for that is the increase in PL for LAPs due to shadowing from obstacles such as high buildings, while the high altitude of HAPs helps to prevent this. RSS decreases with distance and/or shadowing appear as blue dots.

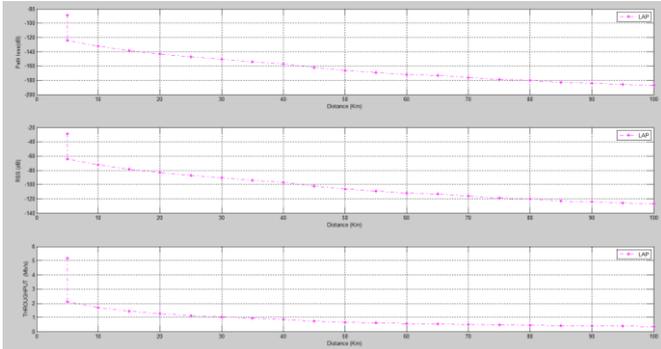


Figure 3: Link budget predictions – LAP

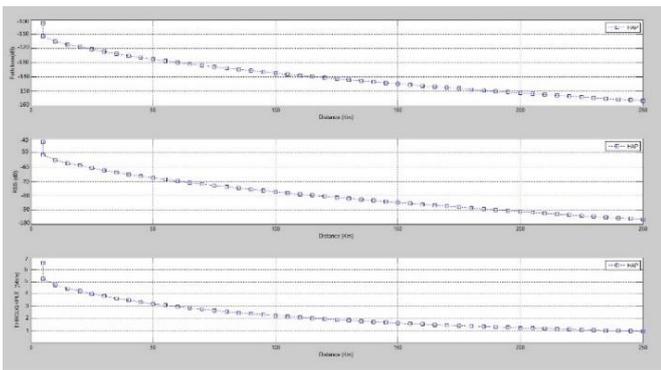


Figure 4: Link budget predictions – HAP

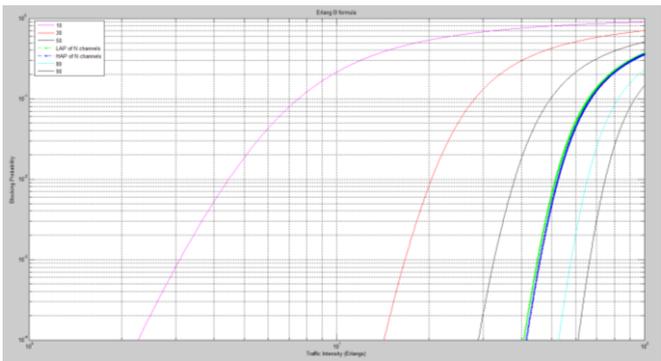


Figure 5: The Erlang B probability of blocking

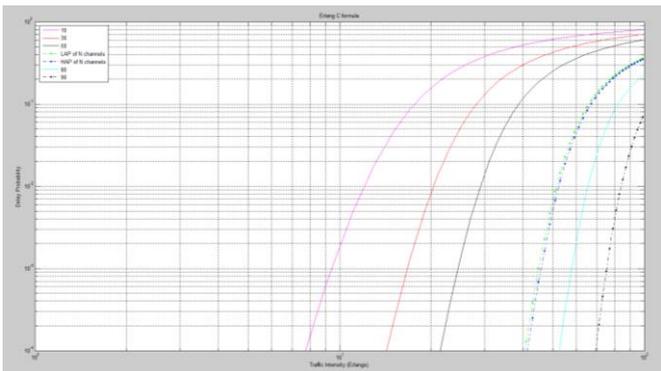


Figure 6: The Erlang C probability of delay

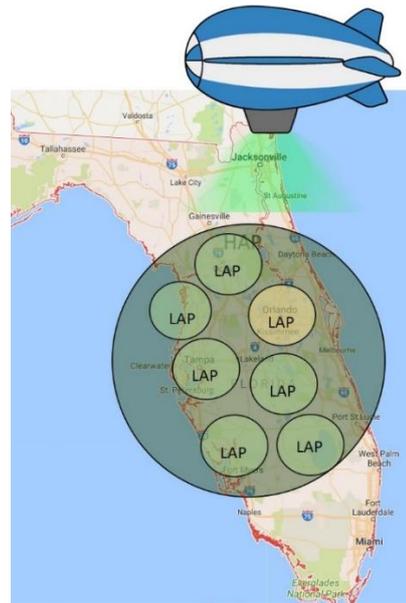


Figure 7: Coverage of Orlando by LAP and 65% of Florida state by HAP

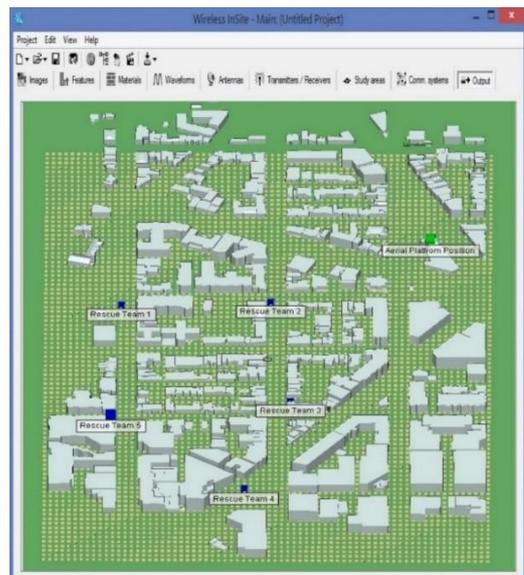


Figure 8: 3D Ad Hoc network architecture

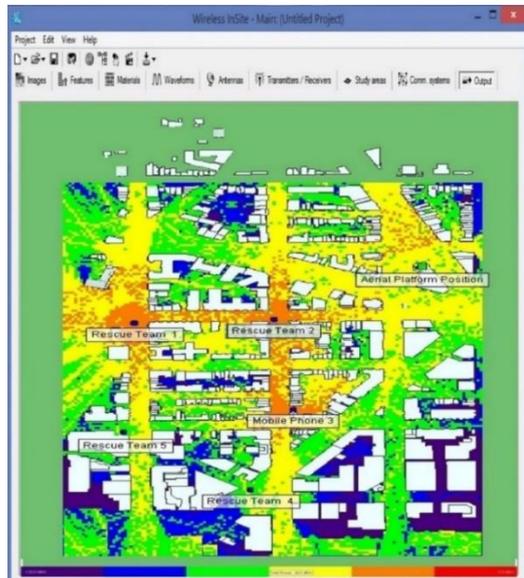


Figure 9: RSS prediction of Ad Hoc network architecture – LAP

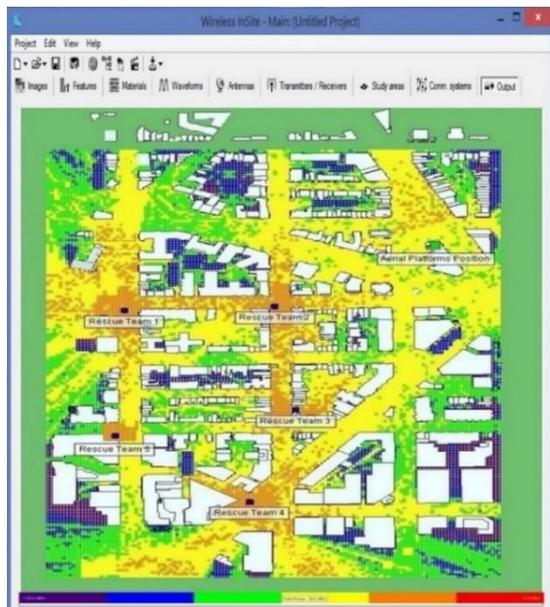


Figure 10: RSS prediction of Ad Hoc network architecture – HAP

## V. CONCLUSION

When disaster strikes, the channels that make up our communications networks often become congested. This is due to an exceptionally high level of data traffic, and emergency calls. RF channel modelling and link budget management of aerial-based communications for disaster recovery has been considered to improve performance and enhance QoS requirements. Our proposed architecture has considered both the Okumura and ATG propagation models with respect to path loss, RSS, frequency, transmitter and receiver height, and coverage. LAP shows improved performance in terms of PL, RSS, and  $P_B$ . HAP offers wider coverage, larger number of users that can be served within the footprint, and higher LoS connectivity to rescue teams.

## REFERENCES

- [1] Y. Wang, Y. Xu, Y. Zhang, P. Zhang "Hybrid Satellite-Aerial-Terrestrial Networks in Emergency Scenarios: A Survey", *IEEE China Communications*, 2017.
- [2] Y. Wang, Y. Xu, Y. Zhang and P. Zhang, "Aerial Base Stations with Opportunistic Links for Next Generation Emergency Communications", *IEEE Communications Magazine*, 2016.
- [3] G. Baldini, S. Karanasios, D. Allen, and F. Vergari, 'Survey of Wireless Communication Technologies for Public Safety', *IEEE Communications Surveys & Tutorials*, vol. 16, pp. 619–641, Jan. 2014.
- [4] F. A. Almalki, M. C. Angelides, "Considering near space platforms to close the coverage gap in wireless communications; the case of the Kingdom of Saudi Arabia", *FTC 2016 - Future Technologies Conference*, San Francisco, USA, 2016.
- [5] H. Hariyanto, H. Santoso, A. Widiawan, 'Emergency broadband access network using low altitude platform', *International Conference on Instrumentation, Communication, Information Technology, and Biomedical Engineering*, Bandung, Indonesia, 2009.
- [6] L. Reynaud, T. Rasheed and S. Kandeepan, "An Integrated Aerial Telecommunications Network that Supports Emergency Traffic", *14th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Brest, France, 2011.
- [7] K. Gomez, T. Rasheed, L. Reynaud, and S. Kandeepan, 'On the performance of aerial LTE base-stations for public safety and emergency recovery', *IEEE Globecom Workshops (GC Wkshps)*, Atlanta, USA, Dec. 2013, pp. 1391 – 1396.

- [8] M. A. Rahman, 'Enabling drone communications with WiMAX Technology', *IISA 2014, The 5th International Conference on Information, Intelligence, Systems and Applications*, Chania, Greece, 2014.
- [9] M. Helmy, T. Baykas, and H. Arslan, "Optimization of aerial base station location in LAP for disaster situations," *IEEE Conference on Standards for Communications and Networking (CSCN)*, Tokyo, Japan, 2015.
- [10] I. Aziz and Iskandar, "Disaster mitigation techniques based on LTE release 8 network employed using HAPS", *8th International Conference on Telecommunication Systems Services and Applications (TSSA)*, Kuta, Indonesia, 2014.
- [11] J. Sae, S. Yunas J. Lempiainen, "Coverage aspects of temporary LAP network," *12th annual IEEE conference on wireless on-demand network systems and services (WONS)*, Italy, 2016.
- [12] J. Kosmerl and A. Vilhar, 'Base stations placement optimization in wireless networks for emergency communications', *2014 IEEE International Conference on Communications Workshops (ICC)*, 2014.
- [13] S. Chandrasekharan, A. Al-Hourani, K. Gomez, S. Kandeepan, R. Evans, L. Reynaud, and S. Scalise, "Performance Evaluation of LTE and WiFi Technologies in Aerial Networks", *IEEE Globecom Workshops (GC Wkshps)*, Washington, USA, Dec. 2016, pp.1-7.
- [14] I. Bucaille, S. Hethuin, T. Rasheed, A. Munari, R. Hermenier, and S. Allsopp, 'Rapidly Deployable Network for Tactical Applications: Aerial Base Station with Opportunistic Links for Unattended and Temporary Events ABSOLUTE Example', *MILCOM 2013 - IEEE Military Communications Conference*, San Diego, USA, Nov. 2013, pp.1116-1120.
- [15] A. Habib and Q. ul-Islam, 'MIMO Channel Modeling for Integrated High Altitude Platforms, Geostationary Satellite/Land Mobile Satellite and Wireless Terrestrial Networks', *Journal of Space Technology*, 2013.
- [16] K. Gomez, A. Hourani, L. Goratti, R. Riggio, S. Kandeepan, I. Bucaille, 'Capacity evaluation of Aerial LTE base-stations for public safety communications', *European Conference on Networks and Communications (EuCNC)*, Paris, France, Jun. 2015, pp. 133 – 138.
- [17] D. Nalineswari and N. Rakesh, 'Link budget analysis on various terrains using IEEE 802.16 WIMAX standard for 3.5 GHz frequency', *IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India, 2015.
- [18] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, 'Modeling air-to-ground path loss for low altitude platforms in urban environments', *2014 IEEE Global Communications Conference*, Austin, USA, Dec. 2014, pp. 2898 – 2904.
- [19] F. A. Almalki, M. C. Angelides, "Empirical Evolution of a Propagation Model for Low Altitude Platforms", *Computing Conference 2017*, London, UK, 2017.
- [20] J. Deaton, "High Altitude Platforms for Disaster Recovery: Capabilities, Strategies, and Techniques for Emergency Telecommunications", *EURASIP Journal on Wireless Communications and Networking*, 2008.

# A Comparative Analysis of MANET Routing Protocols through Simulation

Callum Brill, Thomas Nash  
Department of Computer Science  
College of Charleston, Charleston, USA

**Abstract**— In this paper we investigate the performance of various mobile ad hoc network routing protocols. We use AODV and DSDV as our baseline as these are two of the most commonly used protocols in reactive and proactive routing, respectively. We compare these with the performance of an ant colony optimization implementation AntHocNet. This protocol works as a hybrid protocol, using both reactive and proactive elements. We implement the simulations across a variety of scenarios using the NS2 network simulator and present our findings.

*Keywords*-MANET; wireless sensor network; routing

## I. INTRODUCTION

Mobile ad hoc networks (MANET's) are networks comprised of mobile nodes, which are connected in a flexible manner, allowing dynamic changes. Completely lacking any kind of infrastructure and dedicated router, each node acts as its own router in these wireless mobile networks. Applications for MANETs include emergency search-and-rescue and military data acquisition operations in rough terrain, among others.

For nodes to be able to communicate with each other, routing protocols are necessary to discover connections between each node. MANET protocols fall into two general categories: table-driven (proactive) and source-initiated (reactive). Proactive protocols work to preserve the most up-to-date routing information in tables on each node. Changes in the network topology are propagated through the network to update each of the tables. In this paper, we focus on the destination-sequenced distance-vector (DSDV) routing protocol. Source-initiated routing creates routes only when demanded by the source node, hence the name. A route discovery process finds a route and maintains the route until the destination is inaccessible or the route is no longer desired. The ad-hoc on-demand distance vector (AODV) routing protocol is the reactive protocol examined in this paper [1].

Research has been conducted in creating new MANET routing protocols which incorporate machine learning (ML) techniques. Using a variety of methods and approaches, the eclectic mix of protocols have produced mixed results and each have their own benefits and drawbacks. Whether it be changes in memory or computational requirements, each technique is distributable and aims to outperform traditional methods. We have chosen to examine the AntHocNet protocol which makes use of swarm intelligence. This method has

potential for innovation and fine-tuning with different metrics used by the agents.

The organization of the paper is as follows. Section II describes the related work and provides a description of machine learning approaches to MANET routing. Section III covers the specific algorithms implemented in this paper. The methodology for conducting simulations is described in Section IV and results of previous simulations from related work is described in Section V. The results of our simulations are presented in Section VI and discussed in Section VII. Conclusions and future work are covered in Section VIII.

## II. RELATED WORK

MANET routing protocols have traditionally fallen into the two categories outlined above: proactive and reactive. We will discuss the properties of the protocols examined in this paper in further detail in Section III. Moving beyond reactive and proactive routing, further work has been done to investigate the role of ML in such routing. Förster [2] covers a wide range of ML techniques including reinforcement learning, swarm intelligence, heuristics, and mobile agents. The paper assesses each of the techniques across a variety of metrics including but not limited memory and computational requirements, tolerance to topology changes, and initial costs. Swarm intelligence and reinforcement learning are found to be the most favorable for achieving optimal results. AntHocNet [3] is one such ML protocol which implements a swarm intelligence approach to achieve improved MANET routing. It will be discussed in more detail in the next section.

To further investigate and understand the inner workings, software which can simulate the connections and network activity amongst nodes is a valuable tool. Tuteja et al. [4] use the ns-2 network simulator for comparing protocols AODV and DSDV. This discrete, event-driven simulator allows packets to be sent at specified times in addition to the creation of defined networks with designated connections. Statistics such as packet delivery, end-to-end delay and throughput can be gathered from the output of the simulation.

## III. MANET ROUTING PROTOCOLS

DSDV is a proactive routing protocol which makes use of the Bellman-Ford algorithm and builds upon it by eliminating loops in the routing table. The protocol works by sharing a single table across every node in the network. This table includes the number of hops to reach each destination along with sequence numbers used to identify the age of the record.

This table is updated on a regular basis by either sending either "full dumps" which include all of the available routing information or only the changes in the network structure [1].

AODV is a reactive protocol which actually builds on the DSDV protocol by minimizing the number of broadcasts as routes are only created when demanded. When a node wishes to communicate with a node but does not know the route to it, a path discovery is initiated. The route request is sent to the requesting node's neighbors which then forward the request to their neighbors. This process is repeated until the destination node or a node which a route to the destination node is found. Reverse paths are generated as each node tracks from which node it received the route request. When the destination node receives the request, it sends a reply to the source through this path, incorporating the information discovered [1].

AntHocNet [3], drawing inspiration from the machine learning approach known as ant colony optimization (ACO), is considered a hybrid protocol as it contains both proactive and reactive elements. Routing information is not maintained between all nodes in the network but rather held between the two communicating nodes, in a reactive fashion, and attempts to improve these routes are made while communication occurs, in a proactive fashion. ACO works by reactively sending out "ants", agents looking for paths from the source node to the destination. Each ant collects information about the quality of the path on its way to the destination and upon arrival traverses the path in reverse and updates the routing table. Included in the table is a measure of path's quality, called a pheromone. This pheromone value is used to decide the next hop from a given node. This value is important for the proactive ants, which not only monitor the path being used, but also adds probability to the ants being broadcasted to finding new paths.

Both the reactive ants and the proactive ants in AntHocNet work under different guiding mechanisms which make the protocol a powerful one. The reactive forward ants are sent out in a foraging pattern, allowing them to find different path to the destination should the route information be stale or the pheromone table being empty. The proactive ants are sent along paths in a similar fashion to data packets, where they are capable of reinforcing or degrading the strength of the pheromone, and have a small chance to be broadcaster from a node so that they are capable of finding new paths. A third form of ant, the repair and, is capable of detecting failed nodes and notifies the surrounding neighbor nodes of such a failure.

#### IV. SIMULATION METHODOLOGY

A number of network simulation tools exist, including but not limited to ns-2, ns-3, QualNet, and OpNet. For the purpose of this paper, we have chosen to use the ns-2 simulator due to availability of resources and support for MANET routing protocols. Ns-2 supports other applications such as the network animator nam which is used to visualize

the movement of nodes. We found an implementation of AntHocNet in ns-2.35, but some modifications were needed to make it fully usable, as it was not completely functional, and lacked the ability to send the necessary proactive back ants. We then created a custom awk script to calculate several values from the large trace files, gathering metrics such as packet delivery ratio, protocol overhead, energy consumption and end to end delay.

With these changes made, the three chosen MANET routing protocols are ready to be simulated. The ns-2 simulations are run using scripts written in OTcl, an object-oriented extension of Tcl, capable of simulating TCP protocols, routers and other network objects. These Tcl scripts establish nodes and their positions as well as movement and connections between the nodes. There were provided simulation scenarios, but we incorporated changes which included an energy model which allowed the tracking of energy usage. Further adjustments were made to transmission and receiving power of nodes to create impactful changes in each packet which passes through a node. The output of the simulations is a trace file which contains detailed information of network activity along with timestamps. The information is then read line at a time and grouped into time windows with length of 5 seconds. In addition to the trace files, network animator (.nam) files are created which show the topology and movement of nodes and traffic. Using the command-line tool gnuplot, graphs are made of the collected statistics.

#### V. RESULTS

Simulations were run using DSDV, AODV and AntHocNet in a defined space of 3km x 1km with 100 mobile nodes for a total time of 900 seconds, or fifteen minutes. A total of 10 iterations of the scenario were run and the results were averaged to gather statistics on total packets sent, packet delivery ratio and energy consumption.

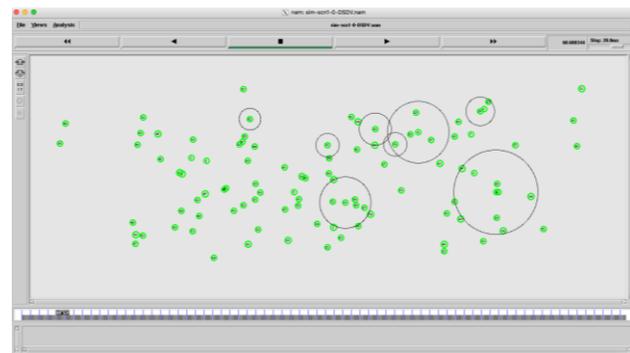


Figure 1. Nam simulation of DSDV

Fig. 1 shows a screenshot of the nam simulation for DSDV from the first iteration. Nodes are in green and move throughout the space over the time window. When a node broadcasts a message, the broadcast range is illustrated with

an expanding circle centered on the sending node, expanding to the maximum range of the broadcasting node.

Fig. 2 shows the average number of packets sent by each protocol during the simulations with each protocol labeled appropriately. Packets sent are counted in five second windows and averaged across all simulations. For this measurement, only application layer "AGT" sends are counted. The results show that AntHocNet sends the fewest number of packets and AODV and DSDV sends considerably more packets over time.

Fig. 3 shows the average packet delivery ratio for each protocol. This measurement is calculated by counting the number of packets that reach each "RTR" destination. We see that AODV and AntHocNet send a comparable number of packets which reach their destination, but DSDV has a significantly lower delivery ratio across the simulations.

Fig. 4 shows the average energy remaining energy across the nodes for each protocol. The results indicate that AntHocNet leaves the system with the lowest remaining energy while AODV consumes the least amount of energy.

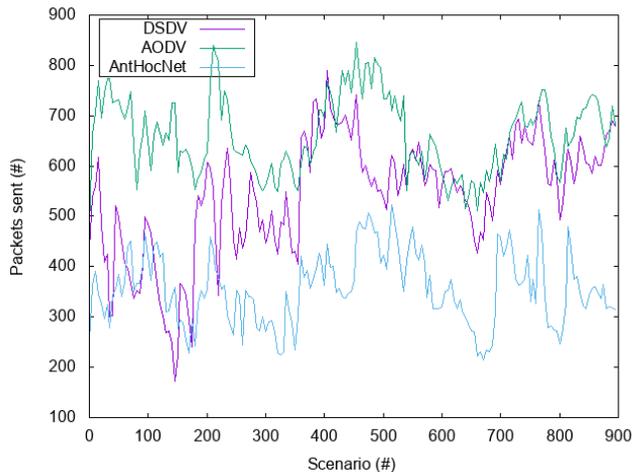


Figure 2. Average number of packets sent by all protocols

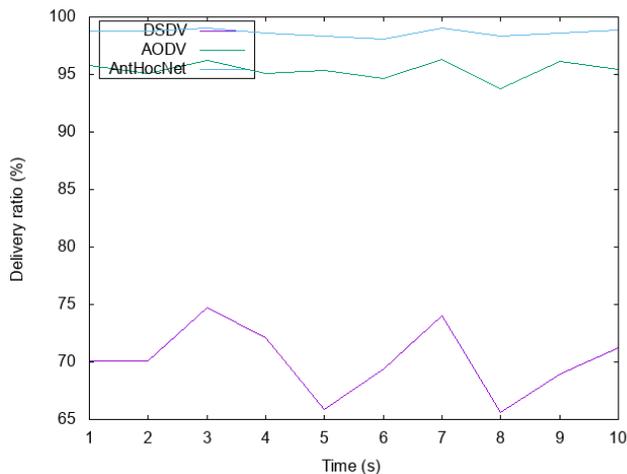


Figure 3. Average packet delivery ratio for all protocols

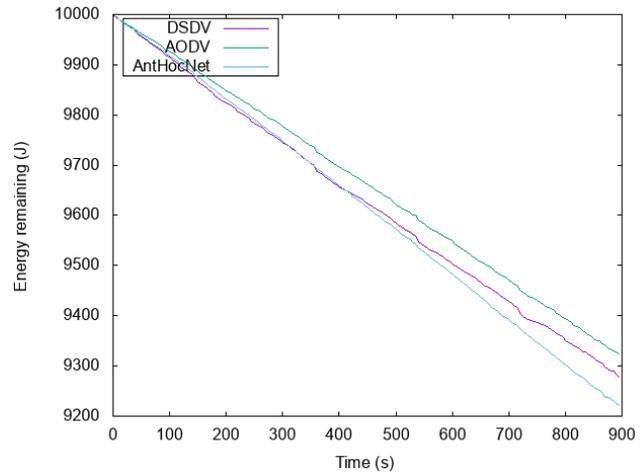


Figure 4. Average energy remaining for all protocols

## VI. DISCUSSION

The findings of the simulations align with the previous findings [4] of similar studies with AODV and DSDV. Due to AODV's reactive nature, it is much more energy efficient as packets are only sent when a route needs to be established. The lower sending and receiving rate keeps traffic low and energy levels high.

The simulations show that AntHocNet is the largest consumer of energy, though. The ACO approach is not only reactively sending out ants like in AODV, but proactively broadcasting ants across the network with greater frequency than DSDV. The protocol does redeem itself with its superior packet delivery ratio, surpassing both AODV and DSDV.

## VII. CONCLUSION

In this paper, we conducted network simulations using varied approaches of MANET routing protocols. Our simulations were carried out using the ns-2 network simulator and compared our findings against previous results from the other existing work.

We focused on two predominant protocols in MANET routing, AODV and DSDV, and looked to an advanced approach in the form of AntHocNet. We attempted to improve the pheromone calculation used in AntHocNet to make it energy-aware, but complications with the implementations found stymied this work. Once these hurdles are overcome, we can implement our method to make AntHocNet account for energy levels of nodes along the path.

Future work will include a similar analysis of other ML MANET routing protocols in hopes of improving the respective metrics. More extensive ranges of scenarios are vital to understanding and testing the effects of changes to objective functions.

## References

- [1] E. M. Royer and C.-K. Toh, "A Review of Current Routing Protocols for Ad-Hoc Mobile Wireless Networks," *IEEE Personal Communications*, vol. 6, no. 2, pp. 46-55, April 1999.

- [2] A. Förster, "Machine learning techniques applied to wireless ad-hoc networks: Guide and survey," in *3rd International Conference on Intelligent Sensors, Sensor Network and Information*, Melbourne, 2007.
- [3] G. Di Caro, F. Ducatelle and L. M. Gambardella, "AntHocNet: an Ant-Based Hybrid Routing Algorithm for Mobile Ad Hoc Networks," in *8th International Conference on Parallel Problem Solving from Nature*, Birmingham, UK, 2004.
- [4] A. Tuteja, R. Gujral and S. Thalia, "Comparative Performance Analysis of DSDV, AODV and DSR Routing Protocols in MANET using NS2," in *International Conference on Advances in Computer Engineering*, Bangalore, 2010.
- [5] M. Er-rouidi, H. Moudni, H. Mouncif and A. Merbouha, "An Energy Consumption Evaluation of Reactive and Proactive Routing Protocols in Mobile Ad-hoc Network," in *2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV)*, Beni-Mellal, 2016.

# A Multi-Channel Steganographic Protocol for Secure SMS Mobile Banking

Omega Obinna<sup>□</sup>, Eckhard Pfluegel<sup>†</sup>, Charles A. Clarke<sup>‡</sup> and Martin J. Tunnicliffe<sup>§</sup>

Faculty of SEC, Kingston University  
Kingston upon Thames, Surrey, KT1 2EE

<sup>□</sup>K1633137@Kingston.ac.uk, <sup>†</sup>E.Pfluegel@kingston.ac.uk, <sup>‡</sup>Charles.Clarke@kingston.ac.uk, <sup>§</sup>M.J.Tunnicliffe@kingston.ac.uk

**Abstract**—The advancement in mobile technologies and wireless communications has led to a rapidly growing number of users benefiting from mobile banking services. SMS banking offers a convenient mobile banking solution which is easy to implement and frequently used in many parts of the world. However, it is only viable under the assumption of secure SMS services. In this paper, a novel secure SMS banking protocol is proposed. The approach is based on a multi-channel security protocol combining low and high entropy steganography. One of the distinct advantages of this protocol is its confidentiality property against the mobile phone operator which, to our knowledge, is a novel feature. Furthermore, the required architecture is simple and only involves GSM services and one additional internet connection which can be insecure. As such it offers security, low deployment costs and would be suitable for example in rural areas or countries without individual secure home internet connections.

**Keywords**—Cryptography, Steganography Encryption, SMS-Banking, Multi-channel Security Protocols

## I. INTRODUCTION

Though traditionally, goods and services have been paid for using physical currency, this has steadily been replaced by electronic payment systems. The ubiquity of mobile phones and wireless technologies such as GPRS (General Packet Radio Service), EDGE (Enhanced Data rates for GSM Evolution) and 3G (Third Generation) have led to the emergence of mobile banking [6], [24] which has been widely accepted by customers and is gradually becoming preferred over traditional electronic banking ([24] and [37]). In some countries, affordable mobile network services such as SMS (Short Message Service) are currently used in mobile banking [25], [5], [12]. SMS is a GSM (Global System for Mobile Communications) that transmits and receives plain text messages up to 160 characters of 7 bit encoding [11]. However the nature of wireless communication leads to a number of issues regarding performance and transaction security; for example SMS messages transmitted over wireless networks are easily eavesdropped, intercepted or modified [32] and the A5 symmetric stream cipher utilized to provide security is vulnerable to attacks [26]. In 3g networks, the International Mobile Subscriber Identity (IMSI) is not encrypted and authenticated. In 4g, the LTE architecture is more IP-centric and flatter, meaning with few steps, an adversary could have access to the core network. The cost of attacking the network infrastructure is < 1000 USD with open

software and hardware (e.g. USRP, Osmocom, OpenBTS and OpenLTE). Additionally, the radio access component home eNode B (HeNB) can easily be affordable by adversaries [8], [14] and [36].

In this paper, we are motivated by the following scenario: Amara wants to implement a secure SMS banking in order to pay money into Ebere's account. However, she is within an insecure and poor infrastructure area, where there is no Internet connection in people's homes. There are public cybercafés in the neighbourhood town which offer access to the Internet. Amara has a (dual-SIM) mobile phone which however is a low powered device, not connected to the Internet, so she cannot install mobile banking apps.

The bank is considered to be trusted. On the other-hand, the cybercafé and the mobile operator environments are not—there could be a range of passive or active attacks such as eavesdropping, spoofing, delay, and replay or man-in-the-middle attacks carried out by adversaries external or internal to the system. Furthermore, steganalysis by passive or active wardens also needs to be considered. Attacks could be done by one individual adversary, or adversaries could conspire together (collaborative attack). The security analysis in this paper will highlight the security properties of our protocol with respect to the most realistic ones among these different attacks.

The main research contribution of this paper is the design of a novel cryptographic protocol based on steganography, achieving covert communication using multiple channels. The steganographic technique of the protocol could be considered as hybrid, as it combines both steganography by cover synthesis and cover modification. When used in a mobile banking setting, compared to previous proposed secure SMS banking protocols, one of its main advantages is that the need for key exchange or trusted third parties does not arise. Another advantage of this protocol is its confidentiality property against the mobile phone operator which, to our knowledge, is a novel feature. An extended version of the protocol could easily allow for additional integrity verification of the SMS transaction.

This paper is organised as follows: Section II presents the security technology relevant for this paper. Section III reviews

related work, and Section IV describes our main contribution, the security protocol. The following section analyses its security, and Section VI concludes with a brief summary and items for future research.

## II. SECURITY TERMINOLOGY AND SETTING

In this section, essential terminologies related to steganography and encryption, are described with a distinction between the two methods. In addition, two steganographic models presented in [7] are examined and adopted. The description of these terminologies sets the foundation for this paper.

### A. Security Goals

In cryptography, the purpose of achieving data confidentiality is to ensure that private information is not disclosed to unauthorized persons. Data integrity is to ensure that a private message transmitted between two participating entities has not been altered, and validation is issued by the sender (e.g. message digest). When an intended recipient is able to identify that a private data is coming from an expected *claimant* (i.e. the sender) and the private data has not been modified by an adversary, this is known as origin integrity. In this paper, the focus will mainly be on confidentiality.

### B. Encryption and Steganography

Encryption and steganography are cryptographic concepts that are used to protect sensitive information. However, their techniques of data protection are distinct from each other. Encryption is a technique that converts a plain-text to cipher-text with a secret key. This technique scrambles or encodes the message from attackers but not the existence of the message. Encryption techniques can be considered in two main categories: “symmetric encryption” and “asymmetric encryption” [35].

Steganography on the other hand, is a cryptographic technique that hides (or “camouflages” the existence of a payload (secret message)  $m$  within an unsuspecting cover medium  $c$ . Both message and the cover medium can take various formats. These include text, image audio and video files for example [19]. Once a secret message has been embedded into a cover medium, it is referred to as a *carrier object*  $o$  (sometimes called a stegogram).

A principle objective of steganography is that the act of sending and receiving a hidden payload is only known to participating senders and recipients. Additionally, a hidden payload should only be accessible and detectable by legitimate parties. Unintended parties should not be aware that secret communications are taking place and ideally a carrier medium should be able to withstand steganalysis (i.e. detailed scrutiny, detection and analysis for a payload or carrier-medium traits).

The key point to note on encryption and steganography, is that each technique has its own advantages and limitations; especially in their effectiveness and utilisation in certain situations. No cryptographic technique is a perfect solution as

one technique may be more useful than another for certain reasons. In the following subsection, descriptions of two steganographic models presented in [7] are described.

1) *High Entropy Model*: When using steganography, the cover-medium must have enough entropy to contain a payload ([13], [15], [2]). This is expressed in the following definition of a high-entropy stego-system:

*Definition 1*: A high-entropy steganographic-system  $S_H$  consists of three efficient algorithms: a probabilistic key generator  $G(\cdot)$ , a probabilistic encoding function  $E(\cdot, \cdot)$  and a deterministic decoding function  $D(\cdot, \cdot)$ .

- $G$  takes as input a security parameter and returns a (symmetric in most cases) key  $k$  – in the simplified preliminary architecture of this paper, encryption is not actually required although its use would make access to a payload more difficult, and hence further increase the security of the protocol.
- $E$  is the steganographic hiding (or embedding) method that receives the secret message  $m$  (payload) and a cover medium  $c$ , returning a carrier-object  $o$ .
- $D$  is a deterministic algorithm that returns the embedded message (payload)  $m$  by extracting it from the carrier medium.

For all messages  $m$  of a size bounded by a polynomial in the size of the cover medium, the following condition holds:  $D(k, E(k, c, m)) = m$

Definition 1 follows the traditional steganographic system where an encoding function  $E$  is used to conceal a private message  $m$  within a specific plausible cover-object  $o$ . The result is a stego-object  $s$  which may or may not require a stego-key  $k$ . It also relates to the ineffectiveness of an adversary to detect the concealed message  $m$ .

2) *Low Entropy Model*: We now define a low-entropy steganographic system.

*Definition 2*: A low-entropy steganographic system  $S_L$  consists of three efficient algorithms as in the high-entropy model but with a few differences:

- The embedding function  $E$  returns a carrier object  $o$  and additionally a secret message  $m'$  (which may be identical to  $m$ ).
- Consequentially,  $D$  has  $m'$  as an additional input parameter.

The following property now holds for all messages  $m$  irrespective of their size: if  $(o, m') = E(k, c, m)$ , then  $D(k, o, m') = m$ .

The low entropy model can be utilized if the cover-object lacks enough entropy to conceal the private message. Rather than encoding a private message in a cover object, a fake unsuspecting text, disguised as cover object could be published on a socially constrained channel. The real private message is concealed by transmission over an out of band channel (OOB). This process might not require a stego-key.

### C. Adversary Model

An adversary model is considered as a set of expectations that explains the aims and limitations of an adversary's computational knowledge and ability. An adversary can be considered as any entity that wishes to hinder, detect or modify concealed content transmitted over covert communication. To identity in totality, the security goal against omnipotent adversaries is impossible. Therefore, particularizing adversary models is essential [30]. For example, if an adversary knows a cover-object  $o(0)$  (an image) for a specific communication channel  $\mathcal{C}$ , a secret message  $m \in \mathcal{M}$  can be detectable with probability  $Prob(m \neq 0)$  by comparing objects  $o(0)$  and  $o(m)$ .  $o(m)$  is a stego-object that contains a secret message  $m$ .

In this paper, it is essential to define two realistic and strong adversary models. The two types of adversaries are based on the amount of information they could have access to.

1) *Passive warden*: A passive warden is a steganalyst who has read only access to a communication  $\mathcal{C}$  only. The goal of the passive warden is to use a detection function  $D^f$  to identify the existence of a secret message  $m$  and decides if a message is to be considered a stego object  $o$ . This could be used to prove to a third party that steganography has been used.

2) *Active warden*: A passive warden is a steganalyst that has the capability to both read and write access to a communication channel  $\mathcal{C}$ . The warden's goal is to hinder covert communication by reducing the covert channel's capacity with a distortion function  $D^*$ . Corrupting stego-objects with a distortion might affect legitimate use of a communication channel adversely (e.g noise).

### III. RELATED WORKS

The initial idea behind our work was to adapt members of a protocol family initially devised for confidential communication through various social media channels [7], [16], [18] and [29] to a mobile banking setting. In this section, we briefly review these papers and in addition, various works more broadly within the context of SMS based secure mobile banking. This is an emerging area of research and to our knowledge the relevant literature base is still quite small.

Short Messages Services (SMS) messages transmitted over mobile networks are transmitted in plain text and therefore, vulnerable to attacks. The paper [32] described an easy SMS protocol that uses MAES (Modified Advanced Encryption Standard) symmetric encryption for secure transmission of SMS messages. The symmetric keys used are stored at an

authentication server (AS) while all information of mobile subscribers is stored at a Certified Authority/Registration Authority (CA/RA). The efficiency and security of a symmetric block cipher depends on the length of the key. A larger key size compared to a smaller key size results to a slower encryption [27]. MAES has a key length of 256 bit key length compared to the 128 bit key length of AES. Some researchers have found vulnerability in AES as explained in [9] and [28].

Most publications on securing SMS banking transactions focus on the use of encryption with third parties as explained in [22], [17], [1] and [10]. However, the idea of using steganography to improve the security of SMS mobile banking was first presented in [33]. Subsequently, a text based steganographic protocol was proposed in [31]. The SMS containing the transaction details is "watermarked" using a "dummy text file" (DTF). However, if the DTF size increases, the covert information is vulnerable to attacks. Further watermarking SMS payment method described in [3] uses text based steganography, using sequences of white-space to encode positions in a shared look-up table which needs to be exchanged secretly.

Security protocols for systematic use of several channels have already been published in early works by [4], [21] and [34] in the context of ad-hoc networks. A new approach was initiated by [23] and later [7] is to disguise user generated content such as posts to an online social networks through a specific form of steganography, which makes it "socially indistinguishable" and hence undetectable (under certain assumptions). The paper gives the formal definition of the communication models already presented in Section II-B the high-entropy model using traditional steganography techniques for cover modification, selection or synthesis and the low-entropy model, based on specialised techniques for the usage in OSNs. In the corresponding protocol, an unsuspecting "fake" short textual message  $t$  is published on a socially constrained channel, scrutinized by the OSN provider, while the secret message is encrypted with a secret key derived from the seed  $t$ . This protocol is further explored in [16] where two-channel protocols are proposed, that enable users to authenticate messages on ad-hoc Social Media Platforms (SMPs). In the first protocol, a message is transmitted to a receiver on the socially constrained channel, while the message digest used for authentication must be published in a contextually relevant manner through an out-of-band (OOB) channel. The second protocol requires a commitment scheme that needs communicating parties to have a shared Authentication Transaction Account (ATA) and password for authenticating transactions. A limitation is that, the use of an ATA may violate the terms and conditions of some SMPs.

The protocols used for OSN security described in this section is closely related to the one presented in this paper. The combination of both low and high steganography in this protocol is a novel feature.

#### IV. PROTOCOL DESIGN

In this section, we design our multi-channel SMS banking protocol. Our approach is, to our knowledge, the first work that achieves secure transaction using a combination of low and high-entropy steganography. Neither a symmetric nor an asymmetric key infrastructure is required, and no trusted third party will be used. The protocol described in this section is inspired by the publications [16] and [7] with a significant difference: the protocol uses three channels.

##### A. System Architecture

The system architecture consists of the core participant customers (in our example, Amara and Eberé), the mobile network operators and the bank. Users may act as payers or payees and are required to register with the SMS banking service through their bank.

The sim cards and phone numbers should be associated with their bank account during banking registration. Customers are also required to subscribe to the GSM network through the mobile network operator, providing customers with the SMS services. In order to authenticate customers, the bank should provide customers with an interactive Point-to-Point Mobile Originated SMS consisting of a four to five digit number [20].

All SMS payment requests are transmitted to the bank through two mobile phones, operated by two distinct mobile service providers. The third channel requires the issuer bank and customer to have a public Internet connection. Our proposed protocol does not require specific software to be installed on the mobile phone, or an expensive mobile phone (smartphone) to be available. In order to be able to use our mobile banking protocol, customers only need access to an affordable dual SIM phone capable of sending SMS. For example, the Nokia 105 DUAL SIM is such a device, available for a reasonable price.

##### B. Protocol Description

Let Amara be a sender wanting to transfer money to Eberé's account through the Bank. Let  $\mathcal{M}$  be the set of all possible messages,  $\mathcal{B}$  a set of bit strings and  $\mathcal{C}$  the set of all channels. The functions  $E$  and  $D$  are encoding and decoding functions respectively. Let  $m \in \mathcal{M}$  be an SMS message and  $b \in \mathcal{B}$  be a binary value used for the protocol.

The channels  $C_i \in \mathcal{C}(i = 1, 2)$  denote two independent mobile wireless channels, implemented through suitable Mobile Network Services. The channel  $C_3 \in \mathcal{C}$  is a (not necessarily secure) internet connection.

The messages  $m_1$  and  $m_2$  are fake, unsuspecting SMS messages and  $m_3$  is a stego-object (e.g. image), using a suitable high-entropy stego-system, e.g. LSB hiding. We may use different mechanisms to implement  $C_3$ , we propose the use of a picture-upload facility on the bank's public website, where

users can share pictures of "happy customers". This does not require a secure connection and would not raise suspicion.

If a payer Amara, wishes to transfer money to a payee Eberé via her Bank, the following steps of our multi-channel protocol are executed:

- 1) She creates an SMS banking instruction message  $m$ .
- 2) She also generates two fake unsuspecting banking instructions  $m_1$  and  $m_2$ .
- 3)  $m_1$  and  $m_2$  are transmitted to the bank over two independent mobile carrier channels  $C_1$  and  $C_2$  respectively.
- 4) Amara computes  $b = m \oplus m_1 \oplus m_2$ .
- 5) The binary value  $b$  is concealed in a stego object (image) using an LSB hiding map  $b \mapsto m_3$ .
- 6) Then,  $m_3$  is sent to the bank as an HTTP post request on channel  $C_3$ .
- 7) On receiving  $m_3$  and extracting  $b$ , the bank can recover  $m$  by computing  $b \oplus m_1 \oplus m_2$ .
- 8) The bank can now transfer Amara's requested amount to Eberé's account.

The following sequence diagram illustrates the protocol:

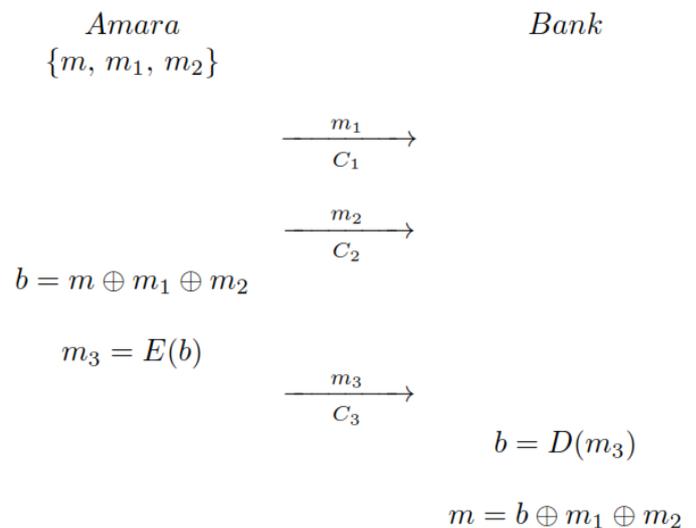


Fig. 1 A Multi-Channel Steganographic Protocol Architecture.

#### V. SECURITY ANALYSIS

In this section, we analyse the security of our protocol against various adversaries that would be to be expected in a real world scenario. The fundamental security assumptions of the protocol are that the bank is considered to be trusted, and we exclude attacks by its customers as well. On the other hand, the cybercafé, the mobile operator(s) and potentially even the government cannot be trusted as they might wish to access communication records. They could carry out passive or active attacks such as eavesdropping, man-in-the-middle attacks or steganalysis. Attacks could be perpetrated by one individual

adversary, or adversaries could conspire together (collaborative attack).

#### A. External SMS Interceptor

We consider this adversary as a passive or active warden that could intercept or modify our SMS banking instruction in an insecure GSM network. Due to the genuine format of the message, he will not suspect its content to be fake, and is led to believe the intercepted information. However, if this adversary pretends to be a legitimate sender and modifies  $m_1$  to  $\hat{m}_1$ , he will at the worst interrupt the scheme. This attack will be detected by the bank, as the reconstructed real message  $m$  will not make sense.

#### B. Mobile Network Operators – Employees and Management

We describe these adversaries as passive and active attackers. If the management of the mobile network operators are genuine and not aware of the malicious activities that could be committed by an employee, then unauthorized modification and interception of transmitted data could be achieved. However, if both adversaries intercept  $m_1$  and  $m_2$  independently or even when colluding,  $m_3$  still remains secret, only causing a (detectable) interruption of the protocol.

#### C. Cybercafé Users and Owner

Both passive and active attacks could be carried out by other users of the cybercafé, or potentially its owner. Negligence could potentially contribute to various threats. These threats could be a malware that could lead to unauthorized disclose in public cybercafés. Amara could be monitored by a malware or hardware device such as a key-logger. However, the impact of these attacks is the same as that of the SMS interceptor attack.

#### D. ISP of Cybercafé and Government

The most powerful attackers however, are the ISP of the Cybercafé or even the Government, who might use suitable legislation, have access to connection details and traffic information. If all three channels  $C_1$ ,  $C_2$  and  $C_3$  can be intercepted simultaneously, the message  $m$  can be reconstructed or altered, and the protocol is compromised.

## VI. CONCLUSION

In this paper, we have proposed a novel secure SMS banking protocol that combines low and high entropy steganography. This approach has been influenced by utilizing OSN multichannel security protocols and protocols that are used in ad-hoc networks. One of the key features of this protocol is its steganographic confidentiality property against the mobile phone operator, malicious employee at the mobile network operators and external SMS interceptor which, to our knowledge, is a novel feature. The advantages and limitations of this protocol are summarised as follows.

#### A. Advantages

- Our steganographic protocol does not use encryption and hence does not require the bank to exchange a secret key with customers. The need for a security infrastructure such as Trusted Third Parties or Public key Infrastructures does not arise.
- The architecture is very simple and cheap to implement, and could be deployed very rapidly. In terms of infrastructure it only requires GSM (2G) services and one additional internet connection which can be insecure. It would be suitable for example in rural areas or countries without individual secure home internet connections, for users who do not have access to modern smart-phones and more sophisticated secure banking mobile apps.
- Short message service is still in use by billions of people around the world especially in developing countries.

#### B. Limitations

- For this protocol to be successful, users require a dual SIM phone (or two mobile devices).
- If only one or two channels are attacked, then the confidential SMS banking instruction  $m$  remains secret. However, if adversaries could collude and intercept or modify all three channels  $C_1$ ,  $C_2$  and  $C_3$ , then the secret message  $m$  can be reconstructed or altered and the protocol is compromised.
- The protocol can be interrupted by modification attacks on the individual channels; however this interruption will be detected by the bank.

Items of immediate next research would be to add message integrity verification to the different channels. This could be achieved relatively easily by using secure hash functions.

In addition, the protocol needs to be strengthened against typical network attacks such as delay or replay.

## REFERENCES

- [1] Meer Soheil Abolghasemi, Taha Yasin Rezapour, and Reza Ebrahimi Atani. A novel protocol for the security of SMS-based mobile banking: Using GSM positioning techniques and parameters. In *IKT 2013 – 2013 5th Conference on Information and Knowledge Technology*, pages 97–101, Shiraz, may 2013. IEEE.
- [2] Shahzad Alam, S. M. Zakariya, and M. Q. Rafiq. Analysis of modified LSB approaches of hiding information in digital images. In *Proceedings - 5th International Conference on Computational Intelligence and Communication Networks, CICN 2013*, pages 280–285, Mathura, India, 2013. IEEE.
- [3] Syed Bahauddin Alam, Md Nazmus Sakib, A. B M Rafi Sazzad, Celia Shahnaz, and Shaikh Anowarul Fattah. Digital security algorithm for GSM incorporated virtual e-banking protocol using watermarking technique. In *2010 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2010*, pages 420–423, Luxor, dec 2011. IEEE.
- [4] D Balfanz, DK Smetters, P Stewart, and HC Wong. Talking to Strangers: Authentication in Ad-Hoc Wireless Networks. In *NDSS*. Citeseer, 2002.
- [5] I. Bankole, Felix Olu and Bankole, Omolola Ola and Brown. Mobile Banking Adoption in Nigeria. *The Electronic Journal on Information Systems in Developing Countries*, 47(2):1–23, 2011.

- [6] Stuart J. Barnes and Brian Corbitt. Mobile banking: concept and potential. *International Journal of Mobile Communications*, 1(3):273–288, jan 2003.
- [7] Kasper B Beato, Filipe and De Cristofaro, Emiliano and Rasmussen. Undetectable Communication: The Online Social Networks Case. In *Privacy, Security and Trust (PST)*, 2014 Twelfth Annual International Conference on, pages 19–26, Toronto, ON, Canada, 2014. IEEE.
- [8] Nicolas Bikos, Anastasios N and Sklavos. LTE / SAE Security Issues on 4G Wireless Networks. *IEEE Security & Privacy*, 11(2):55–62, 2013.
- [9] Alex Biryukov, Orr Dunkelman, Nathan Keller, Dmitry Khovratovich, and Adi Shamir. Key Recovery Attacks of Practical Complexity on {AES}-256 Variants with up to 10 Rounds. In *Advances in Cryptology – EUROCRYPT 2010: 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, number October 2000, pages 299–319, French Riviera, 2010. Springer, Berlin, Heidelberg.
- [10] Sriramulu Bojjagani and V. N. Sastry. SSMBP: A secure SMS-based mobile banking protocol with formal verification. In *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob 2015*, pages 252–259, Abu Dhabi, United Arab Emirates, 2015. IEEE.
- [11] J Brown, B Shipman, and R Vetter. SMS: The Short Message Service. *Computer*, 40(12):106–110, dec 2007.
- [12] Nor Azlina Bt Abd Rahman, Khalida Shajaratuddin Bt Harun, and Yusnita Bt Yusof. SMS banking transaction as an alternative for information, transfer and payment at merchant shops in Malaysia. In *3rd International Conference on Information Technology and e-Services, ICITeS 2013*, Sousse, mar 2013. IEEE.
- [13] Giacomo Cancelli and Mauro Barni. MPSteg-color: A new steganographic technique for color images. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4567 LNCS, pages 1–15. Springer Berlin Heidelberg, 2007.
- [14] Jin Cao, Maode Ma, Senior Member, Ieee Hui Li, Yueyu Zhang, and Zhenxing Luo. A Survey on Security Aspects for. *IEEE Communications Surveys & Tutorials*, 16(1):283–302, 2014.
- [15] Nasir Chandramouli, Rajarathnam and Memon. Analysis of LSB based image steganography techniques. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 3, pages 1019–1022, Thessaloniki, Greece, oct 2001. IEEE.
- [16] Dimitris Clarke, Charles A and Pfluegel, Eckhard and Tsaptsinos. Multi-channel overlay protocols: Implementing ad-hoc message authentication in social media platforms. In *Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), 2015 International Conference on*, pages 1–6, London, UK, 2015. IEEE.
- [17] Surakarn Duangphasuk, Maykin Warasart, and Supakorn Kungpisdan. Design and accountability analysis of a secure SMS-based mobile payment protocol. In *ECTI-CON 2011 - 8th Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI) Association of Thailand - Conference 2011*, pages 442–445, Khon Kaen, Thailand, 2011. IEEE.
- [18] Charles A. Clarke Dimitris Tsaptsinos Eckhard Pfluegel, Joakim G. Randulff and James Orwell. Building Secure ICT through Virtual Private Social Networks: A Multi-Channel Mobile Instant Messaging Approach. In *9th CMI Conference on Smart Living, Cyber Security and Privacy*, Copenhagen, nov 2016.
- [19] Jessica Fridrich. *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge University Press, Cambridge, 1 edition, 2009.
- [20] Philip Garner, Ian Mullins, Reuben Edwards, and Paul Coulton. Mobile terminated SMS Billing - Exploits and security analysis. In *Proceedings - Third International Conference on Information Technology: New Generations, ITNG 2006*, pages 294–299, Las Vegas, 2006. IEEE.
- [21] Christian Gehrmann, Chris J Mitchell, and Kaisa Nyberg. Manual authentication for wireless devices. *RSA Cryptobytes*, 7(1):29–37, 2004.
- [22] Hany Harb, Hassan Farahat, and Mohamed Ezz. SecureSMSPay: Secure SMS mobile payment model. In *2nd International Conference on Anti-counterfeiting, Security and Identification, ASID 2008*, pages 11–17, Guiyang, aug 2008. IEEE.
- [23] Ion, F. Beato, S. Capkun, B. Preneel, and M. Langheinrich. For some eyes only: Protecting online information sharing. In *CODASPY 2013 - Proceedings of the 3rd ACM Conference on Data and Application Security and Privacy*, pages 1–12, San Antonio, Texas, feb 2013. ACM.
- [24] Jesu’s Te’llez Isaac and Universidad De Carabobo. Secure Mobile Payment Systems. *IT Professional*, 16(3):36–43, jun 2014.
- [25] Md. Subrun Jamil and Fouzia Ashraf Mousumi. Short messaging service (SMS) based m-banking system in context of Bangladesh. In *Proceedings of 11th International Conference on Computer and Information Technology, ICCIT 2008*, pages 599–604, Khulna, dec 2008. IEEE.
- [26] Maria Kalenderi, Dionisios Pnevmatikatos, Ioannis Papaefstathiou, and Charalampos Manifavas. Breaking the GSM A5/1 cryptography algorithm with rainbow tables and high-end FPGAs. In *Proceedings - 22nd International Conference on Field Programmable Logic and Applications, FPL 2012*, pages 747–753, Oslo, aug 2012. IEEE.
- [27] Yahya S. Khiabani, Shuangqing Wei, Jian Yuan, and Jian Wang. Enhancement of secrecy of block ciphered systems by deliberate noise. *IEEE Transactions on Information Forensics and Security*, 7(5):1604–1613, 2012.
- [28] Chong Hee Kim. Improved differential fault analysis on AES key schedule. *IEEE Transactions on Information Forensics and Security*, 7(1):41–50, 2012.
- [29] Eckhard Pfluegel, Charles Clarke, Joakim Randulff, Dimitris Tsaptsinos, and James Orwell. A secure channel using social messaging for distributed low-entropy steganography. In Knud Erik Khajuria, Samant, Sørensen, Lene and Skouby, editor, *Cybersecurity and Privacy - Bridging the Gap*. River Publishers Series in Communications, 2017.
- [30] Rainer Bo¨ hme. *Advanced Statistical Steganalysis*. Springer Science & Business Media, 2010.
- [31] Md. Nazmus Sakib, a.B.M. Rafi Sazzad, Syed Bahauddin Alam, Celia Shahnaz, and Shaikh Anowarul Fattah. Security Enhancement Protocol in SMS-Banking using Digital Watermarking Technique. In *2010 Fourth UKSim European Symposium on Computer Modelling and Simulation*, pages 170–173, Pisa, nov 2010. IEEE.
- [32] Neetesh Saxena and Narendra S. Chaudhari. EasySMS: A protocol for end-to-end secure transmission of SMS. *IEEE Transactions on Information Forensics and Security*, 9(7):1157–1168, apr 2014.
- [33] Mohammad Shirali-Shahreza. Improving mobile banking security using steganography. In *Proceedings - International Conference on Information Technology-New Generations, ITNG 2007*, pages 885–887, Las Vegas, NV, apr 2007. IEEE.
- [34] L Shondeep, Clean Energy, Clean Energy, and References Cited. Secure communication method and apparatus, 1995.
- [35] William Stallings. *Cryptography and network security principles and practice*. Pearson, 7 edition, 2017.
- [36] Guan-Hua Tu, Chi-Yu Li, Chunyi Peng, Yuanjie Li, and Songwu Lu. New Security Threats Caused by IMS-based SMS Service in 4G LTE Networks. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS’16*, pages 1118–1130, Vienna, 2016. ACM.
- [37] Hsiao-Cheng Yu, Kuo-Hua Hsi, and Pei-Jen Kuo. Electronic payment systems: an analysis and comparison of types. *Technology in Society*, 24(3):331–347, aug 2002.

## **Session 9: Cyber Security**

Title: On the Cost of Cyber Security in Smart Business

(Authors: Igor Ivkic, Stephan Wolfauer, Thomas Oberhofer, Markus G. Tauber)

Title: Evaluation of Ensemble Machine Learning Methods in Mobile Threat Detection

(Authors: Sanjay Kumar, Ari Viinikainen, Timo Hamalainen)

Title: Evaluation of AV Systems Against Modern Malware

(Authors: Abidullah Zarghoon, Irfan Awan, Jules Pagna Disso, Richard Dennis)

Title: Security Enhancements to TLS For Improved National Control

(Authors: Lamy Alqaydi, Chan Yeob Yeun, Ernesto Damiani)

# On the Cost of Cyber Security in Smart Business

Igor Ivkic, Stephan Wolfauer, Thomas Oberhofer, Markus G. Tauber

University of Applied Sciences Burgenland

Eisenstadt, Austria

e-mail: {igor.ivkic,stephan.wolfauer,thomas.oberhofer,markus.tauber}@fh-burgenland.at

**Abstract**—In a world, as complex and constantly changing as ours cloud computing is a driving force for shaping the IT landscape and changing the way we do business. Current trends show a world of people, things and services all digitally interconnected via the Internet of Things (IoT). This applies in particular to an industrial environment where smart devices and intelligent services pave the way for smart factories and smart businesses. This paper investigates in a use case driven study the potential of making use of smart devices to enable direct, automated and voice-controlled smart businesses. Furthermore, the paper presents an initial investigation on methodologies for measuring costs of cyber security controls for cloud services.

**Keywords**—cloud-services; security controls; cost estimation

## I. INTRODUCTION

The internet boom in the late 90s opened the gates to a new era of the information age [1]. In this era software applications were shaped in world-wide connectivity, look and feel and the time to market [2] decreased significantly. On top of the world wide web, mobile devices along with app stores expanded the boundaries of personal computers (PC). Instead of being bound to the stationary PCs at home, these devices enabled users to operate globally. In addition to that the app stores invited everybody with a computer to professionally develop and sell apps. The latest “game changer” in information technology (IT) transformed the landscape of infrastructure by declaring on premise installations as obsolete in a nearby future. Cloud computing, as Armbrust et al. [3] stated, offers a diverse variety of services, allocates the necessary infrastructure resources and charges in a pay-as-you-go manner. As a result, many of today’s big companies were built on good ideas implemented as apps or services and operated in the cloud.

Besides cloud computing the IoT and Industry 4.0 were among the latest IT trends [4]. IoT in combination with cloud computing promises to turn our computers and automated machines into Cyber Physical Systems (CPS) [5] which leads to Industry 4.0 [4]. In an Industry 4.0 “smart factory” [6], CPS are communicating with other CPS and with human beings in real time over the IoT. Even though these two trends are very promising there are a lot of challenges to overcome before an Industry 4.0 becomes reality. One of them is to guarantee a secure communication in CPS [7] and to measure the effort it takes (cost) in order to do so. Another challenge is the simple lack of prototypes to demonstrate a functioning IoT based, secure and cost-efficient “smart business”.

In this paper, we evaluate whether a smart device could be introduced to a company to communicate via a cloud services and to automate manually performed tasks. Building on this we present a business-model and its architectural design in a use case driven study including security considerations. In a final step, we apply security controls suggested by ISO 27017 [8] in

combination with Six Sigma [9] to first, eliminate security risks in the presented use cases and to second, measure additional expenses resulting from them. Our contribution in this paper towards applying an established security standard and a quality management method in a use case driven study is twofold:

- firstly, we derive its business need and the required technical design including security considerations and
- secondly, we present an initial approach to eliminate security risks and measure the effort it takes to do so.

The rest of the paper is organized as follows: Section II summarizes the related work in the field, followed by the presentation of a use case study in Section III. Finally, in Section IV we introduce a high-level process flow based on Six Sigma for identifying, categorizing, analyzing, eliminating security risks and measuring the resulting costs.

## II. RELATED WORK

Guaranteeing security and measuring the resulting costs of cloud services faces considerable difficulties. Related work in [10] and [11] has developed research results about improving the service quality using Six Sigma. In contrast to that the work in [12] presented a management process for “Security Policy Management” within the Six Sigma framework. In [13] Chen & Sion present solutions of how to protect data and guarantee security in the public cloud. However, none of these works presented solutions of how to apply ISO 27017 controls to secure cloud services and how to measure the effort it takes to guarantee security of a cloud service. This paper presents an approach for applying an established standard to secure a cloud service and to use an established quality management method to measure resulting costs. In this regard, this papers approach offers a method for guaranteeing security of cloud services and quantifying the implementation expenses.

## III. BUSINESS-MODEL & USE CASE

Amazon Echo and the Amazon Web Services (AWS) meet all requirements needed in the following business-model and Use Cases I and II. Figure 1 shows the basic concept of Amazon Echo communicating with the AWS via Amazon Voice Service (AVS) including the processing of the user commands via Alexa service, the Spoken Language Understanding (SLU) and the Amazon Skill Kit (ASK) [14]:

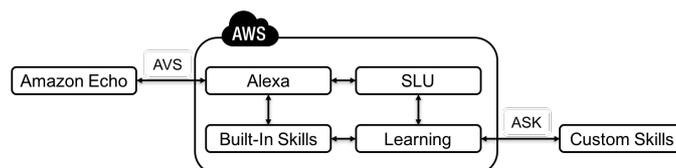


Figure 1. Amazon Echo & AWS Technology. (adapted from Soper, 2016)

We use Amazon Echo and AWS to develop a smart-business case for a fictive company with branches in two cities (City-A and City-B). The company’s core business is delivering packages to customers. At first, an ordered package is delivered to one of the two branches and then transported by delivery trucks to the customers. Usually, the truck is loaded in the morning in City-A before the driver departs to City-B delivering packets on the way. Typically, on the following day, the same truck departs in City-B and heads back to City-A supplying customers on its way, and so on. The key to success for the business model is the organization and direct communication between the two branches and their delivery truck drivers. To save time and costs, voice controlled messaging (active and passive) is considered for automated routine jobs via Amazon Echo. Three deployed devices use AWS for computation (i.e.: convert natural language in computer understandable language, vice versa) and as a dispatching service (i.e.: sending voice mails). For the use case (Figure 2) we also consider ISO 27017 and Six Sigma.



Figure 2. Use Case – Architecture, based on use case description below

A. Use Case I

The finance manager in City-A is planning to improve some accounting tasks, especially the accounting process at the end of a month. Usually, each employee is committed to daily time recording which must be released at the end of the month. Only when the time recordings of an employee have been released the employees of the finance department can start with the accounting process. Unfortunately, it repeatedly happens, that employees forget to release their time recordings in time. This mistake leads to a delay of accounting which means increased costs of administration and effort to finally finish the accounting process. The following sequence diagram shows the Use Case I as described in the next paragraphs:

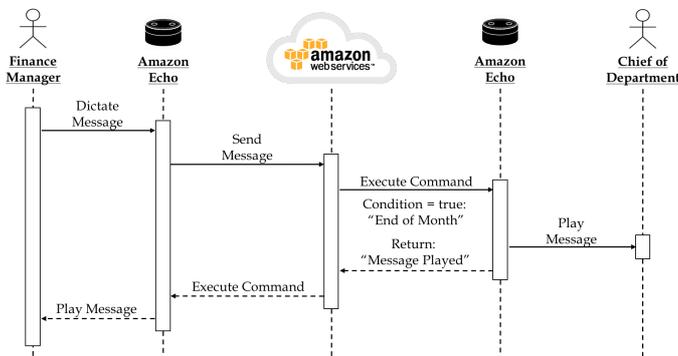


Figure 3. Use Case I – Automated “End-of-Month” Reminder.

So, to fix this issue once and for all the finance manager dictates a voice message to Alexa with the condition that his message should be played to the Chief of Department in City-B at the end of each month. The dictated message is meant to remember the colleague from the other branch to remind the employees to release their time recordings in time.

Use Case I in Figure 3 shows the entire “Reminder Process” starting with the message dictated by the Finance Manager to the message being played to the Chief of Department in City-B. The Use Case I does not show, that the process above is created one time and then repeatedly executed at the end of every month. All the commands necessary to understand and execute the Finance Managers requests are provided by the AWS.

B. Use Case II

The Chief of Department in City-B plans to invite the Finance Manager and a Truck Driver to talk about a customer complaint. Instead of using his email client and wasting time on searching for a date where all three attendees were free, he decides to ask Alexa for help. He uses the key word “Alexa Calendar” to organize an appointment on the next possible date for all attendees. Alexa sends the request to the AWS which combs through all calendars until a date is found and finally sends an invitation to all attendees. The following sequence diagram shows the described Use Case II:

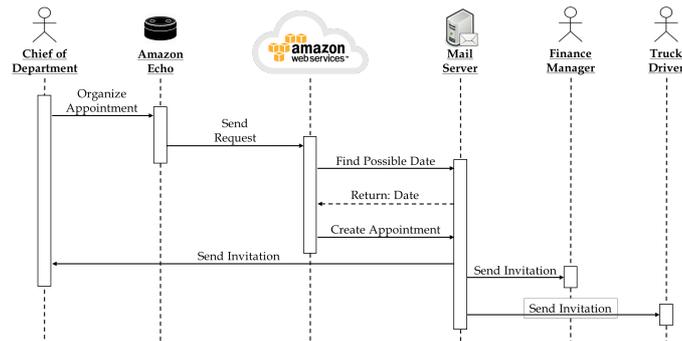


Figure 4. Use Case II – Automated Appointment Scheduling.

Figure 4 shows how the additional features of the business-model can be used to combine the AWS with i.e.: a mail server. In this scenario manually performed tasks like “creating a new appointment”, “searching for free dates in calendars” and “sending an invitation to a meeting” become obsolete. It is another nice example of how simple it is to use cloud services as a building block system to improve working routines or even entire processes.

IV. EVALUATION

After presenting the business-model architecture and the two use cases in the previous chapter, the evaluation in the following is three-fold. First, the Open Web Application Security Project (OWASP) top ten cloud security risks are categorized by their importance to the presented use cases (Section A). Based on that ranking, the ISO 27017 is used then to eliminate the top three risks (Section B). In this step, it is demonstrated how the security controls as proposed by the ISO

standard are applied to eliminate the risks in the use cases. Finally, an adapted version of the Six Sigma method, is proposed to identify, categorize, analyze, eliminate and measure security risks (Section C).

A. Cloud Security Risks

The OWASP is a non-profit organization operating worldwide with the main goal to improve the security of software. This organization is on a mission to point out security risks to individuals and organizations to create a base for informed decision making about security. In order to achieve this goal, the OWASP community started many projects focusing on different issues. One of these projects is the “OWASP Cloud – 10 Project” with the goal to maintain a list of top 10 security risks faced with cloud computing. The following list shows the top 10 cloud computing risks according to OWASP (detailed description of risks in [15]):

- R1 – Accountability and Data Ownership
- R2 – User Identity Federation
- R3 – Regulatory Compliance
- R4 – Business Continuity and Resiliency
- R5 – User Privacy and Secondary Usage of Data
- R6 – Service and Data Integration
- R7 – Multi Tenancy and Physical Security
- R8 – Incidence Analysis and Forensic Support
- R9 – Infrastructure Security
- R10 – Non-Production Environment Exposure

Generally, all ten risks should be analyzed and eliminated to guarantee maximum security for a cloud service. Unfortunately, that would be beyond the scope of this paper. Therefore, a risk analysis [16] [17] [18] was carried out to categorize the top ten risks by their relevance to the use cases and their severity of consequences. Figure 5 shows a risk matrix including the results of the risk analysis:

Relevance	Severity of Consequences				
	Very Low	Low	Medium	High	Very High
Very High				R9	R6
High			R10		R4
Medium			R3, R7, R8		
Low					
Very Low	R2			R1, R5	

Figure 5. Risk Analysis of OWASP Top 10.

In the presented use cases the smart devices (Amazon Echo) and cloud services (AWS) are mainly used to automate tasks which are usually performed manually. So, the cloud service uses input (voice commands) to either inform/remind other users or create appointments automatically. Because all these tasks could be performed manually by the user without creating additional risks (caused by cloud computing) R1 & R5

were categorized as “very low” (relevance) / “high” (severity of consequences). R2 can be left out of the discussion, because only one cloud provider (AWS) was used in the use cases. The next three risks (R3, R7 and R8) would have been significant if almost all cloud providers did not enable the option to choose the location of your data. But, they are still very serious risks which should not be handled carelessly. Just as much prudence is needed with R10, because security is as much important to productivity systems as it is to development environments. Although, in the presented use cases, there is no need to constantly redevelop the services. So, R10 is an important risk, but not the most important one in the given use cases.

The most important issues in Use Case I and II are R4, R6 and R9. The highest ranked issue is R6, because all information transferred from smart device to smart device (Amazon Echo to Amazon Echo, end-to-end) via cloud services (AWS) must be encrypted. Next on the list is infrastructure security (R9) which means that the smart devices need to be physically secured from theft and/or misuse. Finally, the business must go on no matter what. This does not mean that the cloud services need extra treatment. It means that it is important to have backup devices in case of dysfunction to keep the users (and therefore the business) non-stop available.

B. ISO 27017

The ISO 27017 is a code of practice for information security controls for cloud services based on the standards ISO 27001 and ISO 27002 [19]. In this regard, ISO 27017 provides guidance on information security aspects of cloud computing, recommending and assisting with the implementation of cloud-specific information security controls [19]. The standard was published in 2015 [19] and provides cloud-based guidance on 37 of the controls in ISO 27002 [20]. In addition to that it addresses the following seven cloud controls [20]:

- WHO is responsible FOR WHAT (customer/provider)
- removal/return of assets on terminated contracts
- protection and separation of the customer’s virtual environment
- virtual machine configuration
- administrative operations and procedures associated with the cloud environment
- cloud customer monitoring of activity (in the cloud)
- virtual and cloud network environment alignment

When using ISO 27017 in combination with ISO 27001 both the cloud service provider and the cloud service customer benefit of the standards. For instance, when looking at section A.9.4.1 of ISO 27001 in regard of restricting access, the guidance suggests the following controls [19]:

- The cloud service customer should ensure that access to information in the cloud service can be restricted in accordance with its access control policy and that such restrictions are realized.

- The cloud service provider should provide access controls that allow the cloud service customer to restrict access to its cloud services, its cloud service functions and the cloud service customer data maintained in the service.

The following table shows the suggested changes by ISO 27017 to the established standards ISO 27001 / 27002 [21]:

TABLE I. COMPARISON OF CHANGES IN ISO STANDARDS (ADVISERA, 2015)

Section	ISO 27001 / ISO 27002 Control Section	Level of Change in ISO 27017
S5	Information Security Policy	MODERATE
S6	Organization of Information Security	MODERATE
S7	Human Resource Security	LOW-MODERATE
S8	Asset Management	LOW-MODERATE
S9	Access Control	HIGH
S10	Cryptography	MODERATE
S11	Physical and Environmental Security	LOW-MODERATE
S12	Operations Security	MODERATE-HIGH
S13	Communications Security	MODERATE-HIGH
S14	System Acquisition, Development and Maintenance	MODERATE
S15	Supplier Relationships	MODERATE-HIGH
S16	Information Security Incident Management	MODERATE
S17	Information Security Aspects of Business Continuity	LOW
S18	Compliance	MODERATE-HIGH

To eliminate the risks R4, R6 and R9 the following security controls as suggested by ISO 27017 should be applied:

- S17 – Information Security Aspects of Business Continuity (to eliminate R4)
- S10 – Cryptography (to eliminate R6)
- S9 – Access Control (to eliminate R9)

There are two security controls in the ISO standard to eliminate R4. To implement the first, control the continuity of information security must be planned, implemented and reviewed. This control is realized by the proposed Six Sigma process in section C. The second control suggests having sufficient redundancies to satisfy availability requirements. In both use cases the availability of the cloud services is guaranteed by the fault tolerance and high availability of AWS [22]. But, it is also important to have enough hardware reserves in case of smart device failure. So, the delivery trucks and the two offices should each at least have one backup smart device (Amazon Echo) and a process to order a replacement.

To guarantee service and data integration (R6) in cloud services the ISO standard suggests implementing cryptographic controls. These controls take care of secure data transferring

and authentication, so that the messages can neither be intercepted nor sent/received by unauthorized users/devices. Especially the control to ensure that the person who uses the smart device is really who she/he claims to be is a major requirement for both use cases. Although, user biometric authentication can be achieved by available products like ArmorVox [23], they go hand in hand with high expenses.

The last risk (R9) can be eliminated by the access controls. As suggested by the ISO standard the smart devices need physical security controls, as well as authentication controls for the voice control functionality. To ensure physical security the devices should be locked with i.e.: Kensington desk mount cables. This control provides security against theft in the offices and/or the delivery trucks. Additionally, it is important to implement user management processes to take care of creation of users and password policies, as well as allocation of access rights and special restrictions for privileged access rights. The process must include regular reviews and updates of access rights to continuously verify and improve access controls. Furthermore, the implemented access controls should take effect each time the smart device is used. In this regard, it is important to prohibit the usage of a smart device (especially with the AWS skill to access user mail and calendar data) without an authorized user and a password. So, before a user is permitted to use the smart device she/he must authenticate herself/himself with a code or a password or i.e.: the finger scan sensor on a mobile phone.

### C. Six Sigma

Six Sigma was developed in 1986 by Motorola and is a method providing tools for organizations to improve their business processes. The idea behind Six Sigma was to remove causes of errors when detected before they lead to defects in a product or service. This is accomplished by setting up a management system that identifies errors and provides methods for eliminating them. There are two methodologies used within Six Sigma. First the DMAIC (Define, Measure, Analyze, Improve, and Control) method, which is used for improving existing business processes and second the DMADV (Define, Measure, Analyze, Design, and Verify) method for create new processes and new products or services. There are also many different management tools used within Six Sigma [23]. For the sake of completeness, the DMADV method is mentioned here, but will not be discussed in the proposed Six Sigma process.

In the following proposal, the established and well documented DMAIC method is adapted to create a new process. This process includes steps to identify cloud security risks, categorize them, search for a solution, create an implementation plan and finally measure how much effort it would take to apply the elaborated security controls. The DMAIC methodology perfectly suits the task, because of its problem-solving nature and its process-step structure. Figure 6 shows the proposed high-level process flow of the DMAIC method through its five steps [24].

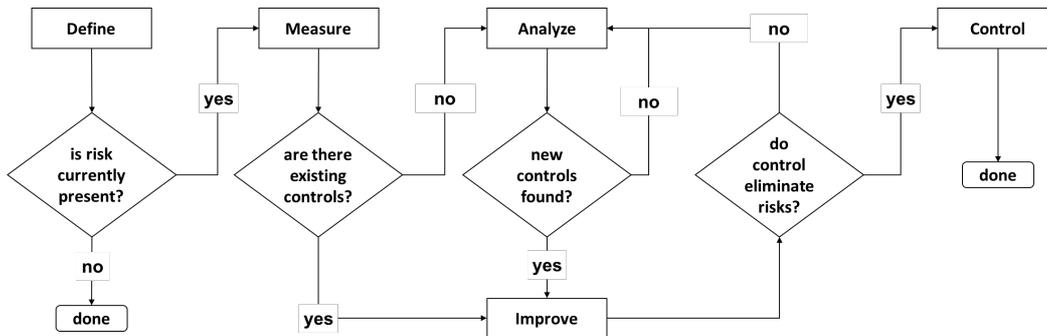


Figure 6. High-Level DMAIC Process Flow. (adapted from Hambleton, 2007)

As previously mentioned the proposed process in Figure 6 makes use of the DMAIC steps which are adapted to serve the purpose of this paper. The entire process, which is designed to take care of security risks from identifying them to preparing a solution including a cost measurement, is a never-ending process. This process should be executed at regular intervals to guarantee that security risks are identified as early as possible. The following list describes the proposed process steps in detail:

1) *Define*: the process starts by identifying and defining security risks for a cloud service. This can be achieved either by implementing a monitoring system [25] or by using any other method to identify security issues. In our use case the OWASP top 10 cloud security risks list was used and categorized by relevance and severity of consequences. Even though this paper only examined the top three risks (R4, R6 and R9) this method can be used to identify all risks as listed by the OWASP.

2) *Measure*: if the risks from the define phase are still present in the current cloud service, the next step would be to categorize them and check whether there were existing security controls to fix the issues. The categorization could be done by a risk analysis as described in Section A and shown in Figure 5. The output of this step is a list of risks grouped by their relevance to the current cloud service / process and the resulting damage, if the issue is not resolved. If there are existing security controls which could be applied the Analyze step can be skipped.

3) *Analyze*: in this step, the main goal is to find solutions to eliminate all identified and categorized security risks which cannot be solved with current security controls. As described in Section B, we used security controls (S17, S10 and S9) as proposed by an established standard (ISO 27017) to solve the top three issues (R4, R6 and R9) identified and categorized in the previous Section A. The output of the Analyze phase is a list of security controls, which can be applied to eliminate the identified issues.

4) *Improve*: so far, we found out WHAT the security risks are, IN WHAT ORDER they need to be solved and WHAT controls could be used to do so. Based on that, the next step is to create a security control implementation plan and to verify whether the control eliminates the risks or not. The plan includes the implementation steps and a list of all resources

(i.e.: all expenses like personnel, software, hardware, etc.) required. Next, these steps are prioritized and tested to see if the solutions resolve the problem.

5) *Control*: in the final step a comparative cost analysis shows how much it would cost to implement the security controls (S17, S10 and S9) to eliminate the risks (R4, R6 and R9). In the following cost analysis software-, hardware-, process-, and operational-factors are considered to calculate the implementation expenses. These factors are based on the presented use cases, the resulting security risks and security controls. Software- and hardware-costs are represented by monetary costs, while process- and operational-costs are measured by man-days (MD). In this regard, an MD is defined as an eight-hour day with an assumed hourly rate of EUR 100. This simplified cost metric will be enhanced in future work by investigating more generic cost metrics. The following table shows the cost analysis measuring the resulting implementation costs:

TABLE II. COST ANALYSIS

Costs	Security Controls / Risks		
	S17 / R4	S10 / R6	S9 / R9
Software, Cloud Service	17 € <sup>a</sup>	17 € <sup>a</sup> + 230 € <sup>d</sup>	17 € <sup>a</sup> + 230 € <sup>d</sup>
Hardware	360 € <sup>b</sup>	-	90 € <sup>c</sup> + 570 € <sup>f</sup>
Process	400 € <sup>c</sup>	800 € <sup>c</sup>	800 € <sup>c</sup>
Operation	400 € <sup>c</sup>	800 € <sup>c</sup>	800 € <sup>c</sup>
<b>monthly</b>	<b>1.177 €</b>	<b>1.847 €</b>	<b>2.507 €</b>

a. Amazon Lambda Pricing. (AWS: Development & Compute Time)

b. Amazon Echo Pricing. (3 x Amazon Echo + 3 x Amazon Echo reserves)

c. Process-/Operation-Costs. (1 MD = 8 hours \* 100 € = 800 €)

d. Voice Biometrics. (Voicelt: 10.000 API Calls per Month)

e. Kensington Desktop Lock for Amazon Echo. (3 x Desktop Lock)

f. Motorola G5 with Finger Scan Sensor. (3 x Mobile Phone)

## V. CONCLUSION AND FUTURE WORK

In this paper, we introduced a use case of for smart business by using smart devices and cloud services. We explored the technical specifications and services of Amazon Echo and explained the ASK. In this regard, we pointed out that the

built-in skills of Alexa could be extended by custom built skills. Furthermore, we have proposed an architectural design to illustrate two points. First, to show how simple it is today to use smart devices and cloud services to build powerful smart businesses with little effort. Second, we have illustrated that IoT can have a big impact both on your business and on security. In the end of section III we presented two use cases showing how Amazon Echo and Alexa can be combined to enable direct, automated and voice-based communication to automate routine tasks.

Finally, in Section IV we proposed a high-level DMAIC process to identify and categorize security risks, analyse and prepare security controls, plan their implementation and in the end, measure the cost involved when implementing them. In this regard, we first presented a method to identify (Define: OWASP) the top ten security risks and label them by their relevance to the use cases and the impact they would have, if the risks were not eliminated (Measure: Risk Analysis). Next, we presented an established standard (ISO 27017) which provides security controls for securing cloud services. Then, we presented solutions according to ISO 27017 to the top 3 security risks from the risk analysis (Analyze). In addition to that we presented how these security controls could be implemented and what resources would be needed to do so (Improve). In the last step, we evaluated the resulting expenses when applying the presented security controls to eliminated the analysed security risks (Control: Cost Analysis).

In summary, we explored the possibilities of smart devices and cloud services to enable smart businesses. In addition to that we proposed methods to identify and categorize security risks, to analyze solutions, plan implementation steps and to measure the costs to do so. The main contribution of this paper is the initial investigation on an approach for how to measure the cost of security in smart business by a combination of the above methods in a simple process. This will be enhanced in future work by considering multiple other use cases and hence more security requirements including a more generic metric for cost calculation. We will furthermore also investigate alternatives to the OWASP, the ISO standard and Six Sigma to verify a broader application of our work.

## VI. ACKNOWLEDGEMENT

Research leading to these results has received funding from the EU ECSEL Joint Undertaking under grant agreement n° 737459 (project Productive4.0) and from the partners national funding authorities FFG on behalf of the Federal Ministry for Transport, Innovation and Technology (bmvit) and the Federal Ministry of Science, Research and Economy (BMWFV).

## REFERENCES

- [1] Castells, M., "The power of identity: The information Age: Economy, society and culture, Volume II (The information age)," 2003.
- [2] Iansiti, M. and MacCormack, A., "Developing products on Internet time," *Harvard business review* 75.5, 1996, pp. 108-117.
- [3] Armbrust, M. et al., "Above the clouds: a Berkeley view of cloud," *Electrical Engineering and Computer Sciences, University of California at Berkeley*, 2009.
- [4] Hermann, M., Pentek, T. and Otto, B., "Design principles for industrie 4.0 scenarios," *System Sciences (HICSS)*, 2016 49th Hawaii International Conference on, IEEE, 2016.
- [5] Lee, E. A., "Cyber physical systems: Design challenges," *Object Oriented Real-Time Distributed Computing (ISORC)*, 2008 11th IEEE International Symposium on, IEEE, 2008.
- [6] Zuehlke, D., "SmartFactory—Towards a factory-of-things," *Annual Reviews in Control* 34.1, 2010, pp. 129-138.
- [7] Banerjee, A. et al., "Ensuring safety, security, and sustainability of mission-critical cyber-physical systems," *Proceedings of the IEEE* 100.1, 2012, pp. 283-299.
- [8] IEC, ISO Std. "ISO 27017." *Information technology-Security techniques (DRAFT)*, 2012.
- [9] Pyzdek, T. and Keller, P. A., "The six sigma handbook", McGraw-Hill Education, 2014.
- [10] Oktadini, Nabila Rizky, and Kridanto Surendro, "SLA in cloud computing: Improving SLA's life cycle applying six sigma," *Information Technology Systems and Innovation (ICITSI)*, 2014 International Conference on. IEEE, 2014.
- [11] Antony, Jiju, "Six sigma for service processes," *Business Process Management Journal* 12.2, 2006, pp. 234-248.
- [12] Anand, V., Saniie, J., and Oruklu, E., "Security policy management process within Six Sigma framework," 2011.
- [13] Singh, Shri RK Bigensana, and Lakshmi Prasad Saikia, "Present Trends of Data Protection Policies in Cloud Computing," *International Journal of Science, Engineering and Computer Technology* 5.10, 2015, pp. 347.
- [14] Soper, T., "Amazon recruiting heavily for Echo and Alexa engineers, hosts big invite-only event to find talent", *GeekWire*, Retrieved March 7, 2017, [Online], Available: <https://www.geekwire.com/2016/amazon-ramps-up-recruiting-efforts-for-engineers/>.
- [15] OWASP, "OWASP Cloud – 10/Initial Pre-Alpha List of OWASP Cloud Top 10 Security Risks", Retrieved May 8, 2017, [Online], Available: [https://www.owasp.org/index.php/OWASP\\_Cloud\\_-\\_10/Initial\\_Pre-Alpha\\_List\\_of\\_OWASP\\_Cloud\\_Top\\_10\\_Security\\_Risks](https://www.owasp.org/index.php/OWASP_Cloud_-_10/Initial_Pre-Alpha_List_of_OWASP_Cloud_Top_10_Security_Risks)
- [16] ISO, ISO31000, "31000: 2009 Risk management—Principles and guidelines," *International Organization for Standardization, Geneva, Switzerland*, 2009.
- [17] Ashley, D.B., Diekmann, J.E., Molenaar, K.R., "Risk Assessment and Allocation for Highway Construction Management," *Federal Highway Administration*, US Department of Transportation: Washington, DC, USA, 2006.
- [18] Treasury Board Secretariat of Canada, "Guide to Corporate Risk Profiles," Retrieved May 8, 2017, [Online], Available: <http://www.tbs-sct.gc.ca/tbs-sct/rm-gr/guides/gcrp-gcpropr-eng.asp?format=print>.
- [19] IsecT, "ISO/IEC 27017:2015 / ITU-T X.1631 — Information technology — Security techniques — Code of practice for information security controls based on ISO/IEC 27002 for cloud services," Retrieved May 3, 2017, [Online], Available: <http://www.iso27001security.com/html/27017.html>.
- [20] BSI Group, "ISO/IEC 27017," Retrieved May 3, 2017, [Online], Available: <https://www.bsigroup.com/en-GB/Security-controls-for-cloud-services-ISO-IEC27017/>.
- [21] Kosutic, D., "ISO 27001 vs. ISO 27017 – Information security controls for cloud services," Retrieved May 3, 2017, [Online], Available: <https://advisera.com/27001academy/blog/2015/11/30/iso-27001-vs-iso-27017-information-security-controls-for-cloud-services/>.
- [22] Amazon Web Services, "Fault Tolerance & High Availability," Retrieved May 8, 2017, [Online], Available: [https://media.amazonwebservices.com/architecturecenter/AWS\\_ac\\_ra\\_ft\\_ha\\_04.pdf](https://media.amazonwebservices.com/architecturecenter/AWS_ac_ra_ft_ha_04.pdf).
- [23] Graves, A., "What Is Six Sigma?," Retrieved April 5, 2017, [Online], Available: <http://www.sixsigmadaily.com/what-is-six-sigma/>.
- [24] Hambleton, L., "Treasure Chest of Six Sigma Growth Methods, Tools, and Best Practices," Pearson Education, 2007.
- [25] Bicaku A., Tauber M., Balaban S., Hudic A, Mauthe A., Hutchison D., "Harmonized Monitoring for High Assurance Clouds," in *CLW: IEEE 2nd Workshop on Legal and Technical Issues in Cloud Computing and Cloud-Supported Internet of Things*, 2016.

# Evaluation of Ensemble Machine Learning Methods in Mobile Threat Detection

Sanjay Kumar\*, Ari Viinikainen<sup>†</sup> and Timo Hamalainen<sup>‡</sup>

Faculty of Information Technology

University of Jyväskylä

Jyväskylä, Finland

Email: \*sanjay.k.kumar@jyu.fi, <sup>†</sup>ari.viinikainen@jyu.fi, <sup>‡</sup>timo.t.hamalainen@jyu.fi

**Abstract**—The rapid growing trend of mobile devices continues to soar causing massive increase in cyber security threats. Most pervasive threats include ransom-ware, banking malware, premium SMS fraud. The solitary hackers use tailored techniques to avoid detection by the traditional antivirus. The emerging need is to detect these threats by any flow-based network solution. Therefore, we propose and evaluate a network based model which uses ensemble Machine Learning (ML) methods in order to identify the mobile threats, by analyzing the network flows of the malware communication. The ensemble ML methods not only protect over-fitting of the model but also cope with the issues related to the changing behavior of the attackers. The focus of this study is on android based mobile malwares due to its popularity among users. We have used ensemble methods to combine output of 5 supervised ML algorithms such as RF, PART, JRIP, J.48 and Ridor. Based on the evaluation results, the proposed model was found efficient at detecting known and unknown threats with the accuracy of 98.2%.

**Index Terms**—Intrusion Detection, Machine Learning, Ensemble Methods, Supervised Machine Learning, Mobile Threats, Anomaly Detection

## I. INTRODUCTION

According to research by Sophos [1], by 2020 more than 6 billion users will be using mobile devices. Mobile devices are rapidly overtaking personal computers from web surfing to mobile banking, due to their portability and smart features. Therefore, the potential usage has caught the attention of cybercriminals who maximize their efforts to obtain user information. Most of the users do not care about the security measure of their devices and thus become the victim of these threats. These applications could lead to several mobile threats, such as theft of financial information, ransom-ware, misuse of premium SMS and theft of personal information. A traditional anti-virus can detect only 50% of the threats and on the other hand around 71% of the smart-phone users do not use any kind of anti-virus [2]. Thus there is a need for an extra layer of security at the network side to protect the users from advanced threats, which a traditional anti-virus could not detect.

Most of the NIDS use signatures to detect attacks and therefore capable of detecting only known attacks [3]. A minor modification in attack can bypass a signature based NIDS and could generate up to 90% of false alarms [4]. Deep packet inspection is also difficult when the traffic is encrypted and computationally expensive [5]. Flow based techniques are useful in combating several issues caused by encrypted

traffic [6]. Machine learning methods are getting popular in the detection of advanced threats. This transition is supported by Arp et al [7].

To build a ML classifier, a dataset is required. The well-known datasets available in the field of intrusion detection, which uses network traffic features are KDD99, DARPA 1998/1999 and ISCX 2012 IDS dataset. However, some shortcomings were observed in these datasets by [8] and [9]. These datasets are quite old and not applicable to mobile attacks. Due to non-availability of public datasets in this domain, we have used a dataset which was created in our previous research [10] for evaluation purpose. This dataset was build using the traffic generated by several benign and malicious samples. The dataset is based on bidirectional flows extracted from real malware traffic, which makes it unique. The dataset contains several threats such as unauthorized premium SMS sender, Spam sender, bots, back-door, root exploit, fake anti-virus, ransom-ware and information theft. These datasets were used to train and build the ML classifiers using several ML algorithms. The classifier is used to make predictions on new data to detect normal or malicious patterns in the traffic. A signature-based NIDS could miss the threats in the traffic, for which the signature is not yet known. However, ML classifiers can detect known and unknown threats by analyzing traffic patterns.

The main focus of this paper is the performance evaluation of ensemble ML techniques that combines output of several ML algorithms. The benefit of using ensemble method is not only to increase the efficiency of the classifier but also to reduce the risk of over-fitting the model. This can also address the problem caused by minor changes in the attack pattern, in order to avoid concept drift situation. The concept drift situation occurs in machine learning methods when the relation between the features used to train the model and target to be predicted changes over an interval of time. The concept drift causes decrease in accuracy when prediction is made on unseen data.

The structure of this paper is as follows. Section 2 focuses on previous research conducted in this area. Section 3 describes the machine learning algorithms used in this study. Section 4 focuses methodology of this research. Section 5 is based on the performance evaluation of ensemble machine learning classifiers using different datasets. Several experiments that are

performed for evaluation purposes are also explained. Finally, in Section 6, conclusion and future work of this research are outlined.

## II. RELATED WORK

Recently, a lot of research has been done in the area of machine learning to solve many cybersecurity issues. Most of the research done in the field of using ML techniques to detect Android-based malware is based on features such as system or API calls. There are only a few studies which have focused on network-based intrusion detection for android malwares [10] [11]. In some of the studies [12] [7] [13] [14], the detection engine or model need to be installed on the mobile phone to detect e.g. intrusions or malicious applications. However, most of the smartphone users do not install security solutions in their phones.

Arp et al. [7] developed Drebin which uses Support Vector Machine (SVM) to detect malicious android applications. Drebin is based on the features such as permissions, API calls and Network addresses. The detection of Drebin is limited when the malwares uses dynamic code or any obfuscation technique. Many researchers [12] [13] used malwares from MalGenome [15] dataset to generate traffic and build classifiers on various traffic based features. The concept drift has been seen in some studies [12] where the classifiers produced significant decrease in the TPR when evaluated on the unseen traffic. Input features used to train the model play important role in the field of machine learning as the attackers change their behavior with time to avoid detection. Features like IP address could lead to produce concept drift in the model. In some of the studies [13] [11], the classifiers were not tested on unseen data which is crucial part in the evaluation of ML classifiers.

Many researchers [16] [17] [6] used flow-based models for network traffic classification. Furthermore, the flow-based techniques to detect botnets were studied by several researchers [18] [19] [20]. Most of network traffic in the malware communication is encrypted [5] and therefore flow-based features seems to be efficient in detecting these threats [6].

## III. MACHINE LEARNING ALGORITHMS

In this research work, we have used the ensemble of several machine learning algorithms such as Random Forest, J48, RIDOR, JRIP and PART. The performance evaluation of these algorithms was already performed individually in our previous work [10]. In this study, we have used ensemble methods to combine the output of these ML algorithms to increase the effectiveness of the ML classifiers. The combination methods used in our study such as majority voting, maximum probability and product of probabilities were adopted from [21] [22]. ML classifiers used in the network traffic analysis becomes less efficient with time as the attackers change their behavior and this situation is known as concept drift. During the model building, each algorithm has different weight-age for each feature. Some of the algorithms have built-in feature

selection algorithms and they make the decisions on limited features. If there is any change in traffic pattern, each algorithm behaves differently. Combining the output of these individual algorithms not only increases the efficiency but also the stability in case of minor changes in the traffic pattern.

J48 is the WEKA Implementation of C4.5 [23] algorithm which was developed by Ross Quinlan in 1993. C4.5 is a decision tree based algorithm, which works on the "divide and conquer" rule. C4.5 first divides the training dataset with highest single class instances, then it checks the feature with the highest information gain in the subset and splits it into further subsets according to that feature. It repeats these steps for each subset [23].

The Random Forest (RF) [24] is one of the most popular ML algorithms used for classification, developed in 2011 by Leo Breiman. Random forest is the collection of decision trees built from random subsets of dataset (bootstraps) with random features selected in each subset. Each tree is trained by 2/3 of the dataset and remaining 1/3 is used to estimate error rate. This 1/3 of the dataset is same as the validation set in other ML algorithms, therefore there is no need for a separate validation dataset. The output of the random forest is based on the majority vote by each decision tree output.

Ridor is the WEKA implementation of Ripple-Down Rule Learner [25], which was developed by Gaines and Compton in 1995. Ridor uses Incremental reduced error pruning (IREP) algorithm [26] to build its rules. Pseudo code for IREP is mentioned in Figure 1 of [27]. Ridor generates its first rule as a default rule for one class and then builds the rules for other classes depending on the weighted error rate known as exception rules. Let's suppose there are two classes "Deny" and "Allow". First, it makes a default rule for "Deny" and then it builds up the rules for "Allow" [25].

JRIP is a WEKA implementation of RIPPER, which was proposed by William Cohen in 1995, as an optimized version of IREP [26], [27]. JRIP divides the training set into two subsets in the ratio of 2:1 in the form of grow:prune.

PART is a partial decision tree algorithm developed by Frank [28] in 1998. This algorithm works on the separate-and-conquer rule and is a combination of C4.5 rules and RIPPER algorithm, excluding the global optimization feature. This algorithm produces rules in ordered sets, which makes a decision list. The rules are based on "Best" leaf of the partial C4.5 decision tree [28].

J48 and Random Forest are the tree based algorithms while RIDOR, JRIP and PART are rule based algorithms. All of these ML algorithms have some internal validation function for tuning to avoid over-fitting e.g Random forest [24] uses 1/3 of the dataset for estimating the error rate. JRIP [27] and RIDOR [25] both use IREP which selects 2/3 of the dataset for training and 1/3 of the dataset for the pruning of the model. J48 [23] has an internal mechanism of pre-pruning and post-pruning to avoid over-fitting and PART [28] is the combination of J48 and JRIP.

Random Forest has several advantages, such as high accuracy and effectiveness on large datasets [29]. Random Forest

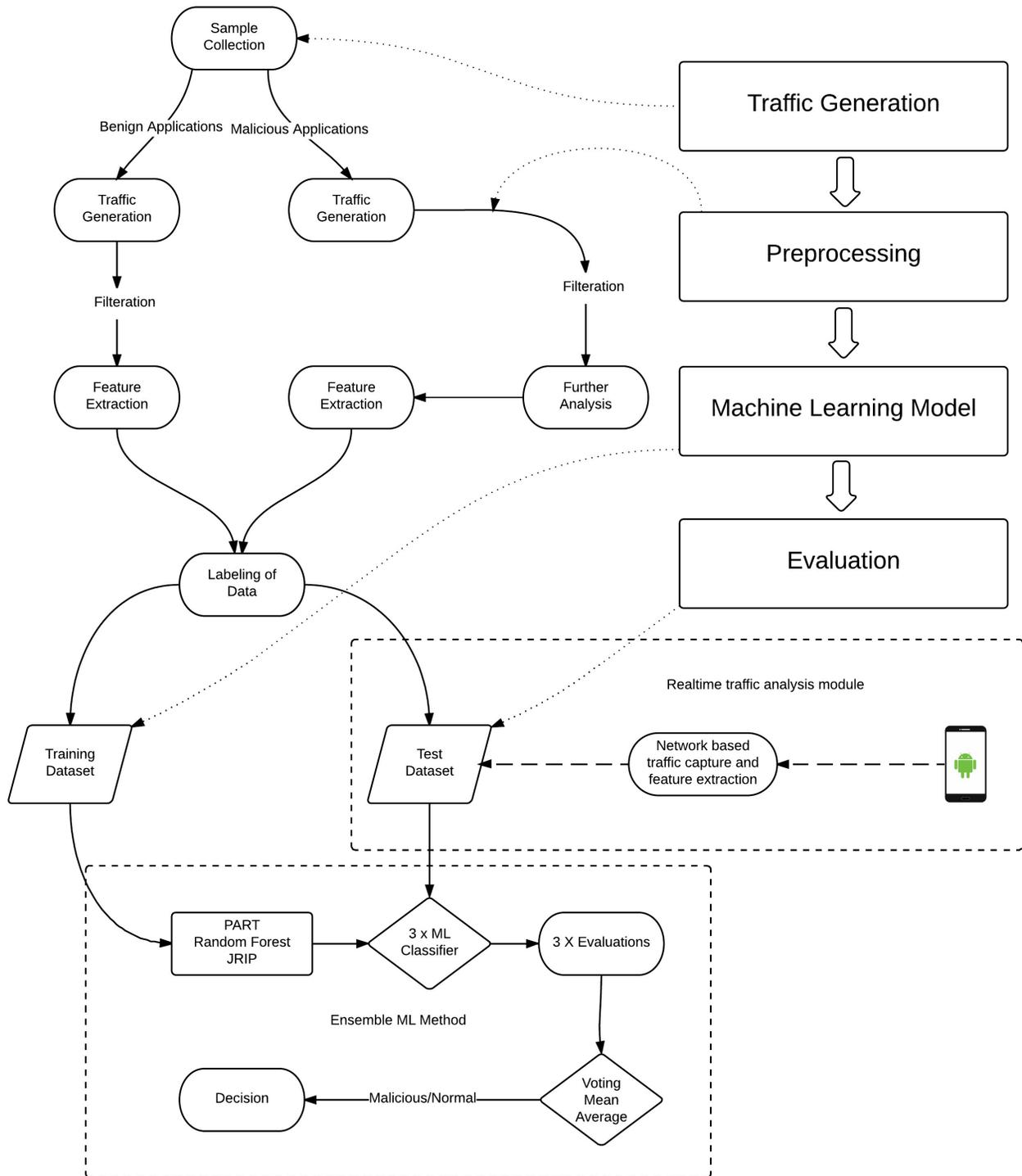


Fig. 1: Ensemble Machine Learning model for Intrusion Detection

is an ensemble of multiple decision trees. Although the output of Random Forest is hard to understand, the performance of this classifier makes it outstanding. J48 provides speed over Random Forest, but the accuracy is not as high as that of Random Forest. J48 is a decision tree so it is easy to understand. RIDOR, PART and JRIP are rule-based algorithms, so the rules generated by these algorithms can be used in any knowledge-based expert system.

#### IV. METHODOLOGY AND IMPLEMENTATION

In this research, an ensemble ML based Network intrusion detection system is proposed and evaluated, shown in Fig. 1. The first step was traffic generation, followed by filtration, feature extraction and labeling of the dataset. The dataset was used to build the ML Classifiers using WEKA [30]. In this study, the focus is only on the evaluation of ensemble ML classification model, as the evaluation of individual classifiers is already done in our previous work [10].

The overall implementation and evaluation of this model performed in four main phases, as shown in Fig. 1, comprised of Traffic generation, preprocessing, model building and evaluation of the ML model.

TABLE I: Feature List

Feature No.	Feature	Description	Value
1	Duration	Connection Duration	Real
2	DP	Destination Port	Real
3	PktSent	Packet Sent	Real
4	PktRcv	Packets Received	Real
5	PLBytesSent	Payload bytes sent	Real
6	PLBytesRcv	Payload bytes received	Real
7	IflagF	Initial Flags in Forward Direction	Nominal
8	IflagR	Initial Flags in Reverse Direction	Nominal
9	UFlagsF	Union of Flags in Forward Direction	Nominal
10	UflagR	Union of Flags in Reverse Direction	Nominal

Traffic was generated for both benign and malicious applications using the method mentioned in our previous research [10]. A number of benign applications were used to generate real traffic which were installed from Google playstore. These applications were executed at a different interval of time. Wire-shark was used to capture the packets on the interface of the virtual machines. The samples of malware families (FakeAV, DroidKungFu, OPFake, GinMaster, FakeInst and Anserver) were downloaded from Virustotal using several conditions. The number of samples completed for the study was around 600. Traffic was generated through a public sandbox "Anubis (Andrubis)" [31] and "Cuckoo" [32].

During Processing feature extraction and labeling of traffic flows was done. The features were extracted using RFC-5103 BiFlow export method [33]. The following features (see Table I) were extracted from the flows of the traffic of benign and malicious applications. Instances were then labeled as normal or malicious respectively.

##### A. Machine learning classifiers

In this study, we have evaluated the combination of 5 decision tree and rule based algorithms. Output of these algorithms can be easily interpreted by security experts and

can be integrated with the traditional NIDS. The classifiers build from these ML algorithms produce rules and trees which can be used to make predictions on new traffic to identify threats. By using ensemble methods, the combination of these ML classifiers was used to increase the efficiency of the classification model as shown in Fig. 1.

#### V. PERFORMANCE EVALUATION AND RESULTS

Several well-known parameters were used to evaluate ensemble ML classifiers.

TABLE II: Confusion Matrix

		Predicted	
		Malicious	Normal
Actual	Malicious	<i>TruePositive</i>	<i>FalseNegative</i>
	Normal	<i>FalsePositive</i>	<i>TrueNegative</i>

True Positive (TP): Malicious instance classified as Malicious.  
 False Positive (FP): Benign instance classified as Malicious  
 False Negative (FN): Malicious instance classified as Normal.  
 True Negative (TN): Benign Instance classified as Normal.

$$TPR = \frac{TP}{TP+FN}$$

$$FPR = \frac{FP}{FP+TN}$$

$$TNR = \frac{TN}{TN+FP}$$

$$FNR = \frac{FN}{TP+FN}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$

ROC (Receiver Operating Characteristic) curve is a plot between TPR and FPR at various threshold settings [34]. Area under ROC Curve (AUC) is also an important parameter in evaluating the ML classifier, this value is derived from the ROC curve and it can tell which model makes best predictions. A higher AUC value shows a better ML Classifier. Accuracy is also an important parameter to consider as it is based on both TPR and FPR.

##### A. Evaluation of Classification Model

We have performed two experiments in order to evaluate the performance of the ML classifiers using ensemble methods. In the first experiment, we have tested the classifiers using cross validation and percentage split of the same data. In the second experiment the evaluation was performed using the new unseen dataset.

1) *Experiment 1 - Ensemble Methods* : Ensemble methods combine output of several ML classifiers by different techniques such as weighted voting or measuring probability as shown in Fig 1. In this experiment, we combined the output of several classifiers by 3 combination rules as shown in the Table III - IV.

These tables show the detailed performance evaluation of ensemble methods using different combination. In Experiment 1a, we have used 10 fold cross validation method which is most widely used validation method. This method splits

TABLE III: Experiment 1a (Ensemble Methods) using Cross Validation

<b>Performance Evaluation by combining J48, RF, JRIP, RIDOR and PART</b>						
<b>Combination Rule</b>	<b>TPR</b>	<b>FPR</b>	<b>TNR</b>	<b>FNR</b>	<b>Accuracy</b>	<b>AUC</b>
Majority Voting	0.995	0.030	0.970	0.005	0.991	0.984
Maximum Probability	0.983	0.005	0.995	0.017	0.985	0.999
Product of Probabilities	0.999	0.005	0.995	0.001	0.998	0.999
<b>Performance Evaluation by combining J48, Random Forest, JRIP and PART</b>						
<b>Combination Rule</b>	<b>TPR</b>	<b>FPR</b>	<b>TNR</b>	<b>FNR</b>	<b>Accuracy</b>	<b>AUC</b>
Majority Voting	0.996	0.041	0.959	0.004	0.990	0.978
Maximum Probability	0.993	0.030	0.970	0.007	0.989	0.999
Product of Probabilities	0.994	0.025	0.975	0.006	0.991	0.987
<b>Performance Evaluation by combining J48, RF and PART</b>						
<b>Combination Rule</b>	<b>TPR</b>	<b>FPR</b>	<b>TNR</b>	<b>FNR</b>	<b>Accuracy</b>	<b>AUC</b>
Majority Voting	0.996	0.023	0.977	0.004	0.993	0.986
Maximum Probability	0.993	0.043	0.957	0.007	0.987	0.999
Product of Probabilities	0.994	0.039	0.961	0.006	0.989	0.993

TABLE IV: Experiment 1b (Ensemble Methods) using Percentage Split Validation

<b>Performance Evaluation by combining J48, RF, JRIP, RIDOR and PART</b>						
<b>Combination Rule</b>	<b>TPR</b>	<b>FPR</b>	<b>TNR</b>	<b>FNR</b>	<b>Accuracy</b>	<b>AUC</b>
Majority Voting	0.995	0.030	0.970	0.005	0.991	0.982
Maximum Probability	0.983	0.005	0.995	0.017	0.985	0.999
Product of Probabilities	0.999	0.005	0.995	0.001	0.998	0.981
<b>Performance Evaluation by combining J48, Random Forest, JRIP and PART</b>						
<b>Combination Rule</b>	<b>TPR</b>	<b>FPR</b>	<b>TNR</b>	<b>FNR</b>	<b>Accuracy</b>	<b>AUC</b>
Majority Voting	0.996	0.041	0.959	0.004	0.990	0.982
Maximum Probability	0.993	0.030	0.970	0.007	0.989	0.999
Product of Probabilities	0.994	0.025	0.975	0.006	0.991	0.978
<b>Performance Evaluation by combining J48, Random Forest and PART</b>						
<b>Combination Rule</b>	<b>TPR</b>	<b>FPR</b>	<b>TNR</b>	<b>FNR</b>	<b>Accuracy</b>	<b>AUC</b>
Majority Voting	0.996	0.023	0.977	0.004	0.993	0.982
Maximum Probability	0.993	0.043	0.957	0.007	0.987	0.999
Product of Probabilities	0.994	0.039	0.961	0.006	0.989	0.981

the original dataset into ten pieces and repeats the training and testing for ten times using holdout technique. Cross validation method reduce the chances of over-fitting the model. In experiment 2a, percentage split was used to divide the dataset into two pieces and 70% of the data was used for

training while the remaining 30% was used for testing. This method is very useful when the model needs to be used for predictions. The best results were obtained by the majority voting of output by Random Forest, PART and J48 classifiers. The TPR of 99.5% was observed using ensemble methods

TABLE V: Performance Evaluation of Experiment 2a - Detecting Unknowns

<b>Evaluation on ML classifiers on unknown dataset</b>						
ML Algorithm	TPR	FPR	TNR	FNR	Accuracy	AUC
Random Forest	1.000	0.029	0.971	0.000	0.975	0.998
PART	1.000	0.117	0.883	0.000	0.900	0.917
JRIP	1.000	0.043	0.957	0.000	0.964	0.979
Ensemble Methods	1.000	0.021	0.979	0.000	0.982	0.989

TABLE VI: Performance Evaluation of Experiment 2b - Detecting Unknowns

<b>Evaluation on ML classifiers on unknown dataset after interval of time</b>						
ML Algorithm	TPR	FPR	TNR	FNR	Accuracy	AUC
Random Forest	0.932	0.038	0.962	0.068	0.955	0.947
PART	0.926	0.112	0.888	0.074	0.897	0.893
JRIP	0.919	0.047	0.953	0.081	0.945	0.937
Ensemble Method	0.939	0.025	0.975	0.061	0.966	0.957

which was better than that of Random Forest. However, the FPR of 4.1% observed which is higher than Random Forest. It can also be seen that the best accuracy is seen by combining output of the five classifiers using the product of probabilities.

2) *Experiment 2 - Detecting Unknown:* This experiment was performed to evaluate the performance of ensemble ML classifiers on a new dataset that contains unknown instances from malicious samples and contains new traffic from different benign applications. We have limit this experiment to 3 ML Algorithms (Random Forest, PART and JRIP) as these ML algorithms produced the best results in individual evaluation. We have combined the output of these ML algorithms by majority vote method as shown in the Table V.

In Table V, performance evaluation of different classifiers can be seen. The RF performed best in individual classifiers with the highest TPR and lower FPR. The FPR produced by PART was significantly high. However, the ensemble classifier outperformed all the individual classifiers. In Table V, it can be seen that ensemble method performed better than RF in detecting unknown threats with the accuracy of 98.2% and the FPR was reduced to 2.1%.

This experiment showed that the ensemble methods are not only able to detect unknown threats but they are also good at identifying benign traffic. True Negative Rate (TNR) produced by these ensemble classifiers is also high which shows the efficiency of the classifier in distinguishing between normal and malicious instances.

Another experiment 2b (see Table V) was performed to check the performance of the classifiers after an interval of time. For that purpose, unseen traffic from some new malicious and benign applications was added to the test dataset. The performance of the individual classifiers decreased a bit with time due to the changes in the traffic patterns generated by the new malware samples. The ensemble methods increased the performance by combining the output of these individual classifiers. It can be clearly seen that the ensemble methods

produced the highest accuracy and AUC value by combining the output of these classifiers.

## VI. CONCLUSION

The ensemble methods used in this study were able to detect known and unknown threats. This study is the first step towards a more advanced ML based intrusion detection system. Ensemble methods not only produce better results but also reduce the chance of concept drift. Intrusion detection systems which rely only on ML techniques need frequent retraining. Otherwise, the decrease in TPR could be seen. In our previous studies, several experiments were performed to compare the ML model with antivirus vendors and we observed that ML classifiers were more efficient than some of the traditional antivirus. Furthermore, the efficiency of the ML classifiers was enhanced by using ensemble methods and these methods also helped with concept drift. Moreover, we have observed that the feature extraction and selection play an important role in the output of the classifier. Wrong features such as "IP Address" could over-fit the model and produce concept drift condition in the system.

The ensemble ML classifiers built were able to detect malicious traffic with a TPR of 99.9%, while the output from individual classifiers was observed between 94%-99.6%. As the ML classifiers are built for predictions, it is also important to evaluate the performance of the ML classifiers on new data. In this research, we have evaluated ML classifiers on unseen data and the accuracy of 98.2% was observed by ensemble methods while the accuracy observed by individual classifiers was between 90% - 97.5%. These results showed that the ensemble methods are more efficient than individual classifiers. Future work in progress aims to integrate the ML classifiers with traditional NIDS and to introduce some innovative methods in order to reduce the chance of concept drift.

## REFERENCES

- [1] SOPHOS, "When malware goes mobile," Tech. Rep., Access Date 26 Sep, 2017. [Online]. Available: <https://www.sophos.com/en-us/security-news-trends/security-trends/malware-goes-mobile.aspx>
- [2] "Malware detection and subscriber protection infographic," Alcatel-Lucent, Tech. Rep., Access Date 10 Sep, 2017. [Online]. Available: <https://www.alcatel-lucent.com/solutions/security-guardian-infographic>
- [3] K. Timm, "Strategies to reduce false positives and false negatives in nids," Tech. Rep., Access Date 10 Sep, 2017. [Online]. Available: <http://www.symantec.com/connect/articles/strategies-reduce-false-positives-and-false-negatives-nids>
- [4] K. Julisch and M. Dacier, "Mining intrusion detection alarms for actionable knowledge," in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD 02*. Association for Computing Machinery (ACM), 2002.
- [5] R. Koch, "Towards next-generation intrusion detection," in *Cyber Conflict (ICCC), 2011 3rd International Conference on*. IEEE, 2011, pp. 1–18.
- [6] J. A. Copeland III, "Flow-based detection of network intrusions," Feb. 27 2007, uS Patent 7,185,368.
- [7] D. Arp, M. Spreitzenbarth, M. Hübner, H. Gascon, K. Rieck, and C. Siemens, "Drebin: Effective and explainable detection of android malware in your pocket," in *Proceedings of the Annual Symposium on Network and Distributed System Security (NDSS)*, 2014.
- [8] M. Tavallae, E. Bagheri, W. Lu, and A.-A. Ghorbani, "A detailed analysis of the kdd cup 99 data set," in *Proceedings of the Second IEEE Symposium on Computational Intelligence for Security and Defence Applications 2009*, 2009.
- [9] J. McHugh, "Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by lincoln laboratory," *ACM transactions on Information and system Security*, vol. 3, no. 4, pp. 262–294, 2000.
- [10] S. Kumar, A. Viinikainen, and T. Hamalainen, "Machine learning classification model for network based intrusion detection system," in *Proc. 11th Int. Conf. for Internet Technology and Secured Transactions (ICITST)*, Dec. 2016, pp. 242–249.
- [11] S. Wang, Z. Chen, L. Zhang, Q. Yan, B. Yang, L. Peng, and Z. Jia, "Trafficav: An effective and explainable detection of mobile malware behavior using network traffic," in *Quality of Service (IWQoS), 2016 IEEE/ACM 24th International Symposium on*. IEEE, 2016, pp. 1–6.
- [12] F. A. Narudin, A. Feizollah, N. B. Anuar, and A. Gani, "Evaluation of machine learning classifiers for mobile malware detection," *Soft Comput*, nov 2014.
- [13] A. Feizollah, N. B. Anuar, R. Salleh, F. Amalina, R. R. Maarof, and S. Shamshirband, "A study of machine learning classifiers for anomaly-based mobile botnet detection," *Malaysian Journal of Computer Science*, vol. 26, no. 4, 2014.
- [14] X. Su, M. C. Chuah, and G. Tan, "Smartphone dual defense protection framework: Detecting malicious applications in android markets," in *Mobile Ad-hoc and Sensor Networks (MSN), 2012 Eighth International Conference on*. IEEE, 2012, pp. 153–160.
- [15] Y. Zhou and X. Jiang, "Dissecting android malware: Characterization and evolution," in *Security and Privacy (SP), 2012 IEEE Symposium on*. IEEE, 2012, pp. 95–109.
- [16] M. Soysal and E. G. Schmidt, "Machine learning algorithms for accurate flow-based network traffic classification: Evaluation and comparison," *Performance Evaluation*, vol. 67, no. 6, pp. 451–467, 2010.
- [17] A. Este, F. Gringoli, and L. Salgarelli, "Support vector machines for tcp traffic classification," *Computer Networks*, vol. 53, no. 14, pp. 2476–2490, 2009.
- [18] D. Zhao, I. Traore, B. Sayed, W. Lu, S. Saad, A. Ghorbani, and D. Garant, "Botnet detection based on traffic behavior analysis and flow intervals," *Computers & Security*, vol. 39, pp. 2–16, 2013.
- [19] M. Stevanovic and J. M. Pedersen, "An efficient flow-based botnet detection using supervised machine learning," in *Computing, Networking and Communications (ICNC), 2014 International Conference on*. IEEE, 2014, pp. 797–801.
- [20] C. Livadas, R. Walsh, D. Lapsley, and W. T. Strayer, "Using machine learning techniques to identify botnet traffic," in *Local Computer Networks, Proceedings 2006 31st IEEE Conference on*. IEEE, 2006, pp. 967–974.
- [21] L. I. Kuncheva, *Combining pattern classifiers: methods and algorithms*. John Wiley & Sons, 2004.
- [22] J. Kittler, M. Hatef, R. P. Duin, and J. Matas, "On combining classifiers," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 3, pp. 226–239, 1998.
- [23] R. Quinlan, "4.5: Programs for machine learning morgan kaufmann publishers inc," *San Francisco, USA*, 1993.
- [24] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [25] B. R. Gaines and P. Compton, "Induction of ripple-down rules applied to modeling large databases," *Journal of Intelligent Information Systems*, vol. 5, no. 3, pp. 211–228, 1995.
- [26] J. Fuernkranz and G. Widmer, "Incremental reduced error pruning," in *Proceedings of the 11th International Conference on Machine Learning (ML-94)*, 1994, pp. 70–77.
- [27] W. W. Cohen, "Fast effective rule induction," in *Proceedings of the twelfth international conference on machine learning*, 1995, pp. 115–123.
- [28] E. Frank and I. H. Witten, "Generating accurate rule sets without global optimization," 1998.
- [29] M. Walker, "Random forests algorithm," Tech.

- Rep., Accessed on 01.10.2014. [Online]. Available: <http://www.datasciencecentral.com/profiles/blogs/random-forests-algorithm>
- [30] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *ACM SIGKDD explorations newsletter*, vol. 11, no. 1, pp. 10–18, 2009.
- [31] M. Lindorfer, M. Neugschwandtner, L. Weichselbaum, Y. Fratantonio, V. v. d. Veen, and C. Platzer, "Andrubis – 1,000,000 Apps later: A view on current android malware behaviors," in *Proc. Third Int. Workshop Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS)*, Sep. 2014, pp. 3–17.
- [32] I. M. Digit Oktavianto, *Cuckoo Malware Analysis*. Packt Publishing, 2013.
- [33] B. H. Trammell and E. Boschi, "Bidirectional flow export using ip flow information export (ipfix) : Rfc-5103," IETF, Tech. Rep., Access Date 01 Jan, 2015. [Online]. Available: <https://tools.ietf.org/html/rfc5103.html>
- [34] F. J. Provost, T. Fawcett, and R. Kohavi, "The case against accuracy estimation for comparing induction algorithms." in *ICML*, vol. 98, 1998, pp. 445–453.

# Evaluation of AV Systems Against Modern Malware

Abidullah Zarghoon, Irfan Awan

Department of informatics  
University of Bradford  
Bradford, BD7 1DP

{A.Zarghoon, i.u.awan}@bradford.ac.uk

Jules Pagna Disso, Richard Dennis

Research and Innovation  
Nettitude Ltd

Leamington spa, CV31 3RZ

{jpagnadisso, rdennis}@nettitude.com

**Abstract**— Countering the proliferation of malware has been for recent years one of the top priorities for governments, businesses, critical infrastructure, and end users. Despite the apparent evolution of anti-virus (AV) systems, malicious authors have managed to create a sense of insecurity amongst computer users. Security controls do not appear to be sufficiently strong to stop malware proliferating. There seems to be a disconnect between public reports on AV tests and what people are experiencing on the daily basis. In this research, we are testing the efficiency of AV products and their ability to detect malicious files commonly known as malware. We manually generated payloads from five malware frameworks freely available to download and use. We use two modes of tests during our experiments. We manually installed a selection of AV systems in one first instance. We also use an online framework for testing malicious files. The findings in this study show that many antivirus systems were not able to achieve a higher score than 80% detection rate. Certain attack frameworks were much more successful in generating payloads that were not detectable by AV systems. We conclude that AV systems have their roles to play as they are the most common first line of defense, but more work is needed to successfully detect most malware the first day of their release.

**Keywords**—component; Malware, AV bypass, Antivirus Systems, Detection Techniques, Payloads, Antivirus Evaluation

## I. INTRODUCTION

Arguably, one of the most dominant threats to computer systems, networks and devices come from malware. Malicious software, known as Malware, can be a software, script used to gather information, damage data, interrupt systems' operation, access sensitive information or other malicious purposes [1].

According to AV-Comparatives, an independent antivirus evaluating organisation stated that detection rate of well-known commercial antivirus system ranges from 99.0% to 99.8% [2]. However, [3] has evaluated 10 common anti-virus systems against a set of ten thousand malware samples and the percentage of detection ranges from around 40% to 80%. The research will select five antivirus systems to manually test the payloads and will use an independent scanning platform known as NoDistribute to scan the payloads with over 35 antivirus systems and choose an attack method that represents the attack surface of recent attack. The objective of this paper is to evaluate the effectiveness of antivirus systems against modern malware using some attack methods commonly known to be used by either malicious actors or pen-testers. In this work, we have focused on Microsoft Windows Operating System.

## II. THE STATE OF ANTIVIRUS SYSTEMS

### A. Signature-based Detection

In signature-based detection, the antivirus scans files and compares files' signature against the known malware signatures in its database. If a malicious signature is found in any file, the file is either deleted, quarantined or cleaned depending on the antivirus system [4]. In signature-based detection, the antivirus software maintains a database of known traits that is keep up-to-date. Most AV companies now used a cloud model to keep their endpoints up-to-date. Several hours can be needed to create an effective signature when dealing with complex malware [5]; [6]. Manual processing of malware carries the risk of human error in signature creation [7]. Moreover, [8] argues that signature-based detection is becoming ineffective because AV companies are slow in extracting latest malware signatures given the increasing large number of malware sample available daily. Additionally, latest malware generation tools used to create new malware employ methods to avoid signature-based detection.

### B. Behaviour-based Detection

As opposed to signature-based detection, where the AV system detects a malicious code based on the byte patterns extracted from it by a malware researcher, the behaviour-based detection observes different programs for any malicious behaviour. In this detection technique, the AV systems look closely at various factors such as file creation patterns, malware hooking techniques, service creations, monitoring keystrokes, modification of hosts file in Windows system, creating autorun.inf files on network drives or removable media, or unpacking malicious code [9]; The malicious behaviours include disabling security controls, calling system APIs, changing Windows registry files, installing rootkits and others [10]. A significant advantage of behavioural based detection over signature based is that the level of packing, encryption, polymorphism or metamorphism employed by the malware does not, in most cases change its behaviour [9, 11].

Behaviour-based detection proves effective in detecting the different variations of a malware or possible detection of new malware; however, it comes with a drawback of increased false-positives. As compared to signature-based detection where known malware are detected with very little possibility of false-positives, due to the known signatures of the malware, the behaviour-based detection falls short because this technique searches for anomalous behaviour and can sometimes avoid a

benign application from execution. Furthermore, another drawback of behaviour-based detection is that malware can evade detection by behaving like a normal application to appear legitimate [7]. Behaviour-based detection is a rather broad concept that consists of different techniques including heuristic detection.

### C. Heuristic Detection

Heuristic detection employs several techniques to analyse the characteristics and behaviour of files. Heuristic analysis tends to combine both static and dynamic analysis. Whilst there is potential to detect more malware, heuristic analysis can generate false positives when using static heuristic. The dynamic heuristic analysis, on the other hand, proves to generate less false positives compared to the static analysis because in dynamic analysis the file performs its intended action in virtual or isolated environment.

## III. TEST METHODOLOGY

The tools used to generate payloads during our experiments are found in the popular Kali distribution namely the Metasploit framework, Veil, Shellter and FatRat.

In this study, we aim to achieve few objectives. Public program or frameworks such as av-test.org use malware that are up to 4 weeks old to test detection rates of antivirus. These tests show that the detections rates are very high, indicating that most virus are detected. Motivated by the number of security breaches observed in recent months and years, we use free to download and use software widely available to anyone on the Internet to challenge the claim by av-test.org. Are antivirus systems able to achieve 90% of detection if they are tested with malware freshly compiled?

A secondary objective of this study is to raise awareness that whilst antivirus have an important role to play in protecting endpoints, they can also be bypassed.

The difference between signature based and heuristic detection is theoretically interesting. However, in this study, we manually scan malicious files at rest to simulate signature and static heuristic detection; we then execute the files to simulate the dynamic heuristic detection. Our results show that some malware were only detected when executed. At rest, whilst some malware were able to bypass detection but this does not necessarily translate to a successful compromise of the endpoint if the malware is executed.

Our experiments are executed in three phases. We use the NODISTRIBUTE malware platform to access a wide range of antivirus. During the test, we have noticed the number of antivirus changing from 37 to 38. This change was out of our control. In the second phase of our test, we manually create malware to test them at rest. In the last round of test, we use execute our malware to test heuristic detection.

### A. Metasploit

Metasploit is one of the most common and well-known frameworks in information security community. It is also very well-reputed and is considered a must have a weapon in a penetration tester's arsenal.

### B. Veil

Veil is a framework that consists of two tools namely Veil-Evasion and Veil-Ordnance that generates payloads. The creators of this framework argue that payloads generated by this framework bypass common antivirus systems [12]. The whole project is currently under support by Christopher Truncer. In a recent update of February 2016, Christopher Truncer states that other functionalities such as obfuscation and random key generation have been added to the framework to avoid antivirus systems. Veil-Evasion has a tool called PyInstaller, that is used to convert the Python code into an executable file. The new PyInstaller or Pysinstaller 3.1 has an added ability of encrypting the bytecode that this installer outputs [13].

### C. TheFatRat

TheFatRat is another tool in penetration tester's arsenal that generates back-doors and post-exploitation attacks. It compiles malware with payloads that can be executed on Windows, Mac or Android platforms. This tool is created by Edo Maland (Screetsec).

### D. AVOIDS

Avoidz is a tool that is used to generate encoded PowerShell with Metasploit payload and convert C and C# templates to EXE files. This tool is dependent on the Metasploit framework, Xtrem, MinGW32 and MonoDevelop. The tool was created by Mascerano Bachir. This tool is a utility that wraps other tools into one. For instance, the generation of payload is achieved through MSFvenom found in Metasploit. One of the encoder implements a polymorphic XOR additive feedback encoder. The decoder stub is generated based on dynamic instruction substitution and dynamic block ordering. The tool was created by Mascerano Bachir.

### E. Shellter

Shellter is a dynamic shellcode injection tool. It is claimed that this tool is probably the first dynamic PE infector to be ever created. It can inject shellcode into the native applications of Windows platform. The shellcode used with this framework can be either user-generated or generated by a framework. To prevent antivirus detection, this tool does not infect or modify the PE file in a way, which will seem suspicious to the antivirus systems, for example, changing memory access permissions. Shellter is compatible with both 64-bit and 32-bit Windows operating systems. This tool is not dependent on other dependencies such as python and .NET [14]. Furthermore, the automatic mode of Shellter makes the process easier for anyone who is using it [15].

## IV. ANTIVIRUS SYSTEMS

During our experiments, we used two sets of antivirus (AV) systems. We manually installed and configured 5 AVs and we also use the NODISTRIBUTE online platform to access over 35 AVs already configured with the lasted signatures. The antivirus were selected as a result of a quick survey amongst students to understand what AV systems they use.

A. Manually Configured Antivirus

The following AV products were manually installed and configure for testing using their latest version as of July 2017. The table below show the results for the test performed by av-test.org in July 2017. However, Panda AV scores are from the tests performed in October 2016.

AntiVirus	0-day protection rate	Widespread / prevalent malware rate
Kaspersky	100%	100%
Panda	100%	99.5%
Avast	100%	99.9%
AVG	100%	99.9%
McAfee	99%	100%

B. NODISTRIBUTE

Nodistribute is public website that offers to scan files using over 30 antivirus products with their latest signatures. This platform was used to allow us to compare more than 30 AV products.

V. EVALUATION AND RESULTS

A. Lab set up

Our experimental lab was made of one physical machine (the host) and two virtual machines, one for the attacker and one for the victim. We used 6 snapshots for our victim system. On 5 snapshots we installed one AV each. One snapshot did not have AV installed to validate the malicious files.

HOST SYSTEM		ATTACKER'S SYSTEM		VICTIM'S SYSTEM	
Computer System	Dell Inspiron 5459	Computer System	Virtual System	Computer System	Virtual System
Operating System	Windows 10 x64	Operating System	Kali Linux x32	Operating System	Windows 7 x32
Processor	Core i7 – 2.50 GHz (4 CPUs)	Processor	2 processors, 2 cores (4 cores)	Processor	2 processors, 2 cores (4 cores)
RAM	12 GB	RAM	4 GB	RAM	2 GB
Storage	1000 GB	Storage	60 GB	Storage	40 GB
IP Address	Not Applicable*	IP Address	192.168.0.2	IP Address	192.168.0.3

Table 1: Testbed

B. Test methodology

1) Method 1 – Nodistribute platform

Using the NODISTRIBUTE platform, we have generated 15 malicious files on our attacker systems. We then uploaded the files one by one to the online platform and recorded the results. The advantage of this method is that the malicious files that we generated were tested by an independent, public and well known-platform to security professionals.

2) Method 2 – manual testing

We started by building the payloads. We used the reverse\_tcp and the reverse\_https during our test. Once the payload generated, we tested them by executing on the victim machine to confirm that the files were malicious. We then proceeded with scanning the files against the different antivirus. The process is summarized in figure 1.



Figure 1: Manual test procedure

C. Results

1) NODISTRIBUTE – detection rate per testing option

In the first set of results, we present how often a malicious file was detected. From the results, the malicious file generated by Shellter has the least detections followed by Vernom and FatRat Powershell Batch. Results are shown in table 2 and figure 2.

Malicious file	AV detection	AV used	Detection rate
FatRat Powershell Batch	2	37	5.4%
Simple FatRat Payload	23	37	62.2%
FatRat DOC Macro	12	38	31.6%
FatRat Powerstager	12	38	31.6%
Metasploit PDF	21	38	55.3%
Metasploit DOC Macro	15	38	39.5%
Shellter File	1	38	2.6%
Veil XLS Macro	8	38	21.1%
Veil and MSF file	5	38	13.2%
Veil Ordnance AES Encrypt	5	38	13.2%
Veil Rev HTTPS Payload	12	38	31.6%
Veil Simple Payload	24	38	63.2%
Venom File	1	37	2.7%
WinPayloads File	15	37	40.5%
AvoidZ Payload	18	37	48.6%

Table 2: AV scanning results for NODISTRIBUTE

NODISTRIBUTE detection rate per tool option used.

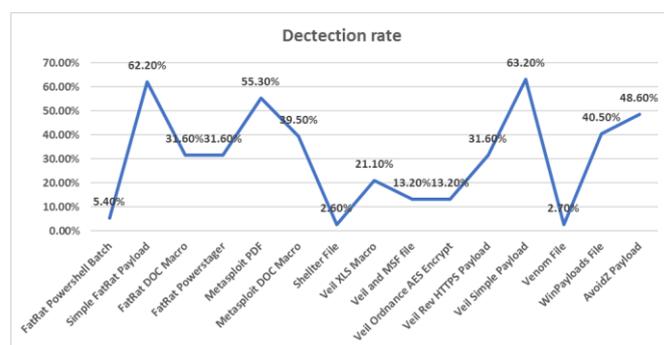


Figure 2: Detection rate per testing option

2) *NODISTRIBUTE Antivirus detection rate results*

We generated 15 files which were uploaded to the NODISTRIBUTE platform. The results show how many files each antivirus was able to detect as malicious. We manually confirmed that all files were malicious.

<b>Antivirus</b>	<b>Files Detected / 15</b>	<b>Detection rate</b>
A-Squared	7	46.7%
AVG Free	9	60.0%
Ad-Aware	9	60.0%
AhnLab V3 Internet Security	4	26.7%
Arcavir Antivirus 2014	1	6.7%
Avast	12	80.0%
Avira	6	40.0%
BitDefender	10	66.7%
BullGuard	11	73.3%
Clam Antivirus	1	6.7%
Comodo Internet Security	1	6.7%
SET NOD32	6	40.0%
F-PROT Antiviurs	3	20.0%
F-Secure Internet Security	6	40.0%
FortiClient	8	53.3%
G Data	9	60.0%
IKARUS Security	10	66.7%
Jiangmin Antivirus 2011	4	26.7%
K7 Ultimate	4	26.7%
Kaspersky Antivirus	11	73.3%
MS Security Essentials	6	40.0%
Malwarebytes Anti-Malware	0	0.0%
McAfee	3	20.0%
NANO Antivirus	2	13.3%
Norton Antivirus	2	13.3%
Outpost Antivirus Pro	0	0.0%
Panda Security	1	6.7%
Quick Heal Antivirus	2	13.3%
SUPERAntiSpyware	0	0.0%

Solo Antivirus	0	0.0%
Sophos	0	0.0%
TrustPort Antivirus	10	66.7%
Twister Antivirus	1	6.7%
VBA32 Antivirus	1	6.7%
VirIT eXplorer	1	6.7%
Zillya! Internet Security	2	13.3%
eScan Antivirus	7	46.7%

Table 3: Performance evaluation of tools against antivirus

Our results are significantly different from those achieved during the tests by av-test.org. One of the major reasons behind the difference is that av-test.org used malware that have been in circulation four weeks prior to the test date. There is a higher chance for a malicious file to be detected during that period as compare to a file that is a newly created. The average detection rate form av-test.org is higher than 99% whilst our detection rate is just over 30%.

3) *Manual scanning – detection rates*

In our test, we have used 5 antiviruses as shown in the table below. Each of the antivirus was installed in each system. The 15 files were loaded using a share folder between the host and the virtual environment.

<b>List of antiviruses manually tested</b>	<b>Version</b>
Avast Internet Security	17.6.2310
Avast Internet Security	18.0.0.405
Kaspersky Total Security	14.0.9029
McAfee Antivirus LiveSafe	17.6.3029
Panda Global Protection	17.0.1

Table 4: List of antiviruses

In this test, we observed that heuristic detection has an impact of the detection of malicious files.

<b>Antivirus</b>	<b>Detection during scanning / 15</b>	<b>Additional detection at Execution / 15</b>	<b>Total detection / 15</b>	<b>rate</b>
Avast	12	0	12	80.0
Kaspersky	12	2	14	93.3
AVG	12	0	12	80.0
McAfee	5	1	6	40.0
Panda	4	5	9	60.0

Table 5: Manual testing results

The results in table 5 show a clear distinction between statistic detection and dynamic detection.

Panda antivirus has shown great capabilities in detecting files when they are executed. This would translate to Panda AV potentially having a good dynamic heuristic detection when compared to other antivirus systems tested. AVG, despite being a free AV was able to detect 80% of the malware. Hypothetically, Panda would have detected more malware if it had incorporated the free AVG engine.

It was also interesting to note that paid antivirus does not necessarily translate to better protection.

Some of the AV did not detect the malicious files until they were executed. In order to read Table 6, we use the following legend: D=detected; BaE=Blocked at execution; B=bypassed

Files	Avast	Kasp	McAfe	AVG	Panda
FatRat Powershell Batch	D	BaE	B	D	B
Simple FatRat Payload	D	D	D	D	D
FatRat DOC Macro	D	D	BaE	D	D
FatRat Powerstager	D	BaE	B	D	BaE
Metasploit PDF	D	D	D	D	B
Metasploit DOC Macro	D	D	D	D	B
Shellter File	B	D	B	B	BaE
Veil XLS Macro	D	D	B	D	B
Veil and MSF file	D	BaE	B	D	BaE
Veil Ordnance AES Encryp	D	B	B	D	BaE
Veil Rev HTTPS Payload	D	D	B	D	B
Veil Simple Payload	D	D	D	D	BaE
Venom File	B	D	B	B	B
WinPayloads File	B	D	B	B	D
AvoidZ Payload	D	D	D	D	D

Table 6: static and heuristic detection analysis

This research has referenced previous evaluation of antivirus systems. However, there is not much details about the depth of the tests. Two categories are generally considered: detection against zero days and detection of know malware. Also, there is much talk in security communities that antivirus systems are not efficient enough in detecting new malware. This research as allowed us to set simple but efficient criteria to evaluate antivirus systems. There are few lessons learnt from these experiments:

- The heuristic analysis claimed by AV companies can be challenged. The detection rate of some AV

remained the same whether the file was analysed at rest or dynamically.

- In some cases, the detection rate was better when the files were executed. At rest, the detection rate of manual analysis was 60% compared to 70.7% when executed.
- The average detection rate using the NoDistribute platform was very low. Only 30.6%
- The method or tools used to test antivirus can greatly influence the final results. If the tests performed by only using Shellter or Venom, the conclusion would that antivirus can only detect less than 5% of attacks. It is therefore important that any statistics reveals the full extent of the method used so that results can be considered in their original context.

- [1] Milošević N. 2013. History of malware. ArXiv Prepr. ArXiv13025392.
- [2] Comparative A-V. 2016. File Detection of Malicious Software.
- [3] Mirza QKA, Mohi-Ud-Din G, Awan I. A cloud-based energy efficient system for enhancing the detection and prevention of modern malware. In: Adv. Inf. Netw. Appl. AINA 2016 IEEE 30th Int. Conf. On. IEEE, pp 754–761.
- [4] Kakad AR, Kamble SG, Bhuvad SS, Malavade VN. 2014. Study and Comparison of Virus Detection Techniques. Int. J. Adv. Res. Comput. Sci. Softw. Eng. ISSN 2277.
- [5] Thengade A, Khaire A, Mitra D, Goyal A. 2014. Virus Detection Techniques and Their Limitations. Int. J. Sci. Eng. Res. 5:.
- [6] Al Amro S, Alkhalifah A. 2015. A Comparative Study of Virus Detection Techniques. World Acad Sci Eng Technol Int J Comput Electr Autom Control Inf Eng 9(6):1566–1573.
- [7] Galal HS, Mahdy YB, Atiea MA. 2016. Behavior-based features model for malware detection. J Comput Virol Hacking Tech 12(2):59–67.
- [8] Del Grosso N. 2002. It's Time to Rethink your Corporate Malware Strategy. Retrieved At :12.
- [9] Bazrafshan Z, Hashemi H, Fard SMH, Hamzeh A. 2013. A survey on heuristic malware detection techniques. pp 113–120.
- [10] Cloonan J. 2017. Advanced Malware Detection - Signatures vs. Behavior Analysis. .
- [11] Jacob G, Debar H, Filiol E. 2008. Behavioral detection of malware: from a survey towards an established taxonomy. J Comput Virol 4(3):251–266.
- [12] Truncer C, HarmJoy, Wright M. 2016. Veil-Evasion - Veil - Framework. Available from: <https://www.veil-framework.com/framework/veil-evasion/>, .
- [13] Truncer C. 2016. Veil-Evasion Archives - Veil - Framework. Available from: <https://www.veil-framework.com/category/veil-framework/>, .
- [14] kyREcon. 2017. Shellter | Shellter. Available from: <https://www.shellterproject.com/introducing-shellter/>, .
- [15] Dieterle D. 2015. Anti-Virus Bypass with Shellter 5.1 on Kali Linux

# *Security Enhancements to TLS For Improved National Control*

Lamya Alqaydi, Chan Yeob Yeun, Ernesto Damiani

Khalifa University of Science and Technology, Information Security Research Center

Abu Dhabi Campus P.O.Box 127788, Abu Dhabi, UAE

{100035547, cyeun, ernesto.damiani}@kustar.ac.ae

*Abstract— Establishing a secure connection is a must nowadays since a lot of transactions are being done online. For instance, it can vary from buying items from small shops into buying extremely expensive equipment online. Hence, the need for securing the sessions and e-commerce is highly required. Furthermore, any government entity will require its communication to be secured from eavesdropping and Man in The Middle (MITM) attacks. Web Threats are spreading around the world and becoming more aggressive every year. Secure Socket Layer (SSL) and Transport Layer Security (TLS) were established as a standard to encrypt the communication between the client and the server. Everyday new vulnerabilities and loopholes in the internet protocols are being discovered. Hence an enhancement to the TLS protocol is a must. This need rises since most transactions and confidential communication is done through the network. Then information security researchers and developers have to test, develop, verify and enhance the security of the network. Throughout this report a study of the TLS protocol will be highlighted and its relative security. Later, a proposed method to test different TLS protocols will be explained. Future work will include developing a prototype that will be used to test different TLS protocol versions.*

**Keywords—***TLS, SSL, Privacy, Security, TLS Handshake, Renegotiation attack, BEAST attack, CRIME and BREACH attack, Heartbleed Attack*

## I. INTRODUCTION

Public key infrastructure (PKI) was designed to relate certain public keys with the corresponding user. It also allows another user to trust that certificate which was issued and included in the SSL certificate. PKI includes the certificate authorities (CA) that can verify the identity of user or server and whether their certificate is valid or not [1] [2] [3].

Secure Socket Layer (SSL) was established as a standard to encrypt the communication between the client and the server [1] [2]. Further, Transport Layer Security (TLS) protocol was established to secure the communication between the clients' browser and the web-server. Additionally, the server's public-key is being used to initiate symmetric session keys. This is done so that the client browser can verify the web-server public-key. These keys must be certified by a Certificate Authority (CA). Nowadays, the focus of attacks to breach the system is by exploiting the weakness in creating and verifying the certificates [4] [5] [6].

Any web browser can trust a huge number of root CAs. As an example, if a certificate Authority was compromised then it can help the attackers to create multiple fake certificates and can cause a breach for certain data depending on the type of

certificate that was compromised. Additionally, if a web server was affected it can also lead to lowering the trust of the public-key infrastructure (PKI) [7].

CIA is an acronym used in information security world that stands for confidentiality, integrity and availability of the data at any time. Confidentiality is making sure that data are secured and protected against eavesdroppers in addition to encrypting the whole session, so it is harder for attacker to understand. While the integrity deals with making sure that data can't be altered, and no modification of data is occurring. Finally, Availability means making data available to the user all time.

There are many ways that can be used to provide security by adding the cryptography concept in the system. RSA (Rivest-Shamir-Adleman) crypto-system which is a public key encryption mechanism. It is being used to secure sensitive data over insecure network. Some of the applications that require security to be applied are the following:[8][9][10]

- 1- encryption of the e-mail in high sensitive areas such as governmental agencies.
- 2- web-site encryption to secure financial transactions as an example.
- 3- Encryption of Applications that are used to communicate securely.
- 4- Digital Signatures to preserve the right of the person who wrote certain document, app, website ... etc.
- 5- Using the latest version of SSL/TLS in communication to prevent attackers from being able to access and eavesdrop the session.

Accordingly, the main motivation of this paper is to suggest some security enhancements to the TLS protocol for improved national control. Where we seek to find and build a prototype of an automatic monitoring tool which could identify the major and minor versions of the TLS protocol used by an external physical machine or even a virtual one, coupled with a mechanism to test and notify the vendor or the owner of web server with the security problems in their system. Section II of the paper will be highlighting related work and literature review of the SSL/TLS protocol. Later on in section III an analysis of different TLS upgrades used in different web browsers was done. Finally, section IV will have a conclusion of the work discussed in this paper and the possible future work that will be done.

II. RELATED WORK

A. The need for secure channel of communication

The TLS protocol has been in active use since 1999. Since then the protocol underwent several changes in implementation. The latest of these was TLS 1.3 published in April 2017. The current state of the protocol is described and explored in this report. It has played a major role in the current internet protocols infrastructure. However, as recent breaches such as the following: Freak attack, Poodle attack, Breach attack, Heartbleed attack and Beast attack have shown that the protocol itself leaves much to be desired [10][11].

As more of the critical services of governments and major companies become largely internet-based, ensuring the security of servers relying on the TLS protocol and its varied implementations become more important. Many companies are still using the older versions of TLS that has been considered as broken due to the legacy devices that they currently using. Hence the issue is exacerbated by the multitude of older versions still in use and many known vulnerabilities each of which is relevant for a subset of versions and implementations. Netscape formed the SSL protocol in order for it to be used in e-commerce and allows having security while doing financial transactions. This was critically needed to protect consumer’s personal data and their credit cards information. In order for this to be establish SSL protocol was implemented at the application layer of the OSI model. Then the IETF standardization entity named it as TLS which is the transport layer security once the SSL protocol was considered as a standard [12][13]. Figure 1 illustrates how SSL/TLS protocol can be found in the OSI model.

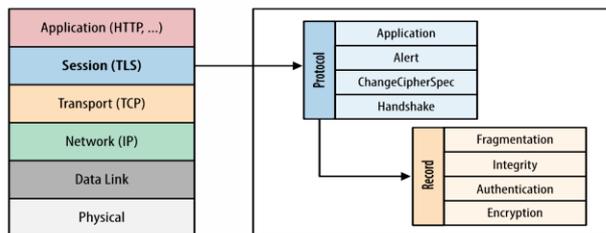


Figure 1: SSL and TLS protocols in the OSI model [1]

At the beginning, SSL v2.0 was the released version of the protocol to the public in 1994. Later on, it was replaced by SSL v3.0 because of the newly discovered flaws in the protocol. Some of the flaws that existed in SSL v2.0 are that it has a weak MAC construction in addition it can force parties to use weaker encryption and there was no protection for the handshakes. Then in 1999 the IETF worked hard in making the protocol as a standard and as a result they formed RFC 2246 that was published and became known as TLS 1.0. the work of IETF to make the TLS as better security protocol continued and in 2006 they have published TLS 1.1 which is RFC 2246 as the new version of the protocol. Whenever a new flaw or a need to add more capabilities to the protocol arises a new version is published. After two years, a new version of TLS which is TLS 1.2 was introduced to be used [14] [15] [16].

Further, TLS protocol as stated earlier was mainly configured to provide the CIA triad. The existing handshake in the TLS protocol allows the users to authenticate their identities. Each party in the communication channel can verify the other. This kind of verification is based on chain of trust. So, in this case both the client and the server can authenticate each other and make sure that they are who they claim to be. TLS protocol signs each message being sent with a message authentication code called (MAC). Once a TLS record is being sent it automatically have a MAC code generated and added on the message. If the receiver wants to check the integrity and authenticity of the received message it can compute and verify the MAC address of it. The more dependency on the internet means more increase of the need to secure the communication. People can range from regular users into application developers who need to make sure that they are protecting themselves by enabling HTTPS everywhere [17] [18].

The SSL/TLS comprises of four fundamental protocols which are:

- 1- Handshake  
It has the negotiation sessions, Cryptographic methods and authentication (One way or two ways)
- 2- Record  
It provides the other party with an optional payload compression in addition to shared transaction layer. Finally, it also makes sure that confidentiality of data is kept and checks the integrity of the payload being transmitted.
- 3- Cipher Spec  
Tells the server/client to use the parameters specified previously
- 4- Alert  
It reports any error occurring during the communication.

At the beginning the client and the server has to agree on the version of TLS protocol, cipher suite and check if the certificates are valid and then it will be used to initiate and encrypt the communication tunnel. After that the client and the server can start exchanging application data over the TLS link. Then the TLS handshake will initiate and start the communication between the client “Sender” and the server “Receiver” as it can be seen in the next Figure [18] [19] [20]. Next an overview of how SSL or TLS handshake can happen will be shown in Figure 2.

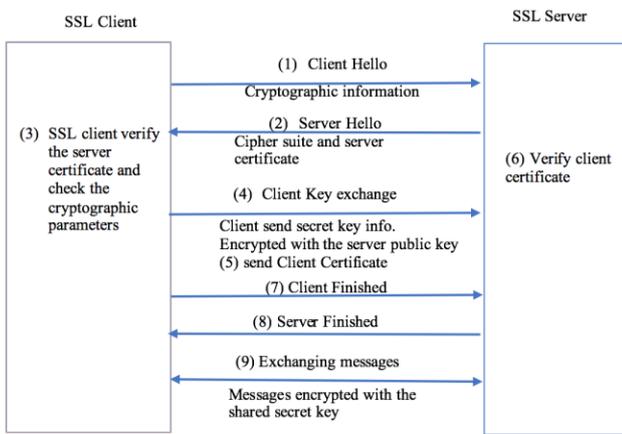


Figure 2: SSL/TLS handshake [20][21]

The following steps are a summary of how the SSL/TLS handshake happens:

Step 1: The SSL/TLS client sends “Client Hello” message that includes the information about TLS/SSL version and the cipher suites that are supported by the client.

Step 2: The SSL/TLS server responds with a “server Hello” that has the cipher suite selected by the server. It also contains session ID in addition to that the server sends its digital certificate to the client.

Step 3: The SSL/TLS client verify the servers’ digital certificate by the server public key.

Step 4: The SSL/TLS client then will send a randomly generated byte that will be used to find the secret key which will be used to encrypt the coming conversations. This randomly generated byte will be encrypted by the public key of the server and the server can decrypt it using its private key.

Step 5: if the SSL/TLS server requested from the client a certificate request. Step 4 will be done in addition to the clients’ digital certificate that will be sent to the server. And in this stage the server can be able to verify the identity of the client.

Step 6: The SSL/TLS server will verify the client certificate.

Step 7: The SSL/TLS client sends “client finished” that is encrypted by the previously generated secret key. This will be an indication to the server that the client finished its part in the handshake.

Step 8: The SSL/TLS server sends “server finished” message to indicate that the server handshake is finished as well.

Step 9: Throughout the session that was initiated once the handshake started both the server and client were able to send and receive messages that uses symmetric encryption mechanism.

Despite that there were major hacks and known vulnerabilities. The knowledge of these issues is important as the implementation of new protocol versions can take time on different operating systems and not all web-connected servers will or can upgrade their version of TLS. The next section goes into some details of the known vulnerabilities of each of the versions of TLS that have been released so far [22] [23].

Additionally, the main objective of my project is to design and build an automatic monitoring tool which could identify the major and minor versions of the TLS protocol used by an external physical machine or even a virtual one. This tool could be offered as a subscription service for subscribers to know the vulnerabilities their system has and proactively upgrade their system or otherwise work to mitigate the influence of the vulnerabilities. Further this tool will try to identify the person of charge of the domain and maybe be able to send them a notification that they have a security problem in their site.

A secondary objective of the project is the testing of the latest TLS version using current and perhaps soon-to-be-released test suites. Any vulnerability found that way would be added to the documented list.

*B. The Vulnerabilities and upgrades of TLS*

Figure 3 illustrates a timeline of the SSL/TLS vulnerabilities and attacks related to it.



Figure 3: TLS timeline and Vulnerabilities [17][18]

In this paper, I will illustrate more on some of the major attacks shown in the previous figure.

1- Renegotiation Attack

The following steps can be used by the attacker to hijack the session and pretend to be a legitimate client [26] [11] [9].

- Step 1: The attacker initiates a session with the server.
- Step 2: It hijacks the connection between the server and the client and collects the TLS handshake packets sent from the client-side.
- Step 3: Attacker sends its own TLS handshake packets to the server.
- Step 4: Then the attacker initiate the renegotiation request with the server.
- Step 5: Attacker now uses the client initial handshake to start communicating with the server.
- Step 6: At this stage, the attacker can start attacks on the application layer protocol and aim to attack the HTTPS protocol when it acts as a proxy in the system.

Figure 4 will show how this attack can happen:

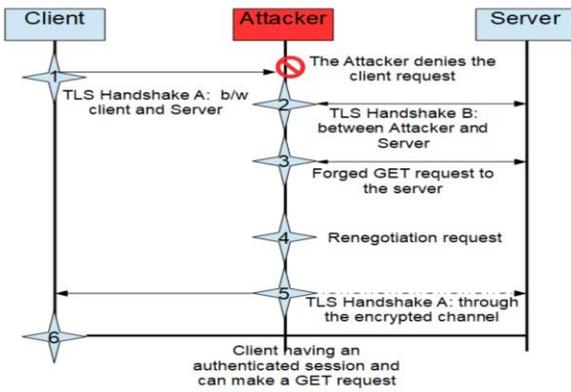


Figure 4: Renegotiation Attack on TLS

2- BEAST Attack

It is an abbreviation of (Browser Exploit against SSL/TLS). It was using a vulnerability found in the Cipher Block chaining (CBC) mode that was used the TLS 1.0. A chosen plaintext attack can be established since the attacker can predict the IV (initialization vector) easily and eventually he/she will be able to decrypt the information [24] [25] [26].

BEAST Attack scenario can be summarized in the following steps:

- a- Charlie tries to force Alice to send NULL [0-14] L [0] to Bob.
  - b- It eavesdrops the communication between Alice and Bob after that it gained the ciphertext  $CL = E(\text{key}, CL - 1 \oplus \text{NULL}[0-14] \ L[0])$
  - c- It assumes H as a guess of L[0]
  - d- Then it forces Alice to send the plain text  $C_{i-1} \oplus CL - 1 \oplus \text{NULL}[0-14] \ H$
  - e- Alice is being forced to send  $C_i = E(\text{key}, C_{i-1} \oplus CL - 1 \oplus \text{NULL}[0-14] \ H)$
- The result will be  $C_i = E(\text{key}, CL - 1 \oplus \text{NULL} [0-14] \ H)$
- f- Then it will be able to check if the value of  $C_i$  is equivalent to  $CL$  then  $L[0]=H$

Next Figure 5 will illustrate how BEAST attack can happen:

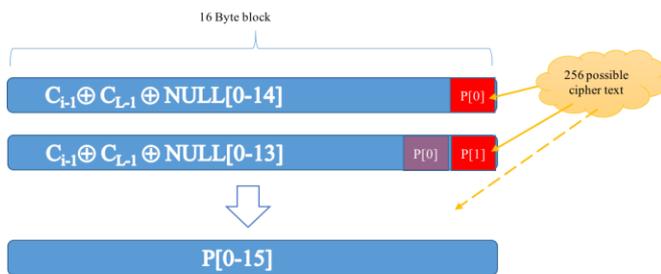


Figure 5: BEAST attack scenario [27][28]

3- Heartbleed Attack

It exists as vulnerability in the openSSL cryptographic library. It allows the attacker to steal confidential data despite the usage of SSL/TLS protocol. It provides anyone the power to have access to the memory of the system that was protected previously by the openSSL. This vulnerability breaks the confidentiality of the secret keys used to encrypt the entire traffic. Additionally, the attacker will be able to eavesdrop the entire communication between the server and the client and might have the ability to impersonate the user and get high confidential data [26] [27] [28].

Heartbeat Normal usage will be illustrated in Figure 6:

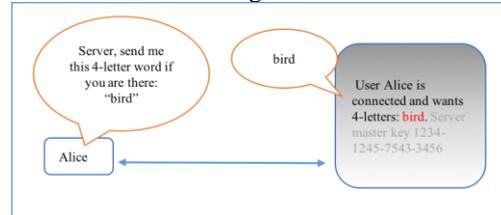


Figure 6 : Heartbeat normal usage

While Heartbeat Malicious usage (Heartbleed Attack) can happen as illustrated in Figure 7:

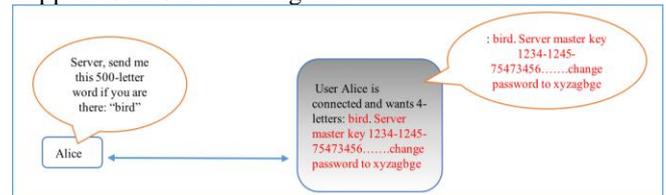


Figure 7: Malicious usage of Heartbeat that lead to discovery of SSL/TLS Heartbleed attack

III. ANALYSIS OF DIFFERENT VERSIONS OF TLS UPGRADES

A. Cipher Security against possible known attacks on different TLS versions

The following table illustrates the Cipher Security against possible known attacks on different TLS versions

Table 1: Cipher security Against possible known attacks on different TLS versions [20] [21] [22] [23] [24] [25] [26]

Type	Algorithm	Nominal strength (bits)	Protocol version					Status	
			SSL 2.0	SSL 3.0	TLS 1.0	TLS 1.1	TLS 1.2		
Block cipher with mode of operation	AES GCM	256, 128	-	-	-	-	Secure	Secure	Defined for TLS 1.2 in RFCs
	AES CCM		-	-	-	-	Secure	Secure	
	AES CBC		-	-	Depends on mitigations	Secure	Secure	-	
	Camellia GCM	256, 128	-	-	-	-	Secure	Secure	
	Camellia CBC		-	-	Depends on mitigations	Secure	Secure	-	
	ARIA GCM	256, 128	-	-	-	-	Secure	Secure	
	ARIA CBC		-	-	Depends on mitigations	Secure	Secure	-	
	SEED CBC	128	-	-	Depends on mitigations	Secure	Secure	-	
	3DES EDE CBC	112	Not Secure	Not Secure	Not Secure	Not Secure	Not Secure	-	
	GOST 28147-89CNT	256	-	-	Not Secure	Not Secure	Not Secure	-	
	IDEA CBC	128	Not Secure	Not Secure	Not Secure	Not Secure	-	-	
	DES CBC	56	Not Secure	Not Secure	Not Secure	Not Secure	-	-	
		40	Not Secure	Not Secure	Not Secure	-	-	-	
RC2 CBC	40	Not Secure	Not Secure	Not Secure	-	-	-		
	40	Not Secure	Not Secure	Not Secure	-	-	-		
Stream cipher	ChaCha20-Poly1305	256	-	-	-	-	Secure	Secure	Defined for TLS 1.2 in RFC
	RC4	128	Not Secure	Not Secure	Not Secure	Not Secure	Not Secure	-	Prohibited in all versions of TLS by RFC 7465
		40	Not Secure	Not Secure	Not Secure	-	-	-	

B. TLS/SSL support history of web browsers

The following table is showing the TLS/SSL support of different web browsers with respect to different parameters :

network and hence give them the opportunity to fix it before any critical situation can occur. A lot of businesses could be saved, and confidential data kept secret.

Table 2: SSL/TLS support in different web browsers [29][30][31][32][33][34]

Browser	Version		Platforms	SSL protocols		TLS protocols				Vulnerabilities fixed					Protocol selection by user	
				SSL 2.0 (insecure)	SSL 3.0 (insecure)	TLS 1.0	TLS 1.1	TLS 1.2	TLS 1.3 (proposed)	BEAST	CRIME	POODLE (SSLv3)	RC4	FREAK		Logjam
Google Chrome (Chrome for Android)	54-60	61	Windows (7+) OS X (10.9+) Linux Android (4.1+) iOS (9.0+) Chrome OS	No	No	Yes	Yes	Yes	Disabled by default (Experimental)	Not affected	Mitigated	Not affected	Disabled by default	Mitigated	Mitigated	Temporary
Google Android OS Browser	Android 8.0			No	No	Yes	Yes	Yes	Unknown	Unknown	Unknown	Not affected	Disabled by default	Mitigated	Mitigated	No
Mozilla Firefox (Firefox for mobile)	55		Windows (7+) OS X (10.9+) Linux Android (4.0.3+) iOS (9.0+) Firefox OS Maemo  ESR only for: Windows (XP SP2+) OS X (10.9+) Linux	No	No	Yes	Yes	Yes	Disabled by default (Experimental)	Not affected	Mitigated	Not affected	Disabled by default	Not affected	Mitigated	Yes
Microsoft Internet Explorer	11		Windows 8.1 Server 2012 R2	Disabled by default	Disabled by default	Yes	Yes	Yes	No	Mitigated	Not affected	Mitigated	Disabled by default	Mitigated	Mitigated	Yes
Opera Browser (Opera Mobile) (Pre-Presto and Presto)	12.18		Windows OS X Linux Android Symbian S60 Maemo Windows Mobile	No	Disabled by default	Yes	Yes	Yes	No	Mitigated	Not affected	Mitigated	Disabled by default	Mitigated	Mitigated	Yes
Opera Browser (Opera Mobile) (WebKit and Blink)	41-46	47	Windows (7+) OS X (10.9+) Linux Android (4.0+)	No	No	Yes	Yes	Yes	Disabled by default (Experimental)	Not affected	Mitigated	Not affected	Disabled by default	Mitigated	Mitigated	Temporary
Apple Safari	11		macOS 10.13	No	No	Yes	Yes	Yes	Unknown	Mitigated	Not affected	Not affected	Disabled by default	Mitigated	Mitigated	No
Apple Safari (mobile)	11		iOS 11	No	No	Yes	Yes	Yes	Unknown	Mitigated	Not affected	Not affected	Disabled by default	Mitigated	Mitigated	No

In Table 2; I have presented the recent versions of different browsers with respect to the support of different TLS versions. In addition, whether an associated vulnerability with that it version is present or not.

IV. CONCLUSION AND FUTURE WORK

An analysis of the SSL/TLS protocol was made throughout the project. Limitations were shown in every version of TLS since it was developed. This raises the need to make sure that the recent version of TLS is secure and can tailor and prevent any hijack on the secure sessions.

The primary objective of my project is to design an automatic monitoring tool that have the capability to identify the major and minor TLS versions used by a remote physical or even virtual machine. In the long run the tool will be able to notify the person of charge of the domain and send them a notification that a security problem exists in their site and needs to be fixed. Additionally, a test of the latest TLS version will be done in the next phase of my project. Any vulnerability will be found would be added to the documented list. This tool will help in making the transactions more secure by checking the website or domain beforehand. Finally, testing any new protocol or enhancement of a protocol will make companies, domain owners aware of the problems that exist in their

REFERENCES

- [1] C.Y. Yeun and T. Farnham 2001, "Secure m-commerce with WPKI," In proceedings of 1st International Workshop for Asian PKI, October 2001.
- [2] K. Han, H. Mun, T. Shon, C.Y. Yeun and J.J. Park, "Secure and efficient public key management in next generation mobile networks," Personal and Ubiquitous Computing, Vol. 16, No. 6, pp. 677-687, August 2012.
- [3] H. Al Housani, J. Baek and C.Y. Yeun, "Survey on certificateless public key cryptography," In proceedings of International Conference for Internet Technology and Secured Transactions, pp. 53-58, 11-14 December 2011.
- [4] A. Freier, P. Karlton, P. Kocher (2011 Aug) "SOCKETS LAYER (SSL) PROTOCOL VERSION 3.0". [Online]. Available: <https://tools.ietf.org/html/rfc6101>
- [5] B. Moller, T. Duong, K. Kotowicz (2014, Sept) "THIS POODLE BITES: EXPLOITING THE SSL 3.0 FALLBACK". [Online]. Available: <https://www.openssl.org/~bodo/ssl-poodle.pdf>
- [6] T. Dierks, C. Allen (1999, Aug) "The TLS Protocol", [Online]. Available: <https://tools.ietf.org/html/rfc2246>
- [7] "Imperial Violet Poodle attacks on SSLv3" (2014, Oct 14). [Online]. Available: <https://www.imperialviolet.org/2014/10/14/poodle.html>
- [8] T. Ohgashi, T. Isobe, Y. Watanabe (2014, May 21), "How to Recover Any Byte of Plaintext on RC4", [Online]. Available: [http://link.springer.com/chapter/10.1007/978-3-662-43414-7\\_8#page-1](http://link.springer.com/chapter/10.1007/978-3-662-43414-7_8#page-1)
- [9] N. J. Alfardan "On the security of RC4 in TLS", [Online]. Available: <http://dl.acm.org/citation.cfm?id=2534793>

- [10] R. Barnes (2014, Oct 14), "Mozilla Security Blog", [Online]. Available: <https://blog.mozilla.org/security/2014/10/14/the-poodle-attack-and-the-end-ofssl-3-0/>
- [11] Information Security, "How exactly does the OpenSSL TLS heartbeat (Heartbleed) exploit work?", [Online]. Available: <http://security.stackexchange.com/QUESTIONS/55116/HOW-EXACTLY-DOES-THEOPENSSL-TLS-HEARTBEAT-HEARTBLEED-EXPLOIT-WORK> Accessed Apr. 10, 2014.
- [12] A. Freier, P. Karlton, and P. Kocher (1996, Nov. 18), "The SSL 3.0 Protocol", Netscape Communications Corp.
- [13] H. Kipp (1995, Feb. 9) "The SSL Protocol", Netscape Communications Corp. Protocol Version 1.1", RFC 4346, April 2006.
- [14] T. and C. Allen, "The TLS Protocol Version 1.0", RFC 2246, January 1999. [TLS1.1] Dierks, T. and E. Rescorla, "The Transport Layer Security
- [15] OpenSSL Security Advisory (15 Oct 2014), "SRTP Memory Leak", [Online]. Available: <https://www.openssl.org/news/secadv/20141015.txt>
- [16] B. Moeller, A. Langley (2015, Feb 20) "TLS Fallback Signaling Cipher Suite Value (SCSV) for Preventing Protocol Downgrade Attacks", [Online]. Available: <https://tools.ietf.org/html/draft-ietf-tls-downgrade-scsv-05>
- [17] "This POODLE bites: exploiting the SSL 3.0 fallback" (2014, Oct. 14), [Online]. Available: <https://googleonlinesecurity.blogspot.jp/2014/10/this-poodle-bitesexploiting-ssl-30.html>
- [18] "Lucky 13 – a new attack against SSL/TLS" (2014, Sept. 28), [Online]. Available: <http://www.infosecurity-magazine.com/news/lucky-13-a-new-attackagainstssltls/>
- [19] G. V. Bard (2014, Sept. 28), "A Challenging but Feasible Block wise Adaptive Chosen-Plaintext Attack on Ssl", [Online]. Available: <https://eprint.iacr.org/2006/136.pdf>
- [20] A. N. Alshamsi, Takamichi Saito. (2014, Sept. 6). "A Technical Comparison of IPsec and SSL" (Doctoral dissertation, Tokyo University of Technology). [Online]. Available: <http://eprint.iacr.org/2004/314.pdf>
- [21] "BlackHat2014USA\_Prompt\_20140808 - Cryptographic Protocol Evaluation toward Long-Lived Outstanding Security Consortium (CELLOS)". (n.d.). [Online]. Available: [https://www.cellos-consortium.org/index.php?BlackHat2014USA\\_Prompt\\_20140808](https://www.cellos-consortium.org/index.php?BlackHat2014USA_Prompt_20140808) [09 SEPT 13 Dr. Kris Gaj ECE646 cryptography GMU 2014 2014]
- [22] A. Delignat-Lavaud., & K. Bhargavan (2014, Sept. 28). "Virtual Host Confusion: Weaknesses and Exploits" (2014), [Online]. Available: [https://bh.ht.vc/vhost\\_confusion.pdf](https://bh.ht.vc/vhost_confusion.pdf)
- [23] D. Goodin (2011, Sept. 19) "Hackers break SSL encryption used by millions of sites". [Online]. Available: [http://www.theregister.co.uk/2011/09/19/beast\\_exploits\\_paypal\\_ssl/](http://www.theregister.co.uk/2011/09/19/beast_exploits_paypal_ssl/) [28 SEPT 2014]
- [24] IBM. (2014, Sept. 12). "IBM WebSphere Developer Technical Journal: Using the Java Secure Socket Extension in WebSphere Application Server." [Online]. Available: [http://www.ibm.com/developerworks/websphere/techjournal/0502\\_benantar/0502\\_benantar.html](http://www.ibm.com/developerworks/websphere/techjournal/0502_benantar/0502_benantar.html)
- [25] Cisco (2014, Sept. 28). "Introduction to Secure Sockets Layer Internet" [Online]. Available: <http://euro.ecom.cmu.edu/resources/elibrary/epay/SSL.pdf>
- [26] N. J. AlFardan and K. G. Paterson, "Lucky thirteen: Breaking the TLS and DTLS record protocols," in 2013 IEEE Symposium on Security and Privacy, 2013, pp. 526–540.
- [27] K. Bhargavan and G. Leurent, "On the practical (in-)security of 64-bit block ciphers: Collision attacks on HTTP over TLS and OpenVPN," Cryptology ePrint Archive, Report 2016/798, 2016, <http://eprint.iacr.org/2016/798>.
- [28] K. Bhargavan, A. Delignat-Lavaud, C. Fournet, A. Pironti, and P.-Y. Strub, "Triple handshakes and cookie cutters: Breaking and fixing authentication over TLS," in 2014 IEEE Symposium on Security and Privacy, 2014, pp. 98–113.
- [29] T. Jager, J. Schwenk, and J. Somorovsky, "On the security of TLS 1.3 and QUIC against weaknesses in PKCS#1 v1.5 encryption," in 22nd ACM Conference on Computer and Communications Security, 2015, pp. 1185–1196.
- [28] M. Kohlweiss, U. Maurer, C. Onete, B. Tackmann, and D. Venturi, "(de-) constructing TLS 1.3," in Progress in Cryptology–INDOCRYPT 2015. Springer, 2015, pp. 85–102.
- [29] M. Fischlin, F. Günther, B. Schmidt, and B. Warinschi, "Key confirmation in key exchange: A formal treatment and implications for TLS 1.3," in 2016 IEEE Symposium on Security and Privacy, 2016, pp. 197–206.
- [30] K. Bhargavan, C. Fournet, M. Kohlweiss, A. Pironti, P.-Y. Strub, and S. Zanella-Béguelin, "Proving the TLS handshake secure (as it is)," Cryptology ePrint Archive, Report 2014/182, 2014, <http://eprint.iacr.org/2014/182/>.
- [31] C. Cremers, M. Horvat, S. Scott, and T. van der Merwe, "Automated analysis and verification of TLS 1.3: 0-RTT, resumption and delayed authentication," in 2016 IEEE Symposium on Security and Privacy, 2016, pp. 470–485.
- [32] B. Dowling, M. Fischlin, F. Günther, and D. Stebila, "A cryptographic analysis of the TLS 1.3 handshake protocol candidates," in 22nd ACM Conference on Computer and Communications Security, 2015, pp. 1197–1210.
- [33] —, "A cryptographic analysis of the TLS 1.3 draft-10 full and pre-shared key handshake protocol," <http://eprint.iacr.org/2016/081>, 2016.
- [34] M. Bellare and B. Tackmann, "The multi-user security of authenticated encryption: AES-GCM in TLS 1.3," in Advances in Cryptology – CRYPTO 2016, 2016, pp. 247–276.

## **Session 10: Information Security**

Title: Towards using Transfer Learning for Botnet Detection  
(Authors: Basil Alothman, Prapa Rattadilok)

Title: Transparent Authentication: Utilising Heart Rate for User Authentication  
(Authors: Timibloudi S Enamamu, Nathan Clarke, Paul Haskell-Dowland, Fudong Li)

Title: Performance vs. Security: Implementing an Immutable Database in MySQL  
(Authors: Thomas Nash, Aspen Olmsted)

# Towards using Transfer Learning for Botnet Detection

Basil Alothman  
Faculty of Technology  
De Montfort University  
Leicester LE1 9BH, Great Britain  
p14029266@my365.dmu.ac.uk

Prapa Rattadilok  
Faculty of Technology  
De Montfort University  
Leicester LE1 9BH, Great Britain

**Abstract**—Botnet Detection has been an active research area over the last decades. Researchers have been working hard to develop effective techniques to detect Botnets. From reviewing existing approaches it can be noticed that many of them target specific Botnets. Also, many approaches try to identify any Botnet activity by analysing network traffic. They achieve this by concatenating existing Botnet datasets to obtain larger datasets, building predictive models using these datasets and then employing these models to predict whether network traffic is safe or harmful. The problem with the first approaches is that data is usually scarce and costly to obtain. By using small amounts of data, the quality of predictive models will always be questionable. On the other hand, the problem with the second approaches is that it is not always correct to concatenate datasets containing network traffic from different Botnets. Datasets can have different distributions which means they can downgrade the predictive performance of machine learning models. Our idea is instead of concatenating datasets, we propose using transfer learning approaches to carefully decide what data to use. Our hypothesis is “Predictive Performance can be improved by using transfer learning techniques across datasets containing network traffic from different Botnets”.

**Keywords**—component; Botnet-Detection; Transfer-Learning; data-distribution; improve-predictive-performance; network-traffic-analysis)

## I. INTRODUCTION

Traditional machine learning algorithms use datasets separately to create predictive models. In transfer learning, which is a subfield of machine learning, a group of datasets are used together to enhance the quality of such predictive models [1]. In more detail, transfer learning attempts to learn one or more tasks (known as the source tasks) and use the knowledge learnt to improve learning in another task (known as the target task). The source and target tasks are usually related.

We can summarise our contributions in the following two points: 1) We demonstrate that the distribution of data that belong to different Botnet types is different, which means concatenating such data without care is not always the right decision. 2) We demonstrate that using Transfer Learning, instead of traditional machine learning, can enhance the performance of predictive models, which leads to more accurate Botnet detection.

## II. THE DATA

### A. Obtaining the Data

We have downloaded the data that was used in [2]. The data is in Packet Capture (PCAP) format. Details of this data can be found in [3]. We have mainly worked with the Testing Dataset because it has more Botnets than the Training Dataset. Because the data is in PCAP format, we needed to transform it into a format that machine learning platforms such as WEKA [7] or SciKit Learn [8] understand (i.e. into Character Separated Values, or CSV, format). Therefore, we have downloaded and used FlowGenerator [4] which reads in a directory that contains one or more PCAP files and transforms them into CSV files. It generates several attributes (features) such as Source Port, Destination Port, Protocol, Flow Duration, Flow Bytes per second and Flow Packets per second. The original number of features generated by FlowGenerator is 26 and the total number of instances we obtained is ~309000.

### B. Labelling the Data

After obtaining the CSV file, we have labelled the data using the source and destination IP address fields as explained in [3]. The distribution of the Botnet and Normal traffic in the data varies. For example, the number of instances that belong to Botnet Osx\_trojan was as little as 28. Also, the number of instances that belong to Botnet Weasel Botmaster was 40. On the other hand, the number of instances that belong to Botnet Virut was 42254 and the number of instances that belong to Botnet Neris was 24071. Furthermore, the number of instances that belong to Normal traffic was 149727.

### C. Data PreProcessing

All features generated by FlowGenerator are Numeric which makes it easy to process data. The fields Source Port, Destination Port and Protocol are essentially categorical so we have used the one hot encoding technique to transform these into Dummy Variables. This has resulted in a very large dataset with > 60000 features.

### D. Outlier Removal

In order to obtain clean data, we have used the One-class SVM [9] method for outlier and novelty detection on the large

dataset we obtained from the previous step. We have removed all the instances that were deemed Outliers.

### III. PRINCIPAL COMPONENT ANALYSIS (PCA)

In data exploration and analysis, Principal Components Analysis [5] is a technique used to identify a smaller number of uncorrelated features (i.e. attributes or variables). These uncorrelated variables are usually known as the “principal components”. Its main objective is to explain the highest possible amount of variance with the smallest possible number of principal components. It is commonly used as a dimensionality reduction procedure as well as an exploratory procedure to examine whether there is separation amongst instances that belong to different classes.

We have Run the PCA algorithm on the dataset we obtained after performing all the preprocessing steps we explained. Fig. 1 shows a scatter plot of the first and second principal components. To preserve display space, we only show points (i.e. instances) that belong to Botnets: TBot, Zero access and Zeus.

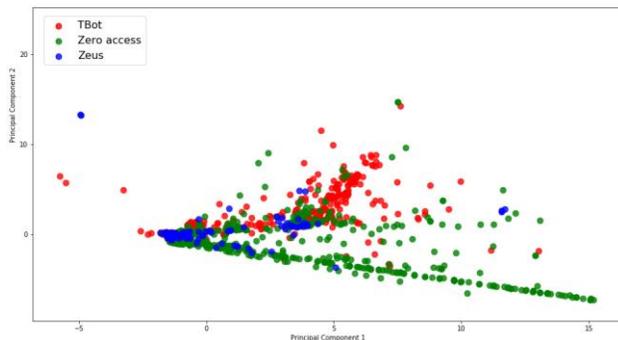


Figure 1: PCA Plot

It can be seen in the Fig.1 that there is separation between instances that belong to different Botnet types. Instances belonging to the same Botnet type tend to form a cluster and this solidifies the argument of this work.

### IV. TRANSFER LEARNING PLAN: IMPLEMENTATION AND EVALUATION

Since transfer learning uses related but separate datasets, we have split the dataset resulting from the steps explained above into smaller datasets. The split is based on the Botnet label. This means the split resulted in a separate dataset for each Botnet. We have split the Normal data (i.e. instances that belong to Normal traffic) into non-overlapping chunks and added one chunk to each Botnet dataset so that we have positive and negative

examples in each dataset (we made sure that the classes in the resulting datasets are balanced). We are currently evaluating the performance of several machine learning models on these datasets and we will choose the best classifier to use in our proposed Transfer Learning experiments. For transfer learning itself, we will use the TrAdaBoost algorithm that can be found in [6]. TrAdaBoost is based on the classical AdaBoost algorithm. It works when the source and target tasks have the same set of features, but different data distributions. In addition, TrAdaBoost assumes that some of the data in the source task can be useful (i.e. leads to positive transfer) and some can be harmful (i.e. leads to negative transfer). The idea is to assign weights to data from the source task in such a way that useful data can have more effect than harmful data. The author has made the java implementation of this approach publicly available. As the java source code of TrAdaBoost is freely available on the internet, we have already downloaded and integrated it into WEKA [7]. We are currently setting up the environment for running experiments with relatively large Botnet datasets as source tasks and smaller Botnet datasets as target tasks.

### V. CONCLUSION

Datasets that contain network traffic data belonging to different types of Botnets should not always be concatenated. In this work we have demonstrated that such data can have different distribution. Therefore, our suggestion to use transfer learning, instead of traditional machine learning, seems to be a reasonable method to enhance the performance of models used for Botnet identification and detection.

### REFERENCES

- [1] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [2] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of A. Shiravi, H. Shiravi, M. Tavallae, and A. A. Ghorbani, “Toward developing a systematic approach to generate benchmark datasets for intrusion detection,” *Computers & Security*, vol. 31, no. 3, pp. 357–374, 2012.
- [3] Beigi, Elaheh Biglar, et al. “Towards effective feature selection in machine learning-based botnet detection approaches.” *Communications and Network Security (CNS), 2014 IEEE Conference on.* IEEE, 2014.
- [4] ISCX/ISCXFlowMeter: ISCX Flow Meter. [ONLINE] Available at: <https://github.com/ISCX/ISCXFlowMeter> [Accessed on 12May 2017].
- [5] I. T. Jolliffe, (2002), *Principal Component Analysis* (Springer Series in Statistics) Hardcover.
- [6] W. Dai, Q. Yang, G. Xue, and Y. Yu, “Boosting for Transfer Learning,” *Proc. 24th Int’l Conf. Machine Learning*, pp. 193–200, June 2007.
- [7] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, Morgan Kaufmann, Fourth Edition, 2016.
- [8] Scikit-learn: Machine Learning in Python, Pedregosa et al., *JMLR* 12, pp. 2825–2830, 2011.
- [9] Larry M. Manevitz and Malik Yousef. 2002. One-class svms for document classification. *J. Mach. Learn. Res.* 2 (March 2002), 139–154.

# Transparent Authentication: Utilising Heart Rate for User Authentication

Timiblouidi S Enamamu  
Centre for Security,  
Communications and Network  
Research, Plymouth  
University, UK  
timiblouidi.enamamu@plymou  
th.ac.uk

Nathan Clarke  
Centre for Security,  
Communications and Network  
Research, Plymouth  
University, UK  
N.Clarke@plymouth.ac.uk

Paul Haskell-Dowland  
Security Research Institute,  
Edith Cowan University,  
Perth, Western Australia,  
p.haskell@ecu.edu.au

Fudong Li  
School of Computing,  
University of Portsmouth, UK  
fudong.li@port.ac.uk

**Abstract**—There has been exponential growth in the use of wearable technologies in the last decade with smart watches having a large share of the market. Smart watches were primarily used for health and fitness purposes but recent years have seen a rise in their deployment in other areas. Recent smart watches are fitted with sensors with enhanced functionality and capabilities. For example, some function as standalone device with the ability to create activity logs and transmit data to a secondary device. The capability has contributed to their increased usage in recent years with researchers focusing on their potential. This paper explores the ability to extract physiological data from smart watch technology to achieve user authentication. The approach is suitable not only because of the capacity for data capture but also easy connectivity with other devices – principally the Smartphone. For the purpose of this study, heart rate data is captured and extracted from 30 subjects continually over an hour. While security is the ultimate goal, usability should also be key consideration. Most bioelectrical signals like heart rate are non-stationary time-dependent signals therefore Discrete Wavelet Transform (DWT) is employed. DWT decomposes the bioelectrical signal into  $n$  level sub-bands of detail coefficients and approximation coefficients. Biorthogonal Wavelet (bior 4.4) is applied to extract features from the four levels of detail coefficients. Ten statistical features are extracted from each level of the coefficient sub-band. Classification of each sub-band levels are done using a Feedforward neural Network (FF-NN). The 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> levels had an Equal Error Rate (EER) of 17.20%, 18.17%, 20.93% and 21.83% respectively. To improve the EER, fusion of the four level sub-band is applied at the feature level. The proposed fusion showed an improved result over the initial result with an EER of 11.25%. As a one-off authentication decision, an 11% EER is not ideal, its use on a continuous basis makes this more than feasible in practice.

**Keywords**—User Authentication, Bioelectrical Signals, Discrete Wavelet Transform, Smart Watch, Smart Phone.

## I. INTRODUCTION

Authentication is the process of accurately authorizing a person to access secured information but it comes with some

inconvenience on the part of the subject because the subject will need to provide the correct credentials to access the information [1-4]. Transparent authentication has been proposed as a possible improvement over these inconveniences by applying biometric modalities in a non-intrusive manner (i.e. the user does not explicitly provide the sample, rather the sample is captured during a user's normal device interactions) [5-8]. These emerging biometric modalities include gait, body odour, ear resonance, lip print and bioelectrical signals [9-12]. The use of emerging biometric modalities is on the increase because of their advantages with respect to reliability, usability and accuracy in a transparent capture mode [13]. Recent research on emerging biometrics applying bioelectrical signals have focussed more on the use of Electrocardiogram (ECG), Electroencephalogram (EEG), Electromyogram (EMG), Mechanomyogram (MMG) and Electrooculography (EOG) with more emphases on EEG and ECG as shown in the work of Faust [14], Mporas [15], Borghini [16], Thomas [17], Suja Priyadharsini [18], Sabow [19], Miramontes [20] and Ito [21].

The direct involvement of a subject in the authentication process brings about usability issues at the point of entry. As stated earlier, in as much as security is the major concern in designing an authentication system, usability still plays an important role in the use of the of the system [3, 4, 22] but it comes with its own issues too [6]. The level of trade-off between security and usability plays a role in the choice of authentication system, a factor worth considering [4, 23, 24]. It is also expected that the security mechanism of a computing device should be robust and adapt to different environment [25]. The application of bioelectrical signals extracted via a smart watch for user authentication should improve usability as well as convenience due to the non-intrusive nature of the technique [26, 27]. In this paper, the authentication system builds the subject's profile by extracting the heart rate through a smart watch. Two experiments were conducted; the first experiment determined the persistency of the signal pattern of

the heart rate while the second experiment determined the viability of using the signal to authenticate a subject.

## II. BACKGROUND

Much of the previous research has used bioelectrical signals obtained from ECG's where the signals were extracted from the heart beat using specialized devices that were often intrusive in nature. Table 1 presents a summary of the studies on ECG bioelectrical signals, their methods for feature extraction, classification and results.

Table I. Study showing the use of Electrocardiogram (ECG) for authentication (MF: Morphological Features; LDA: Linear Discriminant Analysis; QDM: Quartile Discriminant Measurement; QRS :QRS Detection; MD; Mahalanobis Distance; PCA :Principal Component Analysis; KNN; K-nearest Neighbor; LDA; Linear Discriminant Analysis; WT: Wavelet Transform; CC: Correlation Coefficient; ICA: Independent Component Analysis; SVM: Support Vector Machine ; LZ: Lempel-Zil; RBF: Radian Basis Function)

Author	Feature Extractor	Classification	No of Subj.	Length	Success rate
[28]	MF	LDA	29	2 mins.	97 & 98%
[29]	MF	QDM			100%
[30]	QRS	MD	10	30 sec.	
[31]	QRS MF & PCA	KNN and LDA	20		94.47% & 97.8%
[32]	MF	MD	16	2 min.	100%
[33]	WT	CC	50	32-51 ms	89% - 95%,
[34]	QRS	CC	10		99%
[35]	ICA And WT	SVM	47	20 mins.	98.11 -99.33%
[36]	MF	LZ	19	10 mins.	100%
[37]	WT	RBF	16		91%
[38]	QRS	SVM			99.52%

Isreal [28] used the fiducial points from 29 subjects as the feature for authentication. The fiducial collection point includes the neck and the chest. The neck achieved a result of 82% while chest achieved 79%. [29] investigated the possibility of using the normalized time-domain features of Electrocardiogram (ECG) for improving of identification. The ECG signal is measured and extracted between the right and left arm using a Biopack MP-150. The first was to measure the ECG during rest and the second measurement was when the individual is active. The recording is done in 30 seconds on 10 male subjects in two sequences. The reading at a normal heart rate using ECG at a slow rate is 60 ~ 80 and 120 ~ 140 at a fast rate. For Feature Extraction and Classification, after analyzing the sampled data sequence of the ECG beat by beat, the characteristic points of its waveform of P-wave, QRS complex and T-wave are computed as the features for classification of the subjects.

Morphological Features, QRS Detection and Wavelet Transform are among the most used feature extraction methods listed in Table 1. Each of the methods has advantages and disadvantages depending on the type of signal and condition of the features were extracted. The morphological features method is suitable for ECG feature extraction and is suitable for heart rate because this rate varies from one heartbeat to the next [39]. This can show variable fiducial points for feature extraction which will affect the morphological features. QRS Detection has the advantage of efficient extraction of beat-to-beat intervals (RR) from long electrocardiogram (ECG) recordings, it is also suitable for real-time analysis of large datasets but has a disadvantage with regards to its of implementation in software as it is difficult to operate it in real time [40]. Wavelet Transform is chosen for the feature extraction because it has a varying window size, being broad at low frequencies and narrow at high frequencies. It is better suited for analysis of sudden, transient signal changes [41] and irregular data patterns, that is, impulses existing at different time instances [42]. From the works discussed earlier, it shows that the most used classification method is Neural Network and SVM. The two methods have their own advantage depending on the type of bioelectrical signal. Research suggest that neural networks can perform better in nonlinear statistical modeling and is an alternative to logical regression [43] while SVM performs better classification on emotional features which is prevalent in EEG signals [44].

## III. EXPERIMENTAL METHODOLOGY

### A. Data Collection and Experimental Design

Most of the data samples from the previous experiments are control samples [28, 30, 34]. While this is ideal some experimental studies, a typically highly controlled lab environment fails to understand the variance that would be exhibited from a real-life data capture. This study investigates several areas, the viability of the underlying technology to measure the signals successfully, a small-scale study to investigate the nature of the signal given a variety of tasks (e.g. walking, sitting) and also to determine the feasibility of the approach using real-life activity data. This led to the development of three experiments:

1. A technology evaluation of smart watches
2. An activity based experiment to examine the variability in the underlying signal
3. A real-life data capture to determine the feasibility of the approach.

The technology evaluation used three smartwatches Mio Fuse, Fitbit, Microsoft Band. A chest-band Polar H7 heart rate

monitor was used as a reference signal against which the smartwatches were compared. To appraise the accuracy of the signals extracted from the watches, the extracted heart rate signal from a subject wearing all the smart watches and the chest band were captured and analyzed. The signal extracted was compared again the chest band which is more accurate around the chest compared with other parts of the body [40]. An android smart phone with a third party application was installed on the phone to enable it to store the heart rate signals. The extraction from the smartwatch to the phone was via a Bluetooth connection. Taking usability into consideration, the mobile application and the smartwatch communicates without the intervention of the subject when extracting the bioelectrical signal. The application starts as the phone comes on and establishes a connection with the watch. The heart rate is extracted in beat per minute at a rate of 8 samples per second.

To study the variation of signals from one subject to another using the Microsoft band, the heart rate signal was extracted from five subjects. A predefined task was given to the five subjects to be repeated three times. These tasks included a combination of both low and high speed of walking, climbing up and down stairs, standing and sitting. The time between the three tasks ranged from a day to two days between tasks.

In the real-life data capture experiment, the aim was to develop a unique identifier for each of the 30 subjects by extracting features from the heart rate. The subjects were recorded for one hour without a predefined task to make it as natural as possible. As expected in a real life scenario, the possibility of environmental interference like noise (i.e. wireless and other Bluetooth connection) is expected. The data collected through the Microsoft band faced a number of issues including:

- **Disconnection:** The Microsoft band sometimes loses connection with the phone but with the application setting, it can re-establish connection without the intervention of the subject. To make up for this, the data collection time frame is increased makes room for any disconnection gap. The disconnection duration is indicated with a '*Null*' which is deleted in processing the data.
- **Heart rate acquisition:** the heart rate sensor takes some time to start recording the heart rate. At this stage the heart rate output remains constant and it is indicated as "*Acquiring*" until the band is locked to the app. The same remedy for the disconnection is applied to this too.

- **Sampling Rate:** the sampling can be set at 16 Hz, 32Hz and 64Hz. Due to android issues, the sampling rate setting can return to the default rate at the start of each extraction, it can be monitored to make sure the sampling rate is right at the beginning of each extraction. To solve this, after extraction all signals are down-sampled to 8 samples per second.

### B. Feature extraction

The feature extraction algorithm converts bioelectrical signal information into sets of feature vectors. The feature extraction method should be good enough and should meet some properties like repeatability, distinctiveness, quantity, accuracy, and efficiency [45]. However, the extraction technique will need to be carefully considered taking note of the nonstationary nature of bioelectrical signals. There are different types of techniques as earlier discussed which include Wavelet Transform [46,47,48], Independent Component Analysis [49], Morphological Features [31], Discrete Cosine Transform [50]. After investigating the properties of the heart rate signals, the Wavelet feature extraction technique is adopted using discrete wavelet transforms.

The use of discrete wavelet transforms is becoming popular in the measurement and analysis of time-frequency nonstationary signals and the spectral component variation [51,52]. It is widely used in feature extraction as in the case of Mallat [51], Subasi [53] and Jahankhani [48]. Wavelet transform is also useful in processing different types of transient signal analysis [54]. It decomposes a signal into a sub-band of wavelet signals which can be implemented with several wavelet families. The wavelet families include Biorthogonal, Morlet, Symlets, Mexican Hat, Haar, Daubechies, Coiflets, Meyer [55, 56]. Wavelet transform is classified into two types, continuous wavelet and discrete wavelet transform. Existing literature has shown that noise is an issue when processing a signal; this also applies to bioelectrical signals. To achieve an acceptable noise level in a signal, a filter is applied to increase the SNR. As stated earlier, the use of wavelet transform eliminates the direct application of a filter in this work because wavelet transform decomposition is used to implement noise reduction [57, 58]. Discrete wavelet transform decomposition splits the input signal into approximation of coefficients and detail coefficients [59,54]. This depends on the type of wavelet family used as a suitable wavelet can concentrate 90% of the signal energy on the decomposed coefficient [58]. The decomposition enables the signal to be analyzed at the different  $n$  levels [60]. Each  $n$  level is further decomposed into a high and low frequency signal component using a filter bank [54, 61].

Ten statistical features are extracted from each level of the sub-band levels are the Variance, Maximum, Amplitude Minimum Amplitude, Maximum Energy, Minimum Energy, Standard deviation, Peak2peak, Root mean square level(RMS), Mean or Median absolute deviation and Peak magnitude to RMS ratio.

C. Classification

To classify the features extracted, a Nueral Network (NN) is used. The classification evaluation metric calculates the Equal Error Rate (EER) using False Acceptance Rate (FAR) and False Rejection Rate (FRR).

- The Equal Error Rate (EER) is the point at which the False Acceptance Rate (FAR) and False Rejection Rate (FRR) meets also known as Receiver operating characteristic (ROC)
- The False Acceptance Rate (FAR) is the rate at which a subject that is legitimate is falsely refuse access to the system and
- The False Rejection Rate (FRR) is the rate at which an impostor is accepted as a legitimate subject.

IV. RESULTS

From Fig. 1 it is observed that the Fitbit, Mio Fuse and the Microsoft Band perform consistently with the Polar H7 in sequence as shown in Table 2.

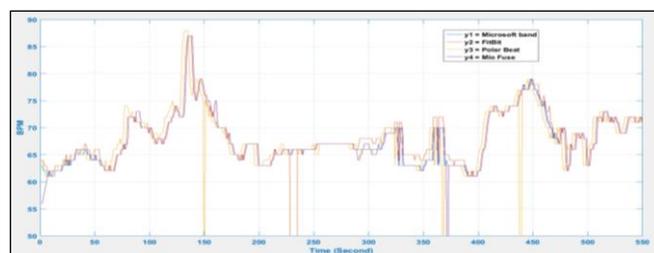


Figure 1. Bioelectrical recording from the Microsoft Band, Fitbit, Polar HR & Mio Fuse

Table II. Fitbit, Mio Fuse and Microsoft Band sensor comparison

Sensors	Microsoft Band	Mio Fuse	Fitbit	Polar H7monitor
Heart Rate	X	X	X	X
Accelerometer	X	-	X	-
Pedometer	X	X	-	-
Walking Speed	X	-	X	-
Calories	X	X	-	-
Distance	X	X	X	-
Gyroscope	X	-	-	-
Magnetometer	X	-	-	-
Altimeter	-	-	X	-
Ambient Light	X	-	-	-
Thermometer	X	-	-	-
Ultraviolet	X	-	-	-
Light Sensor	X	-	-	-
Galvanometer	X	-	-	-
Microphone	X	-	X	-

The result of the variability of subjects as illustrated in Fig. 2 shows that the five subjects have different signals amplitude that are not close and subjects can be differentiated and shows a potential to use this approach for authentication. There are changes depending on the activity carried out by the subjects. This shows that different activities affect the heart rate pattern therefore there is a need to categories the activities into high and low activities for effectivity analyzing the bioelectrical signals that will be extracted from the subjects.

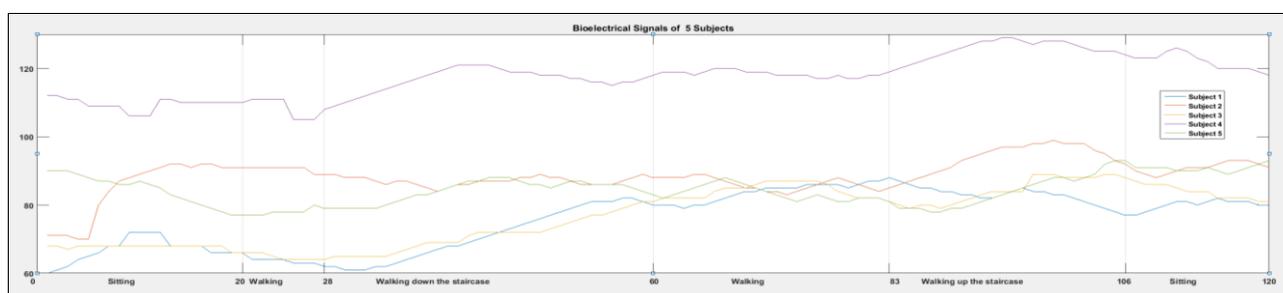


Figure 2. Bioelectrical signal of 5 subjects showing the pattern variance among the subject

The result from the 30 subjects applying the four sub-band classifications are encouraging as illustrated in Fig. 3. The use of a Nueral Network (NN) feed forward classifier achieved 17.2% EER at the first level which is the best result and 21.8 % at the fourth level as the worst. Level 1 and 2 sub-

bands have a higher score compared to level 3 and 4. This means that level 1 with the lowest score has 82.8% of all subjects accurately identified. The continuous reduction as the level increases does not mean that all subjects performed badly at the individual rate.

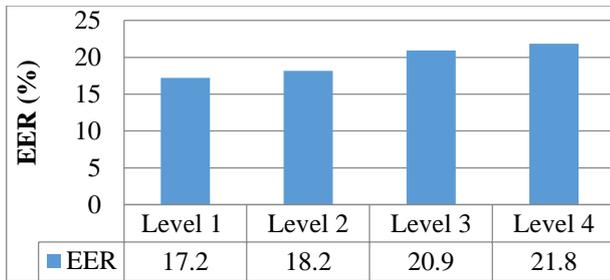


Figure 3. The EER sub-band classifications of subjects from level 1 to 4.

From Fig. 4 the EER of individual results across the four levels of sub-band shows that individual’s performance varies depending on the levels therefore fusion of the feature is undertaken to improve the result. The fusion is done after extracting the feature at various levels. The features are first normalized at each level before the fusion is done. The result at the fusion level has shown an improved EER of 11.25%.

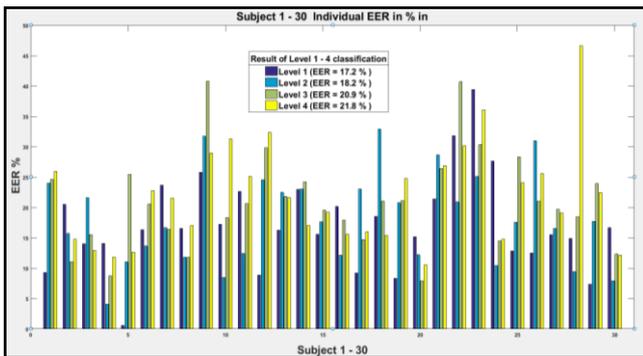


Figure 4. Showing result of individually performance

V. DISCUSSION

A close look at the Table 3 shows the performance between subjects at the different levels of the sub-band. The best individual performance at the first level is subject 5 with an EER of 0.6%, best at the second level is subject 4 with EER of 4.1%. Subject 20 has the best performance at the third and fourth levels with EER’s of 7.9% and 10.6% respectively. This mean performance cuts across difference sub-band levels.

It will be ideal to achieve a system performance of EER below 10% for the system which some subjects achieved. The performance of individual subjects achieving below the EER of 10 % cut across all levels. In level one, subject 1 (9.3%), 5 (0.6%), 12 (8.9%), 17 (9.2%), 19 (8.3%) and 29 (7.4%) achieved less than 10%. Level two results below 10% are recorded for subject 4 (4.1%), 10 (8.5%), 28 (9.4%) and 30 (7.9%). Level three shows subject 4 scoring 8.8% and 10 scoring 7.9% and level four has none though subject 20 achieved 10.6% which is closest to the expected mark.

Table III. Results of EER of Subjects at different levels of the sub-band

Subject’s EER result at different levels (%)									
ID	Level				ID	Level			
	1	2	3	4		1	2	3	4
1	9.3	24.0	24.6	25.9	16	20.1	12.1	17.9	15.6
2	20.5	15.7	11.0	14.8	17	9.2	23.0	14.7	16.0
3	14.0	21.6	15.5	12.9	18	18.6	32.9	21.0	15.4
4	14.1	4.1	8.8	11.8	19	8.3	20.8	21.1	24.8
5	0.6	11.0	25.4	12.6	20	15.2	12.2	7.9	10.6
6	16.3	13.7	20.6	22.8	21	21.4	28.7	26.4	26.9
7	23.7	16.7	16.4	21.5	22	31.8	20.9	40.7	30.1
8	16.6	11.8	11.8	17.0	23	39.4	25.1	30.3	36.1
9	25.8	31.7	40.8	29.0	24	27.7	10.4	14.5	14.8
10	17.2	8.5	18.3	31.2	25	12.8	17.6	28.3	24.1
11	22.6	12.4	20.6	25.1	26	12.5	31.0	21.0	25.6
12	8.9	24.6	29.9	32.3	27	15.5	16.6	19.7	19.1
13	16.3	22.5	21.8	21.6	28	14.9	9.4	18.5	46.6
14	23.0	23.0	24.2	17.0	29	7.4	17.7	23.9	22.4
15	15.6	17.7	19.6	19.2	30	16.7	7.9	12.3	12.1

The use of multiple instance of a biometric can add value to the result but it can also have implications depending on the dataset [62]. Fusion of biometric is done at different levels, the feature extraction level, match score level; and the decision level. The fusion of all level sub-bands is done at the feature level and the results showed an improved EER of 11.25%. This is an improvement of 5.95% which mean 88.75% of all subjects were accurately identified as shown in Fig. 5.

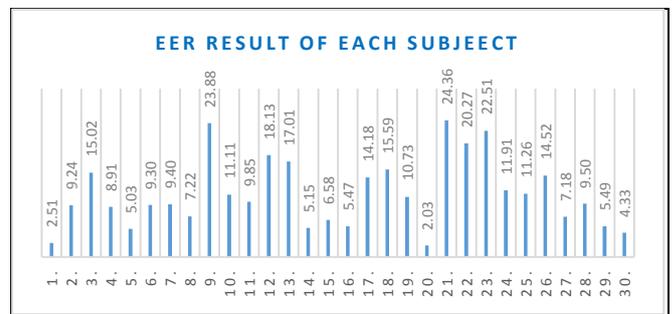


Figure 5. Showing classification result of all 4 levels individually

The experiment showed different subjects performed differently depending on the sub-band levels and the sub-band fusion classifications. Some subjects performed well on both while others on only one of the classification. It is expected that with the fusion of the sub-band, there should be improvement across all subjects but that is not the case. From the result of the sub-band fusion, it shows more subjects perform better while some unexpectedly did not improve in performance. This is seen in subject 3’s performance, there is little change in the sub-band fusion classification where they scored 15.02% which is almost the same on the 3 level sub-band results at 15.02%. It has a better result at level 4 scoring 12.9% compared to the sub-band fusion with 15.02%. The same is for subject 10 with the best result on level 1 at 8.5% compared to the sub-band fusion at 11.11%. Other subjects

scoring the best result at level 1 of the sub-band level include subject 12 scoring 8.90% compared to sub-band fusion scoring 18.13%, subject 13 at 16.3% (sub-band fusion 17.1%), subject 17 at 9.2% (sub-band fusion 14.18%), subject 19 at 8.3% (sub-band fusion 10.73%), subject 21 at 21.4% (sub-band fusion 24.36%) and subject 26 at 12.5% (sub-band fusion 14.52 %). The best results at the sub-band level 2 include subject 24 at 10.4% and subject 28 at 9.4% compared to the scoring at the sub-band fusion at 11.91% and 9.5% respectively. At Level 4, only subject 18 recorded their best performances at 15.4% compare to sub-band fusion at 15.59%. In term of individual performance, the fusion of all levels has shown to be effective in discrimination of subjects. 60% of individual results improved with the fusion introduced. While 40% of the subjects scored a better EER at the sub-band level. The best for each of them showed that subject 4 scored 8.8%, 12 (8.9%), 17 (9.2%), 19 (8.3%) at the 1<sup>st</sup> level, 2<sup>nd</sup> level have subject 10 scoring 8.5%, 28 (9.4%). These subjects individually performed below the expected 10% of EER. This brings to a total of subjects scoring below 10% of EER across the sub-band and the fusion classification to about 66%.

## VI. CONCLUSION

Over 66% of individuals achieved an EER below 10% across the fusion of sub-bands; the performance is promising noting that the overall EER performance of the fusion was 11.25%. The use of one bioelectrical signal is a limitation to the fact that it is affected by aging, emotional factors [63]. Therefore, the use of multi-instance, multi-modal or multi-bioelectrical signals is expected to enhance the performance and overcome these limitations. The use of the Microsoft band will be beneficial in this regard because as stated earlier, the sensors in the Microsoft band 2 can extract other bioelectrical signals like skin temperature, Galvanize Skin Response (GSR), Heart Rate Variability (HRV) and gyroscope and accelerometer for orientation. With these available signals on the Microsoft band, the system can be improved upon by applying multi-bioelectrical signals for Transparent User Authentication.

## VII. REFERENCES

- [1] EL-SAYED, A. 2015. Multi-Biometric Systems: A State of the Art Survey and Research Directions. (*IJACSA International Journal of Advanced Computer Science and Applications*, 6.
- [2] BRAZ, C. & ROBERT, J.-M. Security and usability: the case of the user authentication methods. Proceedings of the 18th International Conference of the Association Francophone d'Interaction Homme-Machine, 2006. ACM, 199-203.
- [3] JOSANG, A., ALFAYYADH, B., GRANDISON, T., ALZOMAI, M. & MCNAMARA, J. Security usability principles for vulnerability analysis and risk assessment. Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual, 2007. Ieee, 269-278.
- [4] BRAZ, C., SEFFAH, A. & M'RAIHI, D. 2007. Designing a trade-off between usability and security: A metrics based-model. *Human-Computer Interaction-INTERACT 2007*. Springer.
- [5] LI, F., CLARKE, N., PAPADAKI, M. & DOWLAND, P. Behaviour Profiling for Transparent Authentication for Mobile Devices. Emerging Security Technologies (EST), 2010 International Conference on, 2010. IEEE, 77-82.
- [6] CLARKE, N. 2011. Transparent User Authentication: Biometrics, RFID and Behavioural Profiling, London, Springer.
- [7] DEUTSCHMANN, I. & LINDHOLM, J. Behavioral biometrics for DARPA's active authentication program. *International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2013. IEEE, 1-8.
- [8] CRAWFORD, H. & RENAUD, K. 2014. Understanding user perceptions of transparent authentication on a mobile device. *Journal of Trust Management*, 1, 7.
- [9] KHUSHK, K. & IQBAL, A. An overview of leading biometrics technologies used for human identity. Engineering Sciences and
- [10] GAFUROV, D. 2008. Performance and security analysis of gait-based user authentication. University of Oslo.
- [11] ADEOYE, O. S. 2010. A survey of emerging biometric technologies. *International Journal of Computer Applications & Information Technology*, 9.
- [12] SHARMA, P., DEO, S., VENKATESHAN, S. & VAISH, A. 2011. Lip print recognition for security systems: an up-coming biometric
- [13] BRAZ, C. & ROBERT, J.-M. Security and usability: the case of the user authentication methods. Proceedings of the 18th International Conference of the Association Francophone d'Interaction Homme-Machine, 2006. ACM, 199-203
- [14] FAUST, O., ACHARYA, U. R., ADELI, H. & ADELI, A. 2015. Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. *Seizure*, 26, 56-64.
- [15] MPORAS, I., TSIRKA, V., ZACHARAKI, E. I., KOUTROUMANIDIS, M., RICHARDSON, M. & MEGALOOIKONOMOU, V. 2015. Seizure detection using EEG and ECG signals for computer-based monitoring, analysis and management of epileptic patients. *Expert Systems with Applications*, 42, 3227-3233.
- [16] BORGHINI, G., ARICÒ, P., GRAZIANI, I., SALINARI, S., SUN, Y., TAYA, F., BEZERIANOS, A., THAKOR, N. V. & BABILONI, F. 2016. Quantitative assessment of the training improvement in a motor-cognitive task by using EEG, ECG and EOG signals. *Brain topography*, 29, 149-161.
- [17] THOMAS, M., DAS, M. K. & ARI, S. 2015. Automatic ECG arrhythmia classification using dual tree complex wavelet based features. *AEU-International Journal of Electronics and Communications*, 69, 715-721.
- [18] SUJA PRIYADHARSINI, S., EDWARD RAJAN, S. & FEMILIN SHENIHA, S. 2016. A novel approach for the elimination of artefacts from EEG signals employing an improved Artificial Immune System algorithm. *Journal of Experimental & Theoretical Artificial Intelligence*, 28, 239-259.
- [19] SABOW, A., GOH, Y., ZULKIFLI, I., SAZILI, A., KAKA, U., AB KADI, M., EBRAHIMI, M., NAKYINSIGE, K. & ADEYEMI, K. 2016. Blood parameters and electroencephalographic responses of goats to slaughter without stunning. *Meat science*, 121, 148-155.
- [20] MIRAMONTES, R., AQUINO, R., FLORES, A., RODRÍGUEZ, G., ANGUIANO, R., RÍOS, A. & EDWARDS, A. 2017. PlaMoS: a remote mobile healthcare platform to monitor cardiovascular and respiratory variables. *Sensors*, 17, 176.
- [21] ITO, K., HARADA, Y., TANI, T., HASEGAWA, Y., NAKATSUJI, H., TATE, Y., SETO, H., AIKAWA, T., NAKAYAMA, N. & OHKURA, M. 2017. Evaluation of feelings of excitement caused by auditory stimulus in driving simulator using biosignals. *Advances in Affective and Pleasurable Design*. Springer.
- [22] BRAZ, C. & ROBERT, J.-M. Security and usability: the case of the user authentication methods. Proceedings of the 18th International Conference of the Association Francophone d'Interaction Homme-Machine, 2006. ACM, 199-203.

- [23] SCHULTZ, E. E., PROCTOR, R. W., LIEN, M.-C. & SALVENDY, G. 2001. Usability and security an appraisal of usability issues in information security methods. *Computers & Security*, 20, 620-634.
- [24] DOURISH, P. & REDMILES, D. An approach to usable security based on event monitoring and visualization. Proceedings of the 2002 workshop on new security paradigms, 2002. ACM, 75-81.
- [25] NAG, A. K., DASGUPTA, D. & DEB, K. An Adaptive Approach for Active Multi-Factor Authentication. 9th Annual Symposium on Information Assurance (ASIA'14), 2014. 39.
- [26] CLARKE, N. L. & FURNELL, S. M. 2005. Authentication of users on mobile telephones—A survey of attitudes and practices. *Computers & Security*, 24, 519-527.
- [27] WAYMAN, J., JAIN, A., MALTONI, D. & MAIO, D. 2005. An introduction to biometric authentication systems. *Biometric Systems*. Springer.
- [28] ISRAEL, S. A., IRVINE, J. M., CHENG, A., WIEDERHOLD, M. D. & WIEDERHOLD, B. K. 2005. ECG to identify individuals. *Pattern recognition*, 38, 133-142.
- [29] SHEN, T.-W. 2005. Biometric identity verification based on electrocardiogram (ECG), University of Wisconsin--Madison.
- [30] KIM, K.-S., YOON, T.-H., LEE, J.-W., KIM, D.-J. & KOO, H.-S. A robust human identification by normalized time-domain features of electrocardiogram. Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the, 2006. IEEE, 1114-1117.
- [31] WANG, Y., AGRAFIOTI, F., HATZINAKOS, D. & PLATANIOTIS, K. N. 2008. Analysis of human electrocardiogram for biometric recognition. *EURASIP journal on Advances in Signal Processing*, 2008, 19.
- [32] GAHI, Y., LAMRANI, M., ZOGLAT, A., GUENNOUN, M., KAPRALOS, B. & EL-KHATIB, K. Biometric identification system based on electrocardiogram data. New Technologies, Mobility and Security, 2008. NTMS'08., 2008. IEEE, 1-5.
- [33] CHAN, A. D., HAMDY, M. M., BADRE, A. & BADEE, V. 2008. Wavelet distance measure for person identification using electrocardiograms. *Instrumentation and Measurement, IEEE Transactions on*, 57, 248-253.
- [34] SASIKALA, P. & WAHIDAUANU, R. 2010. Identification of individuals using electrocardiogram. *International Journal of Computer Science and Network Security*, 10, 147-153.
- [35] YE, C., COIMBRA, M. T. & KUMAR, B. Investigation of human identification using two-lead electrocardiogram (ECG) signals. Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on, 2010. IEEE, 1-8.
- [36] COUTINHO, D. P., FRED, A. L. & FIGUEIREDO, M. A. One-lead ECG-based personal identification using Ziv-Merhav cross parsing. Pattern Recognition (ICPR), 2010 20th International Conference on, 2010. IEEE, 3858-3861.
- [37] SIDEK, K. A. & KHALIL, I. Automobile driver recognition under different physiological conditions using the electrocardiogram. *Computing in Cardiology*, 2011, 2011. IEEE, 753-756.
- [38] LEE, J., CHEE, Y. & KIM, I. 2012. Personal identification based on vectorcardiogram derived from limb leads electrocardiogram. *Journal of Applied Mathematics*, 2012,
- [39] RODWELL, P., FURNELL, S. & REYNOLDS, P. L. 2007. A non-intrusive biometric authentication mechanism utilising physiological characteristics of the human head. *Computers & Security*, 26, 468-478.
- [40] HASHEM, M., SHAMS, R., KADER, M. A. & SAYED, M. A. Design and development of a heart rate measuring device using fingertip. Computer and Communication Engineering (ICCCE), 2010 International Conference on, 2010. IEEE, 1-5.
- [41] SIDEK, K. A. & KHALIL, I. Automobile driver recognition under different physiological conditions using the electrocardiogram. *Computing in Cardiology*, 2011, 2011. IEEE, 753-756.
- [42] AL-FAHOUM, A. S. & AL-FRAIHAT, A. A. 2014. Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains. *ISRN neuroscience*, 2014.
- [43] Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology*, 49(11), 1225-1231.
- [44] RIERA-GUASP, M., ANTONINO-DAVIU, J. A., PINEDA-SANCHEZ, M., PUCHE-PANADERO, R., & PEREZ-CRUZ, J. (2008). A general approach for the transient detection of slip-dependent fault components based on the discrete wavelet transform. *IEEE transactions on Industrial electronics*, 55(12), 4167-4180.
- [45] TUYTELAARS, T. & MIKOLAJCZYK, K. 2008. Local invariant feature detectors: a survey. *Foundations and Trends® in Computer Graphics and Vision*, 3, 177-280.
- [46] SUBASI, A. 2007. EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, 32, 1084-1093.
- [47] TING, W., GUO-ZHENG, Y., BANG-HUA, Y. & HONG, S. 2008. EEG feature extraction based on wavelet packet decomposition for brain computer interface. *Measurement*, 41, 618-625.
- [48] JAHANKHANI, P., KODOGIANNIS, V. & REVETT, K. EEG signal classification using wavelet feature extraction and neural networks. Modern Computing, 2006. JVA'06. IEEE John Vincent Atanasoff 2006 International Symposium on, 2006. IEEE, 120-124
- [49] SUBASI, A. & GURSOY, M. I. 2010. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Systems with Applications*, 37, 8659-8666.
- [50] TAWFIK, M. M. & KAMAL, H. S. T. 2011. Human identification using QT signal and QRS complex of the ECG. *Online J. Electron. Elect. Eng*, 3, 383-387.
- [51] MALLAT, S. G. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11, 674-693.
- [52] ADDISON, P. S., WALKER, J. & GUIDO, R. C. 2009. Time--frequency analysis of biosignals. *IEEE engineering in medicine and biology magazine*, 28, 14-29.
- [53] SUBASI, A., & ERCELEBI, E. (2005). Classification of EEG signals using neural network and logistic regression. *Computer methods and programs in biomedicine*, 78(2), 87-99.
- [54] GOKHALE, M. & KHANDUJA, D. K. 2010. Time domain signal analysis using wavelet packet decomposition approach. *International Journal of Communications, Network and System Sciences*, 3, 321.
- [55] OVANESOVA, A. & SUAREZ, L. 2004. Applications of wavelet transforms to damage detection in frame structures. *Engineering structures*, 26, 39-49.
- [56] TSOU, C., HSIEH, C.-H., LIANG, M.-C., HUANG, P.-W. & LEE, S.-Y. ECG acquisition system with heart rate detection and energy harvesting for drivers. Bioelectronics and Bioinformatics (ISBB), 2015 International Symposium on, 2015. IEEE, 31-34.
- [57] UNSER, M. & ALDROUBI, A. 1996. A review of wavelets in biomedical applications. *Proceedings of the IEEE*, 84, 626-638.
- [58] AGBINYA, J. I. Discrete wavelet transform techniques in speech processing. TENCON'96. Proceedings., 1996 IEEE TENCON. Digital Signal Processing Applications, 1996. IEEE, 514-519.
- [59] BO-ZHI, F. & HONG-BIN, Z. Feature extraction using wavelet packet decomposition based on MPEG-I. Computer Science and Software Engineering, 2008 International Conference on, 2008. IEEE, 1048-1052.
- [60] TSAI, S.-J. S. 2002. Power transformer partial discharge (PD) acoustic signal detection using fiber sensors and wavelet analysis, modeling, and simulation.
- [61] LAINE, A. & FAN, J. 1993. Texture classification by wavelet packet signatures. *IEEE Transactions on pattern analysis and machine intelligence*, 15, 1186-1191.
- [62] ATREY, P. K., HOSSAIN, M. A., EL SADDIK, A. & KANKANHALLI, M. S. 2010. Multimodal fusion for multimedia analysis: a survey. *Multimedia systems*, 16, 345-379.
- [63] CUNNINGHAM, J. J. 1980. A reanalysis of the factors influencing basal metabolic rate in normal adults. *The American journal of clinical nutrition*, 33, 2372-2374.

# Performance vs. Security: Implementing an Immutable Database in MySQL

Thomas Nash, Aspen Olmsted  
Department of Computer Science  
College of Charleston, Charleston, SC 29401  
nashtf@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—Maintaining transactional history is crucial to unraveling the changes any unauthorized user makes to a system, and this logging database is often a prime target for attackers. One common approach to maintaining this security is through an immutable database. There are many ways to implement such a database which maintains appropriate security requirements, each ranging in the complexity and effort required to configure. What remains constant amongst all the methods is that the database is available only to a restricted, defined set of users and that records may only be inserted and not updated or deleted to maintain a proper history. We demonstrate two of these methods in the MySQL database system and compare performance and capabilities amongst them.

**Keywords**- *Immutable databases; security*

## I. INTRODUCTION

Immutable databases, also known as "insert-only" databases, are a secure approach to logging. These databases can provide higher levels of security by not allowing updates to records to occur, thus preserving the full history. Instead, databases create a new tuple every time an update is made to a record as opposed to updating it. Additional columns are needed to flag whether a record is active or not and to keep track of the time window in which the record was active with the associated attributes. While there is additional storage overhead, it comes without a significant reduction in performance.

For this work, we examine how to create such a database in a widely available relational database management system (RDBMS). We have chosen to use MySQL due to its popularity and multiple possibilities for a solution. MySQL presents at least two possible avenues for implementing an immutable database: through restrictions on the default storage engine InnoDB and through an alternative storage engine ARCHIVE. We will cover the specifics of each approach later in Section IV.

The organization of the paper is as follows. Section 2 describes the related work and how we look to make use of or expand upon each of the cited articles. In Section 3 we give a motivating example where such a database could be of use. Section 4 details the implementations and provides performance comparisons in addition to commentary on each of the solutions. We conclude our findings and discuss future work in Section 5.

## II. RELATED WORK

Duncan and Whittington [1] motivate the idea of immutable databases in MySQL as a use case for cloud service audit logging. They identify immutable databases as a secure method to tracking activities in cloud services. They propose a few approaches to how to implement such a database by using available constraints in MySQL, namely removing all user access to modify/delete records, removing the commands from the software or using an archive database. We will expand on their work by implementing the first and third approach to creating an immutable database.

Platner [2] covers insert-only databases in the broader context of OLTP and OLAP operations. The authors identify some benefits of using these databases, notably the reduced number of locks required as records are only inserted. Additionally, the database can be easily horizontally distributed due to the absence of updates. This paper provides a deeper understanding of the underlying mechanics of immutable database and provides insights into the performance gains.

Waters et al. [3] cover the details of secure audit logging with an implementation in SQL. One of the key attributes of a secure audit log is tamper resistant, that is only the creator can create entries, and they cannot be altered. Second, individual records should be created in such a way that the detection of missing records is possible. We use this paper as a basis for the components required for audit logging and how they are used in real-world scenarios.

## III. MOTIVATING EXAMPLE

The idea of an immutable database motivated in [1] for audit logging as described in [3] is one of the most relevant applications. Given the current usage of MySQL in the widely prevalent LAMP architecture, being able to make modifications to the existing backend is essential. Logging databases are often the target of intruders as they can be erased or modified to cover their actions. The ability to easily configure a database which protects the log integrity can be critical to identifying and tracing the actions of a nefarious party. As seen in [3], there can be additional security built on top of the immutable database to provide confidentiality.

## IV. IMPLEMENTATION

We believe that it is possible to create an immutable database using the default MySQL storage engine InnoDB which is comparable in performance and usability to the default InnoDB engine and the ARCHIVE alternative storage engine. The ARCHIVE engine provides immutability out of the box as its DML does not support DELETE or UPDATE. To test this hypothesis, we create three databases: one using the default InnoDB storage engine, one which is built on InnoDB but is given immutability through configuration changes and has some storage enhancements, and one using the default ARCHIVE storage engine.

To achieve immutability, we remove all non-root users' ability to perform operations other than SELECT and INSERT. To attempt to increase storage capabilities, we also set the ROW\_FORMAT to be compressed as seen in Figure 1.

```
USE sys;

CREATE TABLE IF NOT EXISTS Archive (
  EventDate DATETIME NOT NULL,
  ServerName VARCHAR(10) NOT NULL,
  EventMessage VARCHAR(500) NOT NULL,
  Activity VARCHAR(10) NOT NULL,
  FileAffected VARCHAR(100) NOT NULL
) ENGINE = ARCHIVE;

CREATE TABLE IF NOT EXISTS Inno (
  EventDate DATETIME NOT NULL,
  ServerName VARCHAR(10) NOT NULL,
  EventMessage VARCHAR(500) NOT NULL,
  Activity VARCHAR(10) NOT NULL,
  FileAffected VARCHAR(100) NOT NULL
) ENGINE = InnoDB;

CREATE TABLE IF NOT EXISTS Immutable (
  EventDate DATETIME NOT NULL,
  ServerName VARCHAR(10) NOT NULL,
  EventMessage VARCHAR(500) NOT NULL,
  Activity VARCHAR(10) NOT NULL,
  FileAffected VARCHAR(100) NOT NULL
) ENGINE = InnoDB, ROW_FORMAT=COMPRESSED;
```

Figure 1. Table DDL

We test the performance of each database by performing measures of reading and write times as well as storage requirements on each for a defined number of tuples. Figure 2 shows the results of these tests for trickle inserts of 1,000,000 rows and a SELECT query against the table which involves sorting all 1,000,000 rows on EventMessage which was set to be a random string of 108 characters.

While switching to use the ARCHIVE storage engine appears to be the most efficient both in INSERT and SELECT times as well as storage requirements, there are some hindrances. The most significant of which is that ARCHIVE engine does not permit indexing [4], therefore searches over time can be quite lengthy as the logs grow. Using the default InnoDB engine is by far the worst performer, but the simple addition of row compression makes a significant difference in selection times

TABLE 1. Performance metrics of each table

Measure	InnoDB	Immutable InnoDB	ARCHIVE
Storage size	232 MB	103 MB	26 MB
SELECT time	75.891s	11.485s	9.735s
INSERT time	2355.266s	2416.406s	41.343s

## V. CONCLUSIONS AND FUTURE WORK

In this work, we provide an implementation of an immutable database table built using the InnoDB storage engine in MySQL and compare it with a table created with default settings and one created with the ARCHIVE storage engine.

Immutability is easily achieved in readily available software through multiple approaches, but each comes with its advantages and drawbacks. It is up to the administrator to decide the level of security they wish to obtain at what cost. Removing the ability to run certain commands certainly, limits functionality on the server but provides the highest level of security. Simply using a different storage engine is the next step down in security, but privilege escalation can allow users to simply change the storage engine and gain access to other DML commands. The results indicate that using the ARCHIVE engine provides the best performance at the base level, but limits user ability to effectively search logs as they grow. Simply removing privileges on the default storage engine is not enough, but can achieve the desired effects of security and similar storage and read performance to the archive with simple table modifications.

In future work, we will examine more in-depth modifications to the RDBMS software itself to remove any ability to run the UPDATE or DELETE commands which can modify existing records.

## WORKS CITED

- [1] B. Duncan and M. Whittington, "Creating an Immutable Database for Secure Cloud Audit Trail and System Logging," in *Eighth International Conference on Cloud Computing, GRIDS, and Virtualization*, Athens, Greece, 2017.
- [2] H. Platner, "A Common Database Approach for OLTP and OLAP Using an In-Memory Column Database," in *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data (SIGMOD '09)*, New York, 2009.
- [3] B. R. Waters, D. Balfanz, G. Durfee and D. Smetters, "Building an Encrypted and Searchable Audit Log," in *NDSS Symposium*, San Diego, 2004.
- [4] Oracle Corporation, "MySQL :: MySQL 5.7 Reference Manual :: 15 Alternative Storage Engines," 2017. [Online]. Available: <https://dev.mysql.com/doc/refman/5.7/en/storage-engines.html>. [Accessed 14 June 2017].

## **Session 11: Cloud Security**

An Architecture for Privacy-preserving Sharing of CTI with 3rd party Analysis Services  
(Authors: Fabio Giubilo, Ali Sajjad, Mark Shackleton, David W. Chadwick, Wenjun Fan, Rogério de Lemos)

Title: From E-government to Cloud-government: Challenges of Jordanian Citizens' Acceptance for Public Services  
(Authors: Abeer Alkhwalidi, Mumtaz Kamala, Rami Qahwaji)

Title: Data Subject Rights in the Cloud A grounded study on data protection assurance in the light of GDPR  
(Authors: Alaa Altorbaq, Fredrik Blix, Stina Sörman)

Title: Distributed Computing Framework in Security: Case Study of Encryption Method  
(Authors: Shuaiyi Bu, Shuxin Yang, Haoming Ji)

# An Architecture for Privacy-preserving Sharing of CTI with 3<sup>rd</sup> party Analysis Services

Fabio Giubilo, Ali Sajjad, Mark Shackleton  
Security Futures Practice  
British Telecommunications plc (BT)  
Adastral Park, United Kingdom  
{fabio.giubilo,ali.sajjad,mark.shackleton}@bt.com

David W. Chadwick, Wenjun Fan, Rogério de Lemos,  
University of Kent  
Canterbury, United Kingdom  
{w.fan, r.delemos, d.w.chadwick}@kent.ac.uk

**Abstract**—Increasing numbers of Small and Medium Enterprises (SME) are outsourcing or hosting their services on different Cloud Service Providers (CSP). They are also using different security services from these CSPs such as firewalls, intrusion detection/prevention systems and anti-malware. Although for the SMEs the main purpose of using these security services is to protect their cyber assets, either physical or virtual, from security threats and compromises, a very useful and valuable by-product of these security services is the wealth of Cyber Threat Information (CTI) that is collected over time. However, a common problem faced by SMEs is that they lack the resources and expertise for monitoring, analysing and reacting to any security notifications, alerts or events generated by the security services they have subscribed to. An obvious solution to this problem is that the SMEs outsource this problem to a cloud based service as well, by sharing their CTI with this service and allowing it to analyse the information and generate actionable reports or patches. The more CTI obtained from different SMEs, the better the analysis result. In this paper, we try to address some of the privacy and confidentiality issues that arise as a result of different SMEs sharing their CTI with such a third party analysis service for the aggregate analysis scenario we just described. We present the design and architecture of our solution that aims to allow SMEs to perform policy-based sharing of CTI, while also offering them flexible privacy and confidentiality controls.

**Keywords:** data privacy and confidentiality, cyber threat information, analysis services, infrastructure architecture, policy based sharing

## I. INTRODUCTION

According to a recent survey commissioned by the UK Cabinet Office [6], the annual expense in the UK for cyber-crime is £27 billion (1.8% of GDP). This also states an estimate for UK cybercrime losses: £3 billion for citizens and the government, up to £21 billion for companies. Another report by the European Union Agency for Network and Information Security (ENISA) [3] states: “three-quarters of the businesses have seen cyber security as a concern for some time. The majority of respondents believed that their organisation has been the victim of a targeted attack. And almost a third of them reported a significant business impact.”

Information security is also becoming a serious matter to consider for Small and Medium Enterprises (SME). They often

wish to host their services on different Cloud Service Providers (CSP), increasingly deciding to provide their services over the Internet, exposing the services themselves to potential malicious users who attempt to disrupt those services. For this reason SMEs need security services at several layers, such as application (e.g., login portal of their service), network and infrastructure layers (e.g., the servers providing the service, the enterprise network and so on). Thus, for SMEs security is a complex issue to deal with, which often requires cyber security professionals, more expensive facilities and additional costs.

For enterprise customers there are a range of security services available from 3<sup>rd</sup> parties, typically referred to as Managed Security Services (MSS). Depending on the MSS provider (MSSP), the MSS can offer several security capabilities. Examples of MSSPs include BT Assure Threat Monitoring [19], SAP Enterprise Threat Detection [20], HPE SIEM Solutions [21], McAfee Enterprise Security Manager [22] and AlienVault Unified Security Management [23].

The Intelligent Protection Service [1] is an MSS developed by British Telecom which is aimed more towards SME customers. It provides services, such as firewalls, intrusion detection/prevention systems, anti-malware analysis, web reputation protection, log inspection and integrity monitoring. This solution can be managed by the SME itself, if they have sufficient capability and skills, or could in principle be outsourced to a 3<sup>rd</sup> party.

A very useful and valuable by-product of these security services is the wealth of Cyber Threat Information (CTI) collected over time by the MSS. CTI is defined by the National Institute of Standard and Technology (NIST) as any valuable information that can be used to identify, assess, monitor and respond to cyber threats [15]. As cyber threats are increasing considerably, SMEs have to be aware of potential risks, to deal with them in a timely and appropriate manner. This often represents a difficult challenge, due to SMEs’ lack of resources, knowledge and expertise, firstly for handling the large amount of CTI data gathered by an MSS without being overwhelmed by it and secondly for monitoring, analysing and responding to any security notifications, alerts or events generated by the MSS.

An obvious solution for this concern might consist of outsourcing the analysis process to a cloud-based service by sharing the CTI data itself with the service, allowing it to generate actionable reports and/or patches. It is clear that in order to achieve a better, effective and timely analysis outcome the analysis service requires as much CTI data as possible, ideally from different SMEs. In this paper, we aim to introduce a capability for SMEs to be able to allow policy controlled sharing of CTI gathered from an MSS, while preserving privacy and confidentiality of that information. This aggregated CTI from different SMEs will be analysed by a cloud-based third party analysis service and the results will be shared among the SMEs. In our scenario, the above mentioned third party analysis service will be C3ISP [24], which is a flexible framework for carrying out secure data analytics being developed as part of a Horizon 2020 collaborative R&D project.

In the remainder of this paper we discuss related work in section II, the motivation for SME information sharing in section III, and the system architecture in section IV, followed by an analysis of the architecture in section V. Our conclusions are then presented in section VI.

## II. RELATED WORK

Although we are presenting an original proposal for managing collaborative collection, sharing and analysis of CTI, several cyber security solutions exist that are related to our effort. As noted earlier, existing MSS are aimed towards larger enterprises, rather than SMEs. An example is BT Assure Threat Monitoring (ATM), which is a BT security solution oriented to enterprise customers. It is a security event monitoring service running 24x7x365. It operates by re-directing enterprise device data to a central BT Repository, where analysts examine and filter millions of messages from many devices, to discern the irrelevant ones from the suspicious and critical ones, and eventually notify the enterprise of any security concerns. Currently it is not possible to derive additional intelligence and more accurate analysis by sharing/combining the data belonging to different enterprises, since this data is stored in a strictly isolated manner (as it contains sensitive and confidential information).

Coco Cloud [13, 14] was a project that enabled cloud users to securely and privately share their information in the cloud environment. It provided mechanisms to raise trust in cloud services and therefore raise their widespread adoption with consequent benefits for users and for the digital economy in general. The main objectives consisted of:

- facilitating the writing, understanding, analysis, management, enforcement and dissolution of data sharing agreements (DSAs);
- considering the most appropriate enforcement mechanisms depending on the underlying infrastructure and context for enforcing data usage policies;
- addressing key challenges for legally compliant data sharing in the cloud.

By taking a “compliance by design” approach, the project placed an early emphasis on understanding and incorporating legal and regulatory requirements into the DSAs. We are using DSA mechanisms such as these within our framework [5].

## III. MOTIVATION

We allow SMEs to set up federations of information providers/consumers (called prosumers) in order both to obtain the information from other SMEs and to exploit the capacity of analysis of an amount of data such as might be available to a larger enterprise today. The overarching goal is that SMEs should be able to benefit from the MSS, in much the same way that a larger enterprise does. The barrier of the small size of an SME is overcome by allowing SMEs to share and combine security and threat information, so that its value will approximate to that available to a larger enterprise.

The need to share information between SMEs of course brings its own challenges [2-4]. In particular, the information might be considered to contain commercially sensitive or confidential data. Therefore our framework protects the information shared by the SMEs by using a DSA that can be attached to the information and will be enforced to protect such information. Options are provided to pre-process information in order to allow some analytics and, at the same time, to preserve the privacy and confidentiality of the data.

Some other features of our solution include:

- SMEs are able to choose the type of privacy and confidentiality controls that are appropriate for safeguarding their CTI data on the cloud-based analysis service, e.g. to select open access, or to apply data anonymisation techniques or even to use homomorphic encryption based techniques for very sensitive data.
- Through the availability of different data confidentiality and access options, SMEs can confidently share specific types of their CTI data via a gateway platform, even with non-trusted third parties.
- The framework can incorporate diverse techniques for supporting the protection of CTI data, as SMEs do not have to be aware of the inner workings of these techniques.
- The framework can also incorporate diverse techniques for analysing the shared CTI without the SMEs worrying about issues like information leakage, as this process is transparent for the SMEs.

## IV. SYSTEM ARCHITECTURE

The scenario consists of extending the use of a multi-tenant and cloud-based MSS that can be deployed and configured for either public or private cloud environments. SMEs can subscribe to this MSS to enable it to protect their cloud-hosted assets. As SMEs may host their data and applications on cloud platforms different from the one operating the MSS, the MSS can be configured and allowed to acquire the relevant CTI directly from the applications, services, or virtual machines that are being protected by it.

Figure 1 shows a high level view of the system architecture. The SMEs communicate with the MSS to manage the security of the applications and services running on their VMs deployed on different cloud platforms. The MSS enforces the security policies set by SMEs directly on the VMs, which is usually done via a security agent deployed in the VMs. The SMEs delegate the C3ISP framework to collect and process the CTI to a middleware called the C3ISP Gateway, which has the capability of collecting, processing and sending the CTI in a standardised format to the cloud-based C3ISP Service. The SME also delegates the enforcing of the DSA to the C3ISP Gateway, which will process the CTI according to the DSA, before sending it to the C3ISP Service.

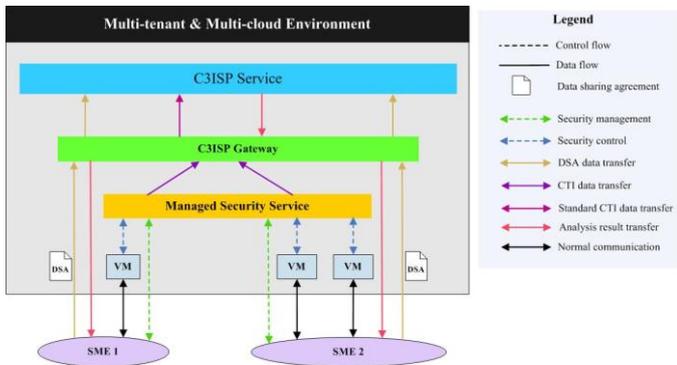


Figure 1. High level view of the system architecture

In the following sections, we discuss the detailed design of the main components of the system architecture using the FMC notation [7], which is a semi-formal but customised framework for describing the concepts and structures of complex informational systems.

**A. Managed Security Service**

SMEs subscribe to the MSS in order to configure, deploy and protect their assets. The process is illustrated below in Figure 2. The SME, by means of a browser, establishes a secure connection with the management portal of the MSS. Thereafter, the SME can subscribe to the MSS, register and customize VM protection and configure the MSS itself according to its requirements. Once the subscription is completed, credentials are issued to the SME to login to the Security Portal, enabling it to configure and activate/enable individual security services (anti-malware, IDS/IPS, firewalls, etc.) on its virtual machines.

The MSS deploys and enforces the security services and their policies by controlling an MSS Agent installed in the SME VMs. The MSS stores the CTI gathered by the agents on each virtual machine into a CTI database, accessible only from the C3ISP Gateway.

**B. C3ISP Gateway**

The C3ISP Gateway is the core component of the system architecture, which collects, processes and shares the CTI with the external C3ISP Service. It is illustrated in detail in Figure 3.

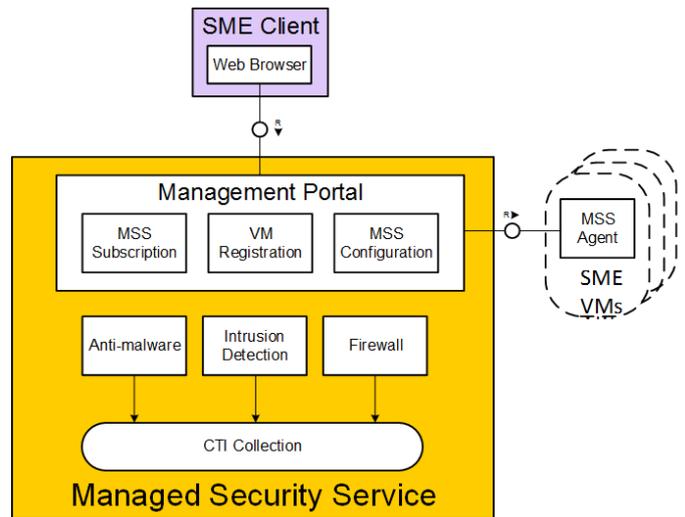


Figure 2. Managed Security Service architecture

The C3ISP Gateway retrieves the raw CTI from the MSS using the Tenant Manager. Once received, the raw CTI has to be formatted and processed by the C3ISP Gateway in accordance with the DSA [8-10] set by the SME. This is where the DSA Manager comes in, by letting the SMEs select DSAs by either configuring a set of pre-formulated policy templates [11,12], or even creating a completely new DSA. The DSA includes details regarding how the data can and cannot be used by the C3ISP Gateway and the C3ISP Service. For example, according to confidentiality levels specified in the DSA, the C3ISP Gateway is able to transform the CTI from raw format into plaintext STIX [15] format, and apply anonymisation [18] or homomorphic encryption techniques [17] on some of the raw CTI before its transformation. Once the DSA has been enforced, the C3ISP Gateway sends the CTI to the C3ISP Service.

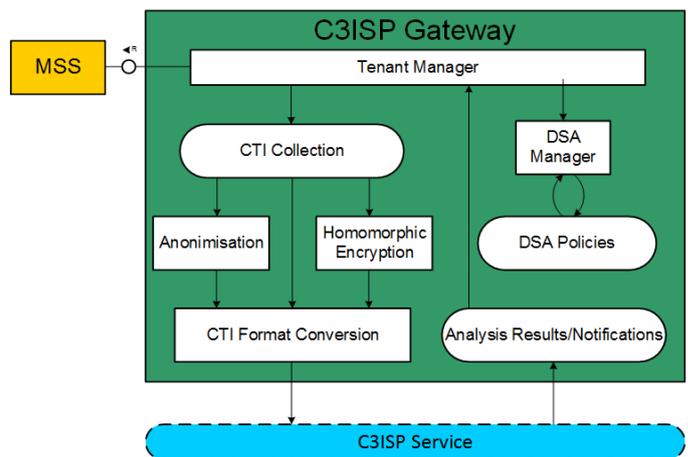


Figure 3. Information Processing Middleware architecture

Once the analysis is completed by the C3ISP Service, SMEs can retrieve the analysis results and outcomes, via the C3ISP Gateway.

## V. ARCHITECTURE ANALYSIS

The C3ISP Gateway acts on behalf of the SMEs and essentially takes care of all the security C3ISP Services for collecting, converting and submitting the CTI to the C3ISP Service. It also provides SMEs with capabilities for defining and customising DSA policies in a user-friendly manner, via a web interface, and provides a front-end for retrieving the analysis outcomes. The C3ISP Service communicates strictly only with the C3ISP Gateway, to make it more difficult for attackers to compromise it.

The MSS also takes care of SME needs, not only for deploying new services and applications but also, and more importantly for security matters. According to our architecture, only the MSS is allowed to handle the security services installed and configured in the VMs, thus avoiding misconfigurations. It also provides CTI to the C3ISP Gateway in a secure way which is transparent to the SMEs.

These platforms are designed according to a modular idea/concept, bringing many advantages. From a security point of view, if the C3ISP Gateway were compromised by attackers, the anonymised and encrypted CTI would not be compromised and the overall damage would be mitigated. Having a modular architecture is also more suitable if the technology regarding specific modules changes in the future, for example regarding how CTI is stored and converted. If the encryption or the anonymisation technique changes, there is no need to change the entire architecture and platform, but only the modules involved.

## VI. CONCLUSION

We have presented an original architectural model aiming to address the current limits of cyber security solutions. The proposal lets SMEs benefit from the same level of security that enterprises do today, via a Managed Security Services (MSS) solution. Our solution for SMEs employs three different platforms, the Gateway, the MSS and the Analytics service. There are several benefits in information sharing for cyber security (including incident notification) as well as several barriers to be removed. The benefits include, earlier detection of attacks, and that the analysis becomes collaborative, which should lead to more precise and informative outcomes. The information of one party can be of benefit to many others (thus leading to an increase of public good). Organisations fear the risk related to business reputation loss if the shared information reveals cyber security incidents which will be reported on public platforms. There are also other concerns related to the compliance with legislation (e.g., sharing of data involving personal information). Barriers of information sharing therefore include lack of trust in sharing data and lack of control over the data which has been shared, so it is important to provide privacy-preserving mechanisms of the type that we have proposed.

## ACKNOWLEDGMENT

We acknowledge financial support for this work provided by the European Commission's Horizon 2020 research and innovation programme under the grant agreement No. 675320 (NeCS) and 700294 (C3ISP).

## REFERENCES

- [1] Daniel, Joshua and El-Moussa, Fadi and Ducatel, Gery and Pawar, Pramod and Sajjad, Ali and Rowlingson, Robert and Dimitrakos, Theo Integrating Security Services in Cloud ServiceStores. In: Trust Management IX. IFIP Advances in Information and Communication Technology, 454. Springer International Publishing, pp. 226-239 (2015)
- [2] Yu, H., Powell, N., Stenbridge, D., Yuan, X.: Cloud computing and security challenges. In: ACM-SE 2012 Proceedings of the 50th Annual Southeast Regional Conference, pp 298-302 (2012)
- [3] Catteddu, D., Hogben, G.: Cloud Computing Risk Assessment. European Network and Information Security Agency (ENISA) (2009)
- [4] Dimitrakos, T.: Cloud Security Challenges and Guidelines. EIT ICT Labs Symposium on Trusted Cloud and Future Enterprises, Oulu, Finland. <http://www.eitictlabs.eu/news-events/events/article/eit-ict-labs-symposium-on-trusted-cloud-and-future-enterprises/> (August 2014)
- [5] European Commission on "C3ISP project". Available at: [http://cordis.europa.eu/project/rcn/202687\\_en.html](http://cordis.europa.eu/project/rcn/202687_en.html)
- [6] Anderson, R., Barton, C., Bohme, R., Clayton, R., Eeten, M.J.G, Levi, M., Moore, T., Savage, S., Measuring the Cost of Cybercrime, WEIS 2012, available at [http://weis2012.econinfosec.org/papers/Anderson\\_WEIS2012.pdf](http://weis2012.econinfosec.org/papers/Anderson_WEIS2012.pdf) (2012)
- [7] Knöpfel A, Gröne B, Tabeling P. Fundamental modelling concepts. Effective Communication of IT Systems, England. (2005)
- [8] Marco Casassa Mont, Iaria Matteucci, Marinella Petrocchi, Marco Luca Sbdio: Towards safer information sharing in the cloud. Int. J. Inf. Sec. 14(4): 319-334 (2015)
- [9] C. Brodie et al. The Coalition Policy Management Portal for Policy Authoring, Verification, and Deployment. In POLICY, pages 247-249 (2008)
- [10] C. Caimi, C. Gambardella, M. Manea, M. Petrocchi, D. Stella. Technical and legal perspectives in Data Sharing Agreements definition. Annual Privacy Forum (2015)
- [11] Matteucci, Iaria, Marinella Petrocchi, and Marco Luca Sbdio. "CNL4DSA: a controlled natural language for data sharing agreements." Proceedings of the 2010 ACM Symposium on Applied Computing. ACM (2010)
- [12] Martinelli, Fabio, et al. "A formal support for collaborative data sharing." Multidisciplinary Research and Practice for Information Systems: 547-561 (2012)
- [13] Deliverable D4.1: DSA Specifications, Methodologies, and Techniques, Coco Cloud EU FP7 Project, GA #610853
- [14] Deliverable D5.1: Enforcement Architecture and Communication Protocol. Coco Cloud EU FP7 Project, GA #610853
- [15] Barnum, Sean. "Standardizing cyber threat intelligence information with the Structured Threat Information eXpression (STIX™)." MITRE Corporation 11: 1-22 (2012)
- [16] NISP WG2 Plenary Report Information Sharing and Incident Notification available at: <https://resilience.enisa.europa.eu/nis-platform/shared-documents/3rd-plenary-meeting-april-2014/> (2014)
- [17] Carpov, Sergiu, Paul Dubrulle, and Renaud Sirdey. "Armadillo: a compilation chain for privacy preserving applications." Proceedings of the 3rd International Workshop on Security in Cloud Computing. ACM, (2015)
- [18] Chen, K., Liu, L., Privacy-preserving Multiparty Collaborative Mining with Geometric Data Perturbation, IEEE Transactions on Parallel and Distributed Computing, Vol XX, (2009)

- [19] BT Assure Threat Monitoring.  
[https://www.globalServices.bt.com/uk/en/products/assure\\_threat\\_monitoring/BT\\_Assure\\_Threat\\_Monitoring.pdf](https://www.globalServices.bt.com/uk/en/products/assure_threat_monitoring/BT_Assure_Threat_Monitoring.pdf)
- [20] SAP Enterprise Threat Detection.  
<https://wiki.scn.sap.com/wiki/display/Security/SAP+Enterprise+Threat+Detection+-+Security+Monitoring+-+Data+Breach+Protection>
- [21] HPE SIEM. <https://saas.hpe.com/en-us/software/siem-security-information-event-management>
- [22] McAfee Enterprise Security Manager.  
<https://www.mcafee.com/us/products/enterprise-security-manager.aspx>
- [23] AlienVault Unified Security Management. <https://www.alienvault.com/>
- [24] Collaborative and Confidential Information Sharing and Analysis for Cyber Protection (C3ISP). <http://c3isp.eu/>

# From E-government to Cloud-government: Challenges of Jordanian Citizens' Acceptance for Public Services

ABEER ALKHWALDI<sup>1,2,\*</sup>, MUMTAZ KAMALA<sup>1,\*\*</sup>, RAMI QAHWAJI<sup>1,\*\*\*</sup>

<sup>1</sup> School of Electrical Engineering and Computer Science  
Bradford University  
Bradford, United Kingdom

<sup>2</sup> Department of Management Information Systems, School of Business  
Mutah University  
AlKarak, Jordan

\* A.F.A.HamidAlkhwaldi@bradford.ac.uk , \*\* M.A.Kamala@bradford.ac.uk , \*\*\* R.S.R.Qahwaji@Bradford.ac.uk

**Abstract**—On the inception of the third millennium, there is much evidence that cloud technologies have become the strategic trend for many governments, not only for developed countries (e.g. the UK, Japan and the USA), but also developing countries (e.g. Malaysia and countries in the Middle East region). These countries have launched cloud computing movements for enhanced standardization of IT resources, cost reduction and more efficient public services. Cloud-based e-government services are considered to be one of the high priorities for government agencies in Jordan. Although experiencing phenomenal evolution, government cloud-services are still suffering from the adoption challenges of e-government initiatives (e.g. technological, human, social and financial aspects) which need to be considered carefully by governments contemplating their implementation. While e-government adoption from the citizens' perspective has been extensively investigated using different theoretical models, these models have not paid adequate attention to security issues. This paper presents a pilot study to investigate citizens' perceptions of the extent to which these challenges inhibit the acceptance and use of cloud computing in the Jordanian public sector and examine the effect of these challenges on the security perceptions of citizens. Based on the analysis of data collected from online surveys, some important challenges were identified. The results can help to guide successful acceptance of cloud-based e-government services in Jordan.

**Keywords**- Acceptance, Challenges, Cloud Computing, E-government, Jordan

## I. INTRODUCTION

One of the significant characteristics of this era is Information Communication Technology (ICT) and, like any new innovation, it has an important influence in terms of

changing people's lives to some extent. The evolution of ICT has dramatically changed citizen-government interactions, developing their expectations in this regard [1]. With the advent of e-government systems, many governments worldwide have moved to an electronic form of public administration to deliver high quality and more efficient services to its citizens [2]. However, with further use of e-government services more adoption challenges have emerged (e.g. technological and financial issues) [2]. Innovative ICTs, such as cloud computing, can contribute to solving these challenges; cloud technology represents a profound change in the technological construction of the whole public sector and the ways that governments conduct their business [3]. In the last two decades, the government sector focused predominantly on traditional web-based services to improve transparency, accountability and accessibility to public services and information. Guided by the pioneering initiatives of many developed countries, such as the UK, the EU, the USA and Japan [4], cloud computing has successfully progressed towards being the next generation of e-government services. The idea is to use highly scalable, ubiquitous, location-independent IT resources to improve organizational processes and reinvent the services that meet citizens' expectations, to improve collaboration between government agencies with more open, flexible, low-cost and unified computing. However, cloud-based e-government is considered a fundamental change within governments, it also represents a user-centric services platform aiming to increase citizens' participation. Recently, public sectors in the Middle East and other developing countries have started to gear toward cloud computing to achieve increased levels of performance and efficiency while offering cost-effective outcomes [5]. However, a number of these governments are still at a rudimentary stage. The Hashemite Kingdom of Jordan (HKJ), a country at the heart of the Middle East region, is in the process of making a complete transformation to cloud-government. Jordan recognised the fundamental role of cloud

computing in the e-government environment and launched the “National Cloud Platform”, designed to enable continued improvements and growth within e-government applications [6]. Currently, a number of ministries and government entities use cloud-based solutions to deliver improved public services to their citizens; however, a high-percentage of Jordanian people do not yet use cloud-based e-government services and still depend on paper printouts of their official transactions. Therefore, services provided using cloud technology, such as “Issuing Certificate of Non-Criminal Record”, are not exploited effectively. A number of models and theories of IT/IS acceptance, such as the Technology Acceptance Model (TAM), the Theory of Planned Behaviour (TPB) and the most recently developed model, the second version of the Unified Theory of Acceptance and Use of Technology (UTAUT2), have been largely used to examine individuals’ low adoption or reluctance to use new e-government services [7]. These models are criticised for not considering the constructs representing the specific perspectives of e-government (e.g. security). This paper takes a step toward answer the question: “Do the barriers and challenges of e-government adoption influence the use of cloud-based e-government services from the Jordanian citizens’ perspective?” This research identified that some e-government challenges still affect the acceptance of cloud-based public services, such as lack of awareness and security. In addition, to determine some of the security concerns relevant to the research context, these can be taken into account when formulating a new theoretical model.

## II. CHALLENGES OF E-GOVERNMENT

E-government initiatives aim to achieve a high level of government performance and provide citizens with improved public services. However, a number of researchers indicate the many difficulties faced by government organizations that obstruct the realization of e-government promised goals and degrade its successful adoption [4]. Therefore, the public sector has a responsibility to overcome these barriers. In spite of government efforts in this regard, the success of e-government initiatives are based significantly on citizens’ willingness to use and accept e-government services utilizing new ICT . One of the major determinants to using e-government services is security [8].

The challenges of e-government adoption experienced most often can be categorised as illustrated in Table 1.

Compared to a large number of studies that address e-government challenges from a citizens’ perspective, relatively limited research exists regarding the influence of these challenges on the adoption and acceptance of cloud-based services in the public sector. Also, the extent to which such challenges are relevant to the security of cloud-based e-government services. Thus, there is a concurrent need to gain empirical investigation for the impact of these challenges on

TABLE I. CATEGORIES OF E-GOVERNMENT ADOPTION CHALLENGES

Challenges	Examples	Ref.
<b>Technological</b>		
• IT infrastructure	<ul style="list-style-type: none"> <li>• Insufficient networking capacity</li> <li>• Inadequate integration across systems</li> <li>• Poorly updated hardware and software</li> <li>• Incompatibility and complexity of the existing systems</li> </ul>	[9]
• Security	<ul style="list-style-type: none"> <li>• Lack of transaction protection</li> <li>• Lack of trust in online and government e-services</li> <li>• Lack of security hardware in public sector</li> </ul>	[9-11]
• Availability	<ul style="list-style-type: none"> <li>• Inability to deliver services and information upon request</li> <li>• Slow response to citizens expectations, making unsuccessful delivery of e-services.</li> </ul>	[9, 12]
• Accessibility	<ul style="list-style-type: none"> <li>• Difficulty accessing the system for people with disabilities</li> <li>• Internet coverage is limited</li> </ul>	[9, 13]
• Website design	<ul style="list-style-type: none"> <li>• Limited languages to present the website content</li> <li>• Perceived ease of use</li> <li>• Perceived usefulness</li> </ul>	[10, 12]
<b>Human-aspects</b>		
• Lack of awareness	<ul style="list-style-type: none"> <li>• Lack of knowledge about e-government services and its benefits</li> <li>• Lack of orientation campaigns to promote e-government</li> </ul>	[10, 12]
• ICT skills	<ul style="list-style-type: none"> <li>• Lack of IT skills among users of e-government (i.e. citizens, employees, IT staff)</li> <li>• Lack of baseline knowledge related to secure online practices</li> </ul>	[9, 14]
<b>Social</b>		
• Culture	<ul style="list-style-type: none"> <li>• Religious and tribal beliefs</li> <li>• Language problems</li> <li>• Change resistance</li> </ul>	[13, 15]
<b>Financial</b>		
• Lack of budget / high cost	<ul style="list-style-type: none"> <li>• High maintenance and operational cost</li> <li>• High budget for security solutions</li> </ul>	[11]

the acceptance and security of cloud computing applications in the public sector. For this paper, an online survey was developed to achieve this aim.

## III. CLOUD COMPUTING

Cloud computing, which provides a highly scalable computing resource, has become a salient milestone in the development of information systems (IS) architecture and, more importantly, IT strategies for governments. According to the National Institute of Standards and Technology (NIST), cloud computing is an emerging model “for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications and services) that can be

rapidly provisioned and released with minimal management effort or service provider interaction” [16].

Cloud technology, perceived as one of the most promising information technologies today, has a number of inherent distinguishing characteristics including broad network access, pooled resources, on-demand self-service, rapid elasticity and measured service [16]. Wang et al. (2016) defined cloud computing as “the delivery of computing as a service rather than a product” [17]. This service is delivered to individuals, businesses and government agencies on three different levels, including Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). In addition, based on the usage range, cloud technology can be classified into four deployment models, namely private cloud, public cloud, community cloud and hybrid cloud [16, 17]. Security concerns become crucial during service delivery and affect users’ acceptance for those services .

#### IV. CLOUD COMPUTING IN THE E-GOVERNMENT CONTEXT

Evidence from the prior literature shows that cloud technology is adopted as a novel delivery channel for public services [18]. It can contribute to significant improvements in the performance of government sectors, also creating novel public services worldwide . Around the world, governments have begun to deliver their services depending on cloud computing applications and platforms to develop the service quality, reduce costs and realize efficiency [19]. Adopting cloud technologies to deliver public services provides several benefits, such as dynamic scalability, security management, distributed storage, accountability, and green IT [20]. However, there are a number of risks associated with cloud-based e-government services, including tangible risk (e.g. availability and infrastructure) and intangible risk (e.g. security) [3]. Security concerns are the main source of risk [21]. Therefore, it is necessary to identify and address such concerns for cloud-based e-government services along with other risks. Thus, government sectors will require the ability to undertake cloud risk-management, which is deemed as the main determinant of cloud computing success and acceptance.

#### V. ACCEPTANCE AND ADOPTION OF NEW TECHNOLOGY: THEORIES AND MODELS

Research on individuals’ IT acceptance covers one of the well-established streams in the field of information systems (IS) [22]. So far various competing models have been developed to understand IT/IS acceptance behaviour. Among these models are the TAM, TPB, the theory of reasoned action and so forth [23]. A comprehensive model that offers a more complete picture of the IT acceptance process of users was needed. Venkatesh et al. (2003) developed the new UTAUT model through integration and consolidation of eight dominant technology acceptance theoretical models [23]. The UTAUT model includes four

core determinants (performance expectancy, effort expectancy, social influence and facilitating conditions) of behavioural intention and actual usage behaviour. While gender, age, experience, and voluntariness have been constructed as moderators to the key relationships (see Fig.1) [23].

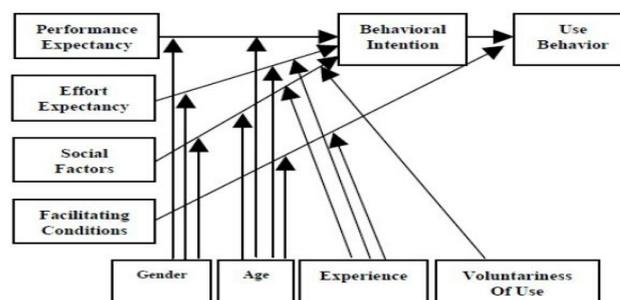


Figure 1. UTAUT [23]

Since its inception, the unified theory’s (UTAUT) relationships have been extensively confirmed in the context of e-government acceptance and use [7]. Hence, the UTAUT model is proper for understanding the acceptance of e-government services using cloud technology. According to Venkatesh et al. (2012), the UTAUT2 model, the extended version of the UTAUT model, is geared towards the consumer use and acceptance of IT [24]. This can serve the aims of this study better as the citizen is the main consumer of e-government cloud-based services . Moreover, compared to the original model, UTAUT2 provides a notable improvement with respect to explained variance ( $R^2$ ) up to 74 percent in behavioural intention (BI) to use technology. However, the unified theory tends to be limited to not taking into account security concerns and trust, which represents one of the specific constructs associated with the e-government context [2].

#### VI. RESEARCH METHODOLOGY

In this study, a quantitative research method using an online questionnaire survey was selected to meet the aims of the research. A survey questionnaire was employed as it is low cost with respect to both money and time, participants’ anonymity can be guaranteed and it has the ability to provide an inflow of data from a large study sample with minimum researcher bias [25]. After the initial design for the questionnaire, a pre-test was done by PhD researchers, academic staff and some of the Jordanian public. This was important to enhance questions and check participant comprehension before the distribution of the actual survey [26]. This resulted in some amendments to the wording of a few questions. The questionnaire was employed to determine what Jordanian citizens perceived to be challenges facing the acceptance of cloud-based e-government services, and the effect of these challenges on the perceived security. Because cloud computing is a relatively new concept in developing

countries, the researchers provided a definition of this new technology on the first page of the online survey. The online survey link was advertised to targeted respondents through various communication channels (e.g. personal emails, social media groups and university mailing lists). In general, the research sample can be classified as Jordanian citizens who are internet users and have a basic understanding of CC and e-government services. The questionnaire includes four sections: (1) general respondent information; (2) respondents' perceptions about challenges and barriers facing acceptance of cloud-based public services; (3) respondents' perceptions of the security of cloud-based public services; (4) respondents' experience in internet and e-government usage. The questionnaire was carried out following the University of Bradford human research ethics.

VII. DATA ANALYSIS AND FINDINGS

The next sections highlight the key results and provide indications to the answer to the research question, drawing on the survey findings. An overview of the online survey is presented in the first section. Then, illustration of the implications of the study question with more details is presented in the second section.

A. Overview of the Online Survey

As outlined above, the online survey consisted of four parts. It was presented in Arabic for better understanding of its questions, as Arabic is the native spoken language for Jordanian citizens. However, an English survey was also designed for the purpose of the research. Involvement in the questionnaire was entirely voluntarily and informed consent was obtained through the first question in the cover sheet indicating that consent is explicit by the "yes" response to the question. The survey could be accessed through the online survey website eSurveyCreator.com. It was available "online" for one month to all Jordanian people worldwide. During that time, 187 responses were received. However, 23 were discarded because of missing or incorrect answers. Thus, a total of 164 responses were valid for the data analysis process, in order to identify the extent to which e-government challenges affect the acceptance of cloud services in the public sector from the Jordanian perspective and its influence on perceived security. People who understood cloud-based services (e.g. university students, ICT sector employees) were qualified participants. This stems from the fact that they are among the adult groups for whom the use of the internet has become an essential part of their daily life, and they have the required knowledge about CC services. Furthermore, since the development of cloud-based services in Jordan is in the early stages, the respondents will be the main potential users. So, knowledge of their understanding and perceptions is crucial to improving cloud-based e-government services. The analysis results are described below.

B. General (Demographic) Information:

Table 2 shows the respondent sample who answered the first section of the questionnaire by providing personal

information. The majority of respondents were male (108), with 56 female respondents. The bias towards males is due to the fact that Jordan is considered a male-dominated society where the majority of elements of women's lives are at the mercy and authority their male relatives, this could affect various aspects such as conducting government transactions.

TABLE II. GENERAL (DEMOGRAPHIC) INFORMATION FROM THE SURVEY

Characteristics	Percentage(%)
<b>Gender</b>	
Male	66%
Female	34%
<b>Age</b>	
18-20	3%
21-30	68%
31-40	16%
41-50	11%
50+	2%
<b>Education Level</b>	
Secondary school or below	7%
Diploma	8%
Bachelor	63%
Postgraduate	22%
<b>Residency Country</b>	
Jordan or developing countries	78%
Developed countries	22%
<b>Security Awareness Level</b>	
Beginner	57%
Medium	32%
Advanced	11%

The subjects' ages mainly ranged between 21 and 30 years old (68%), consistent with the highest majority of internet users in Jordan. Most subjects had Bachelor's degrees (63%) which lies in agreement with the general distribution of Jordanian internet users. Seventy-eight percent of all subjects lived in Jordan or other developing countries, while 22% lived in developing countries such as the UK or USA. This overview of participants' general information will help in the results interpretation and answering the research question in the following sections.

C. Challenges to Cloud-Based E-Government Services Acceptance: Interpretation of the Study Question

According to the results, there are many barriers and challenges to e-government adoption and acceptance (e.g. technological, human, social and financial aspects). These inhibit the acceptance and use of cloud computing in the public sector from the Jordanian citizens' perspective (see Fig. 2). Table 3 lists these barriers, ranked based on the percentage of respondents who consider the challenge as either important or very important.

TABLE III. BARRIERS TO ACCEPTING CLOUD-BASED SERVICES

Rank	Barrier	Percentage of respondents
------	---------	---------------------------

1	Lack of awareness	63.4%
2	Security	61.2%
3	Culture	53.7%
4	IT infrastructure	44%
5	Website design	41%
6	Accessibility	31.8%
7	IT skills	27.6%
8	Availability	25%
9	Lack of budget / high cost	17%

2	Social relations and culture have a significant influence on the security of "cloud-based e-government" services	81.1%
3	Lack of security awareness is one of the main determinants of the user's perception regarding the security of "cloud-based e-government" services.	76.3%
4	Perceived security is a significant resource for public users' trust of "cloud-based e-government" services.	70.6%
5	The design of the "cloud-based e-government" website influences the users' perception of its security.	60.1%
6	There is a lack of security guidelines for using "cloud-based e-government" services, on the government website, social media or other media channels.	56.8%
7	There is a lack of regulations and policies to use cloud public services.	44.9%

Promotion is one of the most important factors for successful initiatives in the e-government context. For any new technology (e.g. cloud computing) there are a number of steps to encourage and convince citizens to use and accept it. Therefore, government's promotional activities will contribute significantly to accomplishing this aim. From the above table, it is evident that more than 63% of respondents cited the lack of awareness of cloud services and its advantages in the e-government context, as the number one barrier to the use and acceptance of cloud-based e-government services.

Security concerns are a serious technical obstacle identified in this study and are a well-documented issue for e-government systems adoption and implementation worldwide [8, 19, 27]. More than 61% of the participants in this research indicated security to be a significant concern, making it the second-ranked challenge to cloud-based e-government use and acceptance.

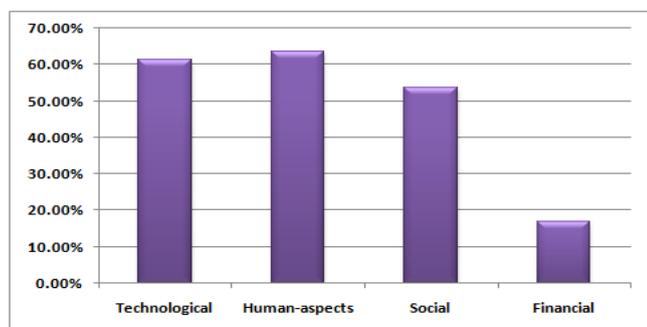


Figure 2. Challenges of Cloud Computing in the Public Sector

#### D. Security Perceptions of Cloud-Based E-Government Services

A number of e-government challenges noted in the survey influence the security perception of citizens towards cloud services. These were ranked based on the percentage of responses that indicated either strongly agree or agree (see Table 4).

TABLE IV. SECURITY BARRIERS TO ACCEPT CLOUD-BASED SERVICES

Rank	Barrier	Agreement
1	High level of security concerns regarding "cloud-based e-government" services are inspired by non-technical aspects (e.g. culture and awareness)	83.6%

Succeeding cultural inertia is one of the major difficulties of e-government use and adoption in developing countries [15]. This cultural issue has a different effect in developing and developed nations. Illustrating this, the percentage of respondents in Jordan and other developing countries "who prefer to conduct online transactions" is unequal to those in developed nations (see Fig. 3). The responses indicate that the primary reason is cultural differences, one of the participants living in Jordan commented: "I've heard about many people whose money was stolen when they tried to buy online" while a second said: "Electronic transactions are not guaranteed, I prefer to contact the other person face-to-face and have a chance of discussion". Therefore, it is important to train and educate citizens and draw their attention to the advantages, benefits and the facts about online transactions, in particular, cloud-based services for the government sector, through many promotional ways.

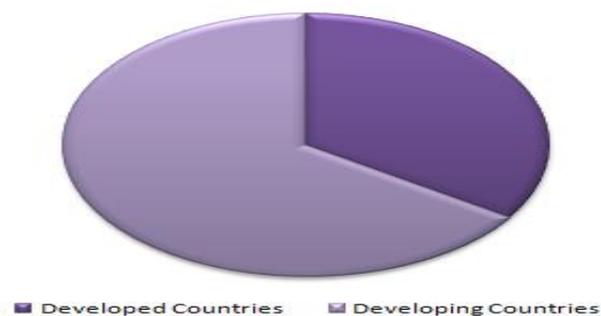


Figure 3. Preference for Using Online Transactions

#### VIII. CONCLUSION AND FUTURE WORK

Cloud-based service use is a current trend for e-government systems around the world. Governments, particularly in developing countries, can refer to the findings of this research to design and promote their services. For example, they can design various promotional programs for various users.

It is suggested to invest money in attractive and widespread awareness campaigns to encourage effective acceptance of cloud services. Also, people who conduct governmental transactions using cloud-services would have a discount on the transaction fees. This will help to increase the usage rate, especially for those who still refuse to use e-government services in general. The main objective of this paper is to draw upon the current perspectives of Jordanian citizens regarding the way in which e-government challenges affect the acceptance of cloud services in the public sector and to what extent these challenges influence the perceived security. This was conducted to support the decision-making process by the stakeholders of e-government, and to enhance cloud-based e-government services outcomes for Jordanian society. The challenges were identified through a relevant literature review and investigated by carrying out a survey. One of the significant findings is that security concerns play an important role in the acceptance of cloud-services in the government context (second rank), and hence an acceptance model would be required to take this into account. In addition, the non-technical aspects have a prominent importance among other security issues in the context of the study. Therefore, consideration of these challenges and their integration into the technology acceptance model is required for future research work. Due to its high explanatory power compared to the previous models and theories of IT acceptance, and also its appropriateness to e-government cloud-based services, the UTAUT2 model would be a superior selection as the base for the future proposed theoretical model to address the acceptance of these services.

## REFERENCES

- [1] E. Abu-Shanab, "Antecedents of trust in e-government services: an empirical test in Jordan," *Transforming Government: People, Process and Policy*, vol. 8, no. 4, pp. 480-499, 2014.
- [2] Y. K. Dwivedi, N. P. Rana, M. Janssen, B. Lal, M. D. Williams, and M. Clement, "An empirical validation of a unified model of electronic government adoption (UMEGA)," *Government Information Quarterly*, 2017.
- [3] S. Paquette, P. T. Jaeger, and S. C. Wilson, "Identifying the security risks associated with governmental use of cloud computing," *Government Information Quarterly*, vol. 27, no. 3, pp. 245-253, 2010.
- [4] T. Almarabeh, Y. K. Majdalawi, and H. Mohammad, "Cloud Computing of E-Government," ed: Communications and Network, 2016, 8, 1- 8, 2016.
- [5] N. N. Almutairi and S. F. Thuwaini, "Cloud Computing Uses for E-Government in the Middle East Region Opportunities and Challenges," *International Journal of Business and Management*, vol. 10, no. 4, p. 60, 2015.
- [6] Microsoft, "The Hashemite Kingdom of Jordan - Microsoft Customer Stories," MSFT\_Business, <https://enterprise.microsoft.com/en-us/customer-story/industries/government/the-hashemite-kingdom-of-jordan/>, [Accessed: 30 June, 2017].
- [7] V. Weerakkody, R. El-Haddadeh, F. Al-Sobhi, M. A. Shareef, and Y. K. Dwivedi, "Examining the influence of intermediaries in facilitating e-government adoption: An empirical investigation," *International Journal of Information Management*, vol. 33, no. 5, pp. 716-725, 2013.
- [8] D. Belanche-Gracia, L. V. Casalo-Arino, and A. Pérez-Rueda, "Determinants of multi-service smartcard success for smart cities development: A study based on citizens' privacy and security perceptions," *Government information quarterly*, vol. 32, no. 2, pp. 154-163, 2015.
- [9] A. M. Odat, "E-GOVERNMENT IN DEVELOPING COUNTRIES: FRAMEWORK OF CHALLENGES AND OPPORTUNITIES," *Journal of Theoretical and Applied Information Technology*, vol. Vol. 46 No.2, pp. 1013-1021, 2012.
- [10] S. A. Alateyah, R. M. Crowder, and G. B. Wills, "Identified factors affecting the citizen's intention to adopt e-government in Saudi Arabia," 2013, p. 904: World Academy of Science, Engineering and Technology (WASET).
- [11] A. Savoldelli, C. Codagnone, and G. Misuraca, "Understanding the e-government paradox: Learning from literature and practice on barriers to adoption," *Government Information Quarterly*, vol. 31, pp. S63-S71, 2014.
- [12] M. Rehman and V. Esichaikul, "Factors influencing the adoption of e-government in Pakistan," in *E - Business and E -Government (ICEE), International Conference*, pp. 1-4, 2011, pp. 1-4: IEEE.
- [13] M. E. Alzahrani and R. D. Goodwin, "Towards a UTAUT-based Model for the Study of EGovernment Citizen Acceptance in Saudi Arabia," *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, vol. 6, no. 4, pp. 376-382, 2012.
- [14] S. Furnell and L. Moore, "Security literacy: the missing link in today's online society?," *Computer Fraud & Security*, vol. 2014, no. 5, pp. 12-18, 5// 2014.
- [15] M. Ali, V. Weerakkody, and R. El-Haddadeh, "The impact of national culture on e-government implementation: A comparison case study," 2009.
- [16] P. Mell and T. Grance, "The NIST definition of cloud computing," 2011.
- [17] N. Wang, H. Liang, Y. Jia, S. Ge, Y. Xue, and Z. Wang, "Cloud computing research in the IS discipline: A citation/co-citation analysis," *Decision Support Systems*, vol. 86, pp. 35-47, 2016.
- [18] K. K. Smitha, T. Thomas, and K. Chitharanjan, "Cloud Based E-Governance System: A Survey,"

- Procedia Engineering*, vol. 38, pp. 3816-3823, 2012/01/01/ 2012.
- [19] D.-H. Shin, "User centric cloud service model in public sectors: Policy implications of cloud services," *Government Information Quarterly*, vol. 30, no. 2, pp. 194-203, 2013.
- [20] A. Tripathi and B. Parihar, "E-governance challenges and cloud benefits," in *IEEE International Conference on Computer Science and Automation Engineering (CSAE), 2011*, 2011, vol. 1, pp. 351-354: IEEE.
- [21] D. Zissis and D. Lekkas, "Securing e-Government and e-Voting with an open cloud computing architecture," *Government Information Quarterly*, vol. 28, no. 2, pp. 239-251, 2011.
- [22] A. Alkhwaldi and M. Kamala, "Why Do Users Accept Innovative Technologies? A critical review of technology acceptance models and theories," *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, vol. Vol. 4 Issue 8, August - 2017, pp. 7962-7971, 2017.
- [23] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS quarterly*, pp. 425-478, 2003.
- [24] V. Venkatesh, J. Y. L. Thong, and X. Xu, "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology," 2012.
- [25] D. E. Gray, *Doing research in the real world*, Third ed. (no. Book, Whole). London: SAGE Publications Ltd, 2014.
- [26] M. Saunders, P. Lewis, and A. Thornhill, *Research methods for business students*, Seventh ed. (no. Book, Whole). Harlow, Essex: Pearson Education Limited, 2016.
- [27] J.-W. Lian, "Critical factors for cloud based e-invoice service adoption in Taiwan: An empirical study," *International Journal of Information Management*, vol. 35, no. 1, pp. 98-109, 2015.

# Data Subject Rights in the Cloud

A grounded study on data protection assurance in the light of GDPR

Alaa Altorbaq<sup>\*</sup>, Fredrik Blix<sup>†</sup>, Stina Sörman<sup>‡</sup>

Department of Computer and Systems Sciences  
DSV, Stockholm University  
Stockholm, Sweden

Email: <sup>\*</sup>alaa@dsv.su.se, <sup>†</sup>blix@dsv.su.se, <sup>‡</sup>stina.sorman@kpmg.se

**Abstract**—The new European General Data Protection Regulation GDPR requires that organisations placing personal data on cloud services can protect certain rights of data subjects, such as their right to access, demand erasure and rectification of their data. Due to the technical complexity and shared environment of cloud services, the flow of personal data in the cloud must be secured and controlled from its initial collection, via processing, to final erasure. This has proved to be both organisationally and technically challenging. This study identifies the related challenges and outlines potential solutions for organisations who need to be able to adequately demonstrate compliance with the regulation as well as to respond to rights requests from data subjects. The study is based on interviews with ten data protection experts. The semi-structured interviews were qualitatively analysed, using an approach informed by Grounded Theory. The contribution of this study is a refined model depicting stages of a personal information life-cycle. Additionally, twelve challenges and fourteen recommendations were identified and presented to the various stages of the model. Both clients and providers of cloud services are expected to benefit from these results, as well as the data subjects, whose rights are protected.

**Keywords;** *Data Protection, Cloud Services, Privacy, Information Security.*

## I. INTRODUCTION

Along with the irrevocable development of information technology and the immense amount of personal data processed in today's interconnected world, legal changes are taking place within the European Union. On 25 May 2018, the European Union's General Data Protection Regulation enters force [1], overriding the current Directive 95/46/EC [2]. The repeal holds a more rigorous approach to the safeguard of personal data, to ensure that technical and organisational measures are implemented in an efficient and adequate manner. The approach holds the controller of the data (the organisation which determines the purposes and means of the processing of personal data) accountable for demonstrating compliance with the Regulation. Non-compliance with the Regulation may result in an administrative fine reaching up to EUR 20 000 000 or 4% of the total global annual turnover, whichever is highest [1, Article 83(5)]. In parallel, businesses that experience an increasing demand for constant availability and accessibility, are increasingly adopting the cloud computing technology as an alternative to traditional outsourcing models, to enjoy

advantages such as broad access, elasticity and lower total cost of services [3]. However, the adoption of cloud services for the processing of personal data comes with additional data protection risks. With significant business processes transitioned to a third party, the risk of declined control over the data flow becomes substantial [4]. With technical solutions facilitating the outsourcing of data storage and new legal requirements for data controllers to demonstrate protection of data subject rights, organisations relying on the cloud computing technology are put into a challenging situation. On the one hand, the control over data is increasingly transferred to a third party. On the other hand, the controller is increasingly accountable for ensuring that the data subjects are in control of their personal data, from collection to deletion. Through all stages of a personal information lifecycle, from collection through storage, deployment, use and finally deletion of information, the control over data shall remain intact. However, cloud computing can be viewed as one of the greatest threats to privacy. By outsourcing the personal data to a cloud service provider (CSP), the physical control of the hosted data will reside with the CSP, but the accountability for non-compliance will fall on the CSP client [4, p. 14].

Previous research has highlighted the data protection risks tied to outsourcing personal data to the technically complex cloud computing environments (e.g. [4]-[6]). As the GDPR introduces new rights and requirements on organisations, new challenges are introduced to organisations deploying cloud computing as a business model for the processing of personal data. While there is an abundance of research on privacy in relation to the cloud in general [4]-[11], there is a lack of knowledge regarding what organisations in practice find challenging and how they plan to solve issues emerging when trying to achieve compliance with the new Regulation in a cloud computing environment.

The purpose of this paper is to provide organisations with basic guidelines on how to ensure that data subject rights requests are effectively handled in a cloud computing environment so that they can adequately demonstrate compliance with the General Data Protection Regulation. In other words, the research question is "How should data subject rights be managed in a cloud computing environment to comply with the General Data Protection Regulation?"

## II. BACKGROUND

### A. Data protection principles and goals

The GDPR aims to reinforce the protection of personal data across all countries of the European Union (EU). The objective is to strengthen the fundamental human rights to self-determination and protection in relation to the processing of personal data, corresponding to the right to respect for private life mentioned in Article 8 of the European Convention on Human Rights (ECHR) [12]. Such an approach will ensure that the data subjects remain in control over their personal data [1, Article 4(7)]. To achieve data protection of individuals, the regulation emphasizes the following core data protection principles: ‘lawfulness, fairness and transparency’ of processing; ‘purpose and storage limitation’, ‘data minimisation’, ensured ‘data quality and accuracy’, and; ‘integrity and confidentiality’ presented in [1, Article 5-7]. To enhance the EU resident’s data protection rights, third country transfers are restricted to countries fulfilling an adequate level of data protection, allowing only for a few well-specified deviations [1, Article 44-49].

As a cross-cutting principle, ‘accountability’ expresses that the controller is responsible for demonstrating compliance with the principles mentioned previously. Moreover, this implies that controllers are obliged to only contract processors (organisation that processes data on behalf of data controller) providing sufficient guarantees to implement appropriate technical and organisational measures as declared in [1, Article 29]. In return, the processor is obliged to act only upon the documented instructions of the controller, reflecting [1, Article 28 (3)(a)]. Consequently, engagement of sub-contracted processors should be governed by the written authorisation of the controller. Thus, although the controller is responsible for demonstrating compliance with the principles mentioned above, the processor remains responsible for providing security and data protection safeguards throughout the processing activities. Moreover, while GDPR adheres to the traditional security goals of confidentiality, integrity and availability [5], these goals do not entail the protection of the data subjects’ ability to exercise their right of control over their data, regarding *e.g.* notification, portability and erasure. Therefore, three complementary data protection goals were added: unlinkability, transparency and intervenability. [13]

### B. Personal Information Lifecycle

The GDPR gives rise to the need for information lifecycle management, by asserting the right of the data subject to control what personal data is processed about him or her, in a transparent and accessible manner [1, Article 12]. In sum, the data subjects are the owners of their personal data and shall be granted the right to remain in control of their information from collection to deletion. The personal information lifecycle was defined by [5] as a comprehensive model, including all components “throughout the entire lifecycle of the information, from birth until death, [covering] its path from the source entity of information to diverse destination entities”

[5, p. 388]. In Table I, we suggest an adjusted taxonomy in the column presented as *Personal Information Lifecycle (PIL)*. The suggested PIL in the far-right column synthesises the “Information Lifecycle” [16] and fits it to the essentials of the processing of personal data as found in the GDPR: Collection, Storage, Disclosure and Distribution, Use and Retention, and Destruction.

TABLE I. THE PERSONAL INFORMATION LIFECYCLE

PERSONAL DATA PROCESSING MAPPING		
Information Lifecycle	Processing in GDPR	Personal Information Lifecycle (PIL)
Collection and Receipt	Collection, Recording	Collection
Storage	Organisation, Structuring, Storage	Storage
Distribution and Transmittal	Consultation, Disclosure of Transmission, Dissemination	Disclosure & Distribution
Access and Use	Use, Alignment/Combination	Use and Retention
Maintenance	Adaptation/Alteration, Retrieval, Restriction	
Disposition and Destruction	Erasure/Destruction	Destruction

### C. Cloud Service Providers as Processors

Cloud computing can be divided into different deployment and delivery models [9]. According to Article 29 Working Party [14], in all the delivery models; IaaS, PaaS and SaaS, the cloud service provider (CSP) is considered a data processor as defined in the GDPR. From a data protection perspective, the CSP serves as a processor when it processes personal data for and under the authority of the cloud service customer’s instructions (ISO/IEC 27018: 2014) [15]. In many organisations, cloud computing services are delivered through a third-party provider, which, to a varying extent, owns the entire infrastructure. Regardless of the cost, time and market benefits received from the on-demand computing service, the cloud has been recognised as tied to numerous data protection risks [8].

## III. METHODOLOGY

### A. Research Approach

The study undertakes a qualitative approach to address the full complexity of the research aim and to generate exploratory, descriptive knowledge on the data protection challenges controllers face, as well as suggested management efforts when ensuring individual rights in a cloud environment. Due to the new character of the studied issue, and an absence of a solid scheme on how to handle the upcoming regulatory challenges, most notably the subjects’ rights under the regulation, Grounded Theory was found to be most appropriate approach for data gathering and analysis, particularly adhering to the Straussian strand of grounded theory [19], [20].

## B. Data Collection and Analysis

Data was collected through Semi-Structured Interviews. For a Grounded Theory study, open-ended questions are advised, as they focus the interview questions on inviting a detailed discussion, allowing the researchers to attain in-depth, personalised information [18]. The interview questions were designed to first capture the context of the interview subject, through broad questions, such as what role and relation to cloud services the interviewee had. These were followed by questions specified for four roles: controller, processor, consultant and data protection authority. In total there were ten experts interviewed. The questions were continuously reflected upon, and throughout the project adjusted, based on answers. Theoretical sampling directed the data collection process where participants were selected to include as many members associated with the situation, to maximise the representativeness of the sample.

The sample included five controller representatives from organisations in different sectors (retail, airline, telecommunication), as well as three processor representatives, a cloud consultant and representatives from a national Data Protection Authority (DPA). We reached saturation when new data no longer added or improved the overall “theory”. [19]. To achieve a systematic analysis, without interfering with the rigour of the data, data collection and analysis proceeded concurrently involving a constant comparative analysis. The data analysis followed the steps defined by Strauss and Corbin [20], where data was manually coded through open, axial and selective coding.

## C. Ethical Considerations

This study was designed to include pseudonymity, confidentiality and informed consent into standpoint through the collection up to reporting. Participation relied upon the participant’s willingness and awareness of the research’s purpose, the researchers involved and their role in it.

## IV. FINDINGS

The outcome of our study highlights *twelve challenges and fourteen recommendations*, acknowledged by the interviewed experts, to fit into the Personal Information Lifecycle phases (See *Figure 1*).

### A. Prior to Processing

Certain aspects need to be taken into consideration already before the controller undertakes the responsibility for processing someone else’s data. Still, controllers are faced with a few of challenges, which include:

**C.1 Understanding and Engagement:** To be able to grant the data subject’s rights to data protection, a prerequisite is to be aware of that the rights exist, what they mean, and what risks non-compliance imply. Although most rights to the data subject are not new, several of the respondents emphasised that, for those who did not adhere to the data protection

principles under the current legislation, the journey towards compliance would be extensive. Organisations not adhering to similar regulations may thus face larger compliance challenges, due to lacking understanding and awareness regarding data protection rights.

**C.2 Legacy and Unstructured Data:** When the GDPR enters force, the Regulation applies to personal data also collected before its enforcement, implying that data subject right shall be performed on all personal data, regardless of when it was collected. Therefore, it is recommended prior to processing to ensure the following are in place:

**R.1 Effective Privacy Governance:** To successfully respond to data subject requests, several changes across the organisational structure are needed. Prior to addressing those updates, the controller must maintain an effective Privacy Governance Framework embedded within their existing processes and procedure. Such governance will contribute to administering and managing all privacy activities required to achieve and maintain compliance successfully. Additionally, an ownership model can be developed tied to a purpose, helps to ensure that the responsibility for data protection spans beyond the Data Protection Officer to be mainstreamed into the organisation.

**R.2 Overview and Control:** To make sure that the data rights are adequately adhered to, it was further emphasized that, if data was collected before an existing data protection governance structure, this data should be mapped to enable overview and control over all personal data. Structured as well as unstructured data, should be incorporated into the data protection governance structure. If large volumes of unstructured data were collected without legitimate purposes, the drastic action of starting from scratch, in some cases, might be required as the most effective option.

**R.3 Awareness:** Operational decisions taken on a day-to-day basis may jeopardise the rights of the data subject if they are not adhering to the data protection principles. Therefore, it is imperative that whoever comes into contact with personal data are aware of what the requirements are and that employees adhere to data protection policies.

### B. Collection and Storage

While the importance of structures and strategies applies to all states of the lifecycle, certain aspects become imperative during the collection phase, as well as for the decision on where to store personal data. During these stages, companies are generally faced with the following challenges:

**C.3 Consent and Opt-out Rights:** When storing the data in the cloud, one of the key issues is the communication between the controller, the processor and the data subject when intervening with processing. Although the data subject approves to data processing, consent may be withdrawn at any time, adding another layer of difficulty especially when processing is outsourced for instance, to a cloud service provider. A related concern is how to design the consent forms for customers to adequately be informed and satisfied with the

conditions, as withdrawal of consent implies deletion when consent constitutes the legal ground for processing.

**C.4 Service Availability and Storage Location:** To achieve service availability, cloud service business continuity depends heavily on continuous data copy and backup, which may entail personal data. While the cloud service does so to provide integrity and availability, backups add complexity to the request process, as the change must reflect on all copies – in every location.

Regardless of the designated challenges, the collection and storage phase are requiring following recommendations:

**R.4 Data Protection Risk Assessments:** In the context of cloud computing, the first question to consider is whether to embrace the cloud as service storage or not. The decision should depart from a risk-based approach concerning the potential impact for the data subject. For instance, the possible privacy impact and sensitivity of the data are assessed when considering data storage. The sensitivity of the data and the consequences for the data subject if the rights are not granted should be attentively assessed.

**R.5 Data Minimization:** A key practice explicitly stated as a data protection principle in the GDPR is to minimize data. Simply put, all data that is not necessary is not to be collected, and if it is already collected, it is required to be erased.

**R.6 Granularity:** As affirmed by several respondents, the data subject is granted the right to repetitively opt-in and opt-out to processing. If consent conducts the legal ground for processing, deletion of personal data shall consequently occur as soon as a person opt-out to processing. If processing is based on consent, the consent forms are therefore suggested to be designed in such a manner that the data subject can allow processing for very specific purposes while objecting to processing for other purposes. Therefore, all processed data needs to be tagged and consented with an elevated level of granularity.

**R.7 Anonymization:** Due to the scalability of the cloud, questions regarding big data and profiling tend to emerge. If making use of cloud services for analytical or profiling purposes, an important question to ask is whether the data, on which the analysis is conducted, needs to be personal. In many situations, organisations are not interested in the individual behaviour when performing big data analysis. Therefore, it might be advisable to perform anonymization right from the start to avoid profiling personal information.

**R.8 Choosing the Right Service Model and Location:** When deciding to store data in the cloud, it is imperative to choose the right model for the data. For instance, while the control over the infrastructure is distributed to the cloud provider if deploying an IaaS solution, the control over the platform layer and the application layer is also transferred in a SaaS solution. On the other hand, the more control is distributed to the cloud provider the more support can be expected from the cloud provider, if constituted by a trustworthy and competent entity. Eventually, it's important to note is that the controller remains

accountable for granting the data subject's rights regardless of what solution is chosen.

### C. Disclosure and Distribution

If the decision to process personal data in the cloud is taken, certain risks and challenges materialise and should be taken into consideration, to ensure that data subject rights can be responded to in an accurate manner. The below are the reported challenges:

**C.5 Flexibility and Standard Agreements:** A key concern when migrating to the cloud, is that the cloud service, unlike traditional outsourcing, cannot be tailored to specific needs. An additional concern related to the large CSPs standard agreements is a tendency to lack the degree of details necessary for getting insight in how personal data is processed.

**C.6 Trust, Transparency and Verification:** To prove compliance, controllers should deploy mechanisms for verification and reporting requests performed by data subjects. However, demonstrating compliance, especially with negative rights, such as deletion, may be obstructed by the lack of transparency of the CSP.

**C.7 Roles, Responsibilities and Expectations:** All digital processing of data relies on the interaction between hardware, middleware and software. However, the responsibility for securing the components is shared in the cloud. While the responsibility for securing the hardware relies upon the cloud provider, the responsibility for the application layer remains at the controller. Unfortunately, the relationship between the processor and the controller are recognised as ambiguous.

**C.8 Maturity:** Stated by several controllers, the data protection maturity may vary between the cloud providers. Some Cloud providers don't understand what efforts are required to comply with the GDPR or it isn't of their concern.

**R.9 Due Diligence, Clear Instructions and Audit Rights:** If deciding to migrate to the cloud, transparency between the controller and processor requires detailed control mechanisms. Efforts, such as communicating findings from external audits and the location of data centres are further suggested as relevant for enhancing the transparency and trust between the CSP and the customer on how personal data is processed. Making sure that clauses exist regarding the processing conditions (limitations), auditing terms, liabilities and terms of confidentiality, along with a detailed description of data location, data control, safeguards, guarantees and expectations are essential.

**R.10 Confirm Responsibilities and Expectations:** As certain anomalies in the interpretation between where the responsibilities of the processor and the controller meet or overlap are noted, ensuring that expectations, roles and responsibilities are consolidated and agreed upon from both the controller side and the processor side can facilitate the execution of data subject rights and mitigate the risk of negligence.

#### D. Use and Retention

When deploying a cloud solution for the processing of personal data, challenges specific to the cloud emerge.

**C.9 Conflicting Interests of Usage:** When deploying a cloud solution for the processing of personal data, it becomes imperative to understand how the data will be used in the cloud, either by the CSP or by its subcontractors. Processor state that data is processed only under the instructions of the controller. However, the DPA stated that, what might be an issue when using cloud services for the processing of personal data, is that CSPs “tend to be eager to make use of the data they store to develop their services”. While CSP lacks the motivation to illicitly process data for purposes other than those requested by the controller, the CSP may seek legal grounds for secondary use, for instance, in its standard agreements.

**C.10 Big Data and Profiling:** One of the main advantages of using the cloud environment is that it offers businesses the opportunity to store significant amounts of data for analysing it through big data queries, leading eventually to profiling. To perform big data analysis, organisations need to address the technical feasibility of ‘tagging’ every unit of personal data in the big volume of accumulated data to remove it subsequently.

To ensure usage of personal data does not result in non-compliance with the GDPR. Controllers are advised to:

**R.11 Specify Access Rights and Instructions for Processing:** Despite that CSPs presumably lack incentive to illicitly process personal data, the cloud provider may seek legal grounds for secondary use of data in its contractual agreements. Controllers must provide the CSP with clear instructions on the acceptable use of data and who should have access to it. Seeing that the data subject is entitled to know for what purposes, where, how and by whom the processing will be performed. Ensuring transparency is vital by regarding how and by whom, for instance, which sub-contractors are used and how they are entitled to use and process data was. If the cloud provider or the subcontractor(s) of the CSP is entitled to use the data for other purposes than the original purpose, make sure to communicate for what purposes the data is used and that the purpose rests upon legal grounds, such as consent.

**R.12 Ensure that Profiling is ‘Done Right’:** Principles such as data minimisation and transparency of processing also apply to big data pools. One way of circumventing issues emerging when collecting extensive amounts of data is to provide anonymization. If the data loses its value when no longer personal, other measures should be put in place. Encrypting data stored in the cloud is one solution, which however puts requirements on adherence to the current state of the art of cryptography.

#### E. Destruction

As the last stage of the Personal Information Lifecycle, personal data should be deleted. This stage may, in a cloud computing environment occur for multiple reasons, such as part of an exit strategy or the lack of legitimacy of processing.

While data destruction is a crucial part of maintaining compliance, it comes with several challenges, including:

**C.11 Technical Complexity:** Different platforms require different data erasure actions towards different types of requests. If no legal ground remains for the processing of personal data, deletion is required. In a simple case, delinking data subject information can be sufficient to fulfil the request. In other situations, a complete wipe out is legally required, especially in cases where the data was unlawfully processed or collected. Regardless, adhering to data erasure requests demands track of each instance of personal data collected at distributed locations. Interconnectivity of systems is further seen as a challenge, as deletion of data in one system may cause ripple effects in other systems.

**C.12 The Right to Erasure:** It has been established that the term ‘erasure’ is ambiguous. Some companies are not sure what erasure means or what it entails. It was also emphasized by the DPA that the right to erasure is a right, which may add inconsistency to what it implies: “the right to erasure is a right, implying that it’ll depend somewhat on what the data subject requests.” Furthermore, verifying deletion to demonstrate compliance may become even more challenging in the cloud, much relating to transparency issues. Organisations collecting personal information need to be able to handle constant requests to maintain compliance. The following is suggested:

**R.13 Automation:** Data deletion can be performed in two ways. Either manually, which requires extensive allocation time and human resources, or automatically. The latter, is recommended using a rule-based deletion process and with as little human interaction with the data as possible. This will allow organisations to handle constant requests to maintain compliance in a much efficient manner.

**R.14 Redundancy:** to be aware that an automation requires tagging, data mapping, testing and monitoring to achieve a structured process where appropriate policy rules are applied and followed. Furthermore, as deletion should be an irrevocable action, redundancy was acknowledged as a vital principle.

## V. CONCLUDING REMARKS

This paper, address key themes acknowledged when ensuring data protection rights required by the GDPR in a cloud computing environment. The study contributes to a better understanding of the current situation, by outlining challenges as well as recommendations that can be utilised mainly by controller organisations as well as data protection experts, authorities and consultants. Moreover, we assemble suggestions on how to overcome the challenges and provide an updated view of what data protection professionals assess as important best practices throughout the personal information lifecycle. From our results, we furthermore conclude that there appears to be a gap regarding responsibilities, expectations, perceived capabilities and trust between the processor's role and the role of the controller. Although the regulation states what the roles and responsibilities are when granting the data

subject rights and comply with the regulation, a tendency of conflicting interpretations are noted.

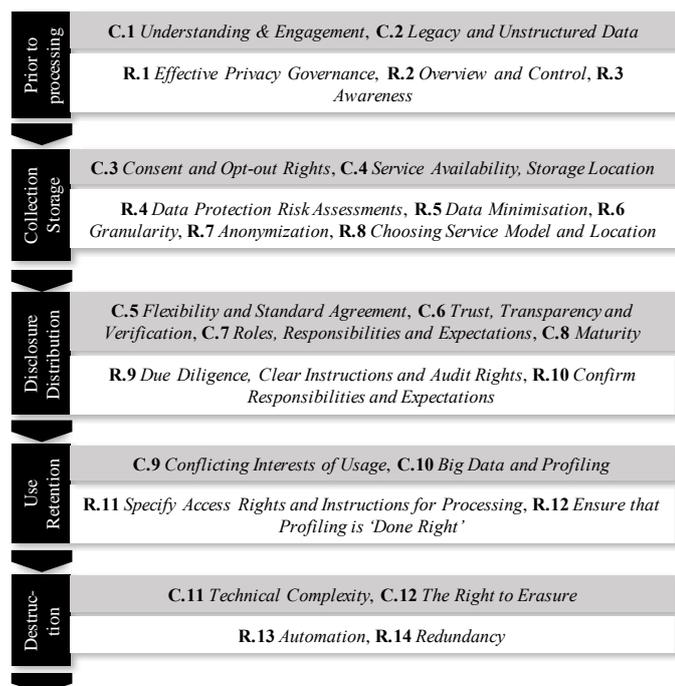


Figure 1. Challenges and recommendations throughout the lifecycle

With less than a year left before the GDPR enters force, this could increase the risk of noncompliance. The results suggest that there is a need for closer communication between the CSPs and their clients. Moreover, there appears to be an urge for the Data Protection Agencies community or the advisory bodies, such as the Article 29 Working Party, to state more clear guidelines on how the responsibilities are shared between the controller and the processor operating in the cloud, especially when it comes to what responsibilities the controller has towards the processor in terms of transparency.

#### ACKNOWLEDGMENT

The authors would like to thank Dr Rasika Dayarathna, senior lecturer in data protection at the University of Colombo, Sri Lanka, for reviewing an earlier version of this paper and the reviewers from the ICITST-2017 conference for valuable feedback which helped improve the paper prior to publication. The authors would further like to acknowledge Peter Lind - Senior Manager at KPMG Sweden, who provided hours of support and guidance throughout the research.

#### REFERENCES

- [1] European Regulation (EU) 2016/679 of the European Parliament and of the Council, "General Data Protection Regulation," 2016. [online] Available: <http://data.europa.eu/eli/reg/2016/679/oj>
- [2] European Parliament, "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data", Oct. 1995. [Online]. Available: [http://www.wipo.int/wipolex/en/text.jsp?file\\_id=313007](http://www.wipo.int/wipolex/en/text.jsp?file_id=313007)
- [3] L. A. Sosunova, I. V. Yakhneeva, and A. N. Agafonova, "Outsourcing Model Evaluation for Business Process Management in Digital Economy." *Mediterranean Journal of Social Sciences*, 2015.
- [4] J. Ruiter, M. Warnier, "Privacy regulations for cloud computing: Compliance and implementation in theory and practice," *Computers, privacy and data protection: an element of choice*. Springer, p. 361–76.
- [5] D. I. Curiac, & M. Pachia, "Controlled information destruction: the final frontier in preserving information security for every organisation", *Enterprise Information Systems*, 9(4), 2015. 384-400.
- [6] K. M. Ramokapane, A. Rashid and J.M Such, "Assured deletion in the cloud: requirements, challenges and future directions," *Proceedings of the 2016 ACM on Cloud Computing Security Workshop*. ACM, 2016.
- [7] S. Pearson, "Taking account of privacy when designing cloud computing services," 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing, 2009.
- [8] S. Paquette, P. T. Jaeger, and S. C. Wilson, "Identifying the security risks associated with governmental use of cloud computing," *Government Information Quarterly*, vol. 27, no. 3, pp. 245–253, 2010.
- [9] K. Popović and Ž. HuHocenski, "Security Issues and Challenges of Mobile Cloud Computing," *The 33rd International Convention MIPRO*, pp. 344–349, 2010
- [10] A. Cavoukian, "Privacy by design: the definitive workshop. A foreword by Ann Cavoukian, Ph.D.," *Identity in the Information Society*, vol. 3, no. 2, pp. 247–251, 2010.
- [11] R. Ko, B. Lee, and S. Pearson, "Towards Achieving Accountability, Auditability and Trust in Cloud Computing," *Advances in Computing and Communications Communications in Computer and Information Science*, pp. 432–444, 2011.
- [12] European Convention on Human Rights, Dec. 1950. [Online]. Available: [http://www.echr.coe.int/Documents/Convention\\_ENG.pdf](http://www.echr.coe.int/Documents/Convention_ENG.pdf)
- [13] F. Bieker, M. Friedewald, M. Hansen, H. Obersteller, and M. Rost, "A Process for Data Protection Impact Assessment Under the European General Data Protection Regulation," *Privacy Technologies and Policy Lecture Notes in Computer Science*, pp. 21–37, 2016.
- [14] Article 29 Data Protection Working Party, 'Opinion 05/2012 on Cloud Computing' 01037/12/EN, WP 196, 2012.
- [15] P. D. Hert, V. N. Papakonstantinou, and I. Kamara, "The New Cloud Computing ISO/IEC 27018 Standard Through the Lens of the EU Legislation on Data Protection," *SSRN Electronic Journal*, 2014.
- [16] R. Bernard, "Information Lifecycle Security Risk Assessment: A tool for closing security gaps," *Computers & Security*, vol. 26, no. 1, pp. 26–30, 2007.
- [17] J. Mason, *Qualitative researching* (2nd edition). London sage publications, 2002.
- [18] K. Charmaz, *Constructing grounded theory a practical guide through qualitative analysis*. Los Angeles: Sage Publications, 2006.
- [19] J. Corbin and A. L. Strauss, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Newbury Park, SAGE, 1998.
- [20] J. Corbin and A. L. Strauss, *Basics of qualitative research: techniques and procedures for developing grounded theory*. Los Angeles: SAGE, 2015.

# Distributed Computing Framework in Security: Case Study of Encryption Method

Shuaiyi Bu

College of Computer Science and Technology,  
Jilin University  
Changchun, China  
15843126572@163.com

Shuxin Yang

College of Computer Science and Technology,  
Jilin University  
Changchun, China

Haoming Ji

College of Computer Science and Technology,  
Jilin University  
Changchun, China

*Abstract*—Cloud computing has garnered increasing attention from researchers who have presented much work on performing massive computing tasks efficiently. Many security vulnerabilities have simultaneously occurred; therefore dealing with security problems in cloud computing has become an urgent issue. The purpose of this paper is to create a cloud computing framework based on the Map Reduce platform, a Google cloud-computing platform, and solve some security problems in the process of distributed computation.

Inspired by SMC (Secure Multi-Party Computation), a protocol naturally suitable for distributed computation, we adopted homomorphic encryption, which could be used for processing a large amount of data securely in cloud computation. We also find that order preserving encryption(OPE), an encryption algorithm put forward in 2004, can be used in our secure framework. Cognizant of the applicability of SMC and OPE in cloud computing, we combine them with Map Reduce to design a security framework of distributed computation.

Our major contributions consist of designing an innovative cloud computation framework in security based on Map Reduce, applying the order preserving encryption(OPE) algorithm, homomorphic encryption, and constructing a real distributed computation platform.

*Keywords*-homomorphic encryption, Map Reduce, order preserving encryption, SMC.

## I. INTRODUCTION

Cloud computing is a calculation model that was designed to deal with a large amount of data and resources. Our research mainly focused on how to calculate data in distributed circumstances more efficiently and effectively in the early stages of cloud computing without considering security problems in the computation. Unfortunately, some malicious attackers could easily get input information and data from the insecure distributed computing model. Even the data that is calculated during the process of computation can be monitored or stolen by some hackers. For example, we have a pool of private data, such as salary data, examination grades, or passwords, to be processed in the insecure cloud computation model. Whenever we submit our data to the cloud sever, the attackers who want to spy on our privacy could monitor the

submission and get data from the channel used to transmit the data from our computer to the cloud sever as the data is transmitted in plaintext. Moreover, the hackers could also get the data from the cloud sever since the model has not provided us any kind of security mechanism. There is no need for us to list all the circumstances in which the insecure model can be attacked by malicious attackers.

As more and more attention is paid to the security problems, some security mechanisms have been put forward. However, many of the theories merely solve part of the problems. Some traditional approaches to deal with the data privacy are based on randomization (1) (3).

However randomization has been proven to provide little privacy protection in most conditions, barring the introduction of a large amount of input noise, which can sacrifice the accuracy of the processing result. Others use cryptographic techniques (2) (4) (5) to deal with the privacy problem.

Though all previous work has its own merits, they all have certain caveats. Some security models cannot guarantee the accuracy of data, such as the model we discussed above, which others cannot provide strong privacy protection. What's more, some of the models pale in efficiency when compared with other secure models, making them unusable in real circumstances.

Considering such circumstances, our work focuses on creating an effective cloud computation framework that can be used in practice. The contributions of our work can be summarized as follows:

- We create an new, efficient, and pragmatic cloud computing model based on the Map Reduce platform, a prevailing distributed computation model, aimed at two particular issues.

Specifically, we make contributions as follows:

- We employ the order preserving encryption algorithm in the Map Reduce platform to solve one type of data processing problem. The algorithm can provide some kind of privacy protection for cloud computing.

- We employ the Paillier encryption algorithm to the Map Reduce platform. This algorithm can be used to solve the other type of security problem.

Particularly, we apply two simple and basic examples to verify the correctness and capability of our security framework. One is sorting a large number of integers, and the other is calculating the sum of a large pool of integers. Moreover, we use some particular methods to do measurements.

## II. RELATED WORK

With the development of cloud computing, people increasingly focus on the security of the distributed computation. Some work related to Map Reduce security has been done. Citation (6) presented a prototype concentrating on demonstrating the flexibility of their security framework based on Map Reduce, and avoids using fully homomorphic encryption.

There is also research focusing on some security issues for cloud computing (7). What's more, there is a research that proposes introducing a Trusted Third Party (8).

There is some research that aims to provide a security framework for Map Reduce as well. One methods provides a decentralized replication-based integrity verification scheme to ensure the integrity of Map Reduce in open systems (9). The work, whose method differs from our method by mainly using encryption, designs five security components and uses the connections and interactions among the five modules to achieve the goal of security. There is also work that uses an encryption method to add security to Map Reduce. However, the main achievement of this work is combining the public key cryptography with identity authentication for the security target (10). In comparison, our work uses the idiosyncrasy of the homomorphic encryption algorithm to protect the privacy of calculating data. We also find work in which the method is similar to ours. They use the OPE algorithm and homomorphic encryption to gain the security in DBMS (Database Management System). In contrast to this model (11), we apply the two algorithms to cloud computing and put forward a new framework. We also found work that applies the RSA cryptosystem and the Paillier cryptosystem to guarantee the security of cloud computing (12). However, they do not consider the tasks required to sort the data.

Different from some related works, our work introduces the OPE algorithm and Paillier cryptosystem into the Map Reduce platform. What's more, we designed a model that prevents the servers from knowing what they are calculating.

## III. PROBLEMS & ASSUMPTIONS

### A. Problem Raised

Initially, we tried to design an unprecedented security interface based on Map Reduce. That means we desire to

combine a specific SMC protocol with Map Reduce to create a brand new cloud-computing interface in security. If such work could be done, the users could merely do a little programming with our interface based on Map Reduce in order to do some calculations on security.

The convenience of the work in creating a security interface is self-evident. For example, suppose there is a task that aims to calculate the average grade of one particular subject for all the students. On the one hand, the amount of data is so enormous that we could only use cloud computing to achieve our goal without wasting a large amount of time. On the other hand, the grade of every single student should not be leaked, which means we want to protect the privacy of the data. Under these circumstances, a programmer can apply the security interface. All he or she needs to do is inherit the security interface and add the calculation function to the Map Reduce platform. Accordingly, we could finish our task conveniently and securely. The advantage of the interface is that whatever type of calculation we want to do, we can just use the single interface to achieve privacy protection.

However, ignoring some details in parallel programming based on the Map Reduce model, we find that it is almost impossible to create such a secure interface. In the following paragraph, we will clarify what the problems of creating such an interface are, and how we solve them and change our work to create a framework.

We came to the conclusion that such an interface is almost impossible according to certain conditions that confine us. Firstly, we want to make full use of the existing Map Reduce framework rather than changing it drastically or creating a new distributed computation framework. The main purpose of our work is to introduce the security features without modifying the existing cloud-computing framework. Secondly, we find that a single security interface cannot deal with all types of calculation tasks. The reason for this is obvious. If we design a function that could do the arithmetic operations securely, the function cannot complete operations like sorting data securely. As we are interested in designing an efficient framework, we do not want to use fully-homomorphic encryption that is not currently efficient.

Admittedly, there are surely some methods to create a single security interface based on Map Reduce, such as changing numerous details in Map Reduce like DFS (distributed file system), or using fully-homomorphic encryption. However, almost all of the approaches are contradictory to the rules and conditions confining us. Because of this, we need to find a solution and create a framework that includes the two programming models by analyzing how the problem is generated.

We find that it is difficult to use a single security interface to deal with different types of data processing in cloud computing. We distinguish two types of data processing as follows. One is called DRP (data-related

processing), and the other is called DNP (data-unrelated processing). The definitions of the two types are as follows:

- **DRP:** Whenever the data processing of cloud computing is related to the content of the input data, or the meaning of input data is indispensable for the processing, we call it **DRP**.

Data sorting is a typical example of **DRP**. It means we can only sort the input data correctly until we know the real content of the input data or we know size relations of the input data. Accordingly, we cannot simply choose an encryption algorithm to deal with this type of problem.

- **DNP:** Whenever the data processing of cloud computing is not related to the content of the input data, or the meaning of input data is not essential to the processing, we call it **DNP**.

Calculating the sum of the input data, such as the average grade of a class's examination, or multiplying the input data are simple but compelling examples of this type.

The two definitions we give above can cover most of the types of data processing issues. Thus, we design two secure cloud-computing models for the two types. One is aimed at **DRP** and the other is aimed at **DNP**. The final distributed computing framework in security includes the two models. For clarification, the framework is more like a scheme for the secure cloud computing based on Map Reduce. The core of our work is the scheme for security that we put forward for each type.

#### B. Assumptions Raised

Briefly, we assume that all nodes in computing clusters are unconvincing, and the client can be trusted. That means we cannot trust every single node in the cloud cluster. We can only trust the client and the user that owns the data. Under these circumstances, the data that the cloud servers process cannot be shown in plaintext, or we come up with an approach to randomize the data saved in the nodes in cloud clusters to make it less meaningful for attackers. We make these assumptions based on the possibility that the cloud servers may be taken advantage of by some malicious attackers attempting to leak the data and private information to the public.

Our work is based on what we assume so that the models we design cannot violate the assumptions that we have made above. We finally select two traditional but representative examples to test the correctness and capability of our models. As for **DRP**, we choose data sorting as our test example. For **DNP**, we select the process of calculating the sum of input data as our example. Admittedly, the model we design for each type cannot satisfy all the examples belonging to it because of the differences among all kinds of calculations, such as addition, division, multiplication, and subtraction, have something to do with the security model. However, it also makes an improvement in the security of cloud computing.

## IV. DEFINITIONS

### A. General Clarification

According to findings that we have analyzed above, on balance, our work design two security models. (One is for **DRP**, and the other is for **DNP**). Specifically, we find that **OPE** is really suitable for some examples of **DRP**, especially for the data sorting process. Additionally, we find that homomorphic encryption is naturally apt to create a privacy protection as for cloud computing. The latter is the idea that we get from the thoughts of **SMC** protocols, and particularly we choose Paillier encryption algorithm, which has additive privacy homomorphism property. We combine each the encryption algorithm with Map Reduce platform based on the assumptions that we have put forward in the preceding part. Before I clarify the details of our secure framework, we will show the definitions of three basic constitutions of our work: order preserving encryption algorithm, Paillier encryption algorithm, and Map Reduce. In detail, we will clarify the rationale of each constitution in this part to help establish the verification of correctness and capability about our work.

### B. Order Preserving Encryption

The algorithm presents an encryption technique called **OPE** that makes some kinds of comparison operations be possible. The operations can be directly used on encrypted data, with no need of decrypting the operands. Basically, **OPE** works in three phases as follows. Essentially, the **OPE** encryption is an injective mapping from plaintext to ciphertext.

Let

$$m \leq n, P = \{i | 1 \leq i \leq m\} \quad (1)$$

is the set of plaintext.

$$C = \{i | 1 \leq i \leq n\} \quad (2)$$

is the set of ciphertext.

$$BE = (KI, ENCR, DNCR) \quad (3)$$

**KI** is a key generation algorithm, **ENCR** is a encryption algorithm, and **DNCR** is a decryption algorithm. The scheme **BE** applies the three algorithms to help do the function (13).

### C. Paillier Cryptosystem

The Paillier cryptosystem consists of three stages as follows. Key generation:

- 1) Choose two large Prime numbers  $p$  and  $q$ .

Note that  $p$  and  $q$  must meet the command that

$$\gcd(pq, (p-1)(q-1)) = 1 \quad (4)$$

- 2) Define a number  $n$

$$n = p \times q \quad (5)$$

- 3) Define a number  $m$

$$m = \text{lcm}(p-1, q-1) \quad (6)$$

- 4) Choose a random number  $g$ .

$$g \in Z_{n^2}^* \quad (7)$$

- 5) Do the calculation

$$u = (L(g^m \bmod n^2))^{-1} \bmod n \quad (8)$$

$$L(s) = \frac{s-1}{n} \quad (9)$$

The public key is (n,g), and the private key is (m,u)

Encryption:

Definitions: “m” is the plaintext, “r” is a random number, and “c” is the ciphertext.

$$c = g^m \times r^n \text{ mod } n^2 \quad (10)$$

Decryption:

$$m = L(c^m \text{ mod } n^2) \times u \text{ mod } n \quad (11)$$

The additive homomorphic property of Paillier cryptosystem can be proved as follows.(m1, m2 are two different plaintexts. r1, r2 are two random numbers, the remainders have been defined above)

$$E(m_1) = g^{m_1} r_1^n \text{ mod } n^2, E(m_2) = g^{m_2} r_2^n \text{ mod } n^2 \quad (12)$$

$$E(m_1) E(m_2) = g^{m_1+m_2} (r_1 r_2)^n \text{ mod } n^2 \quad (13)$$

$$D(E(m_1) E(m_2)) = m_1 + m_2 \quad (14)$$

From the three formulas we can see that after decrypting the sum of two ciphertexts, the result is equal to the sum of corresponding plaintexts (14).

#### D. Map Reduce

Map Reduce is a programming model and a practical tool which could be used for processing and calculating a large amount of data. It uses parallel and distributed algorithm. Although the purpose of Map Reduce framework does not accord with their initial forms, this model is encouraged by map and reduce functions which are always used in functional programming. The Map Reduce framework’s key contributions are scalability and fault-tolerance. In general, the effectiveness and improvement can usually merely be seen with multi-threaded implementations. It is essential to optimizes the communication cost for a good Map Reduce algorithm. Map Reduce libraries have been written in many programming languages. Apache Hadoop is a popular open-source platform that roots for the distributed computing Apache Hadoop. The name Map Reduce traditionally referred to the exclusive technology of Google.

The general details of working process in Map Reduce are shown in Figure 1. JobTracker is responsible for dividing the original data into several input splits and communicate with TaskTrackers. TaskTracker is responsible for coordinating some several specific phases such as Map phase, shuffle, and Reduce phase.

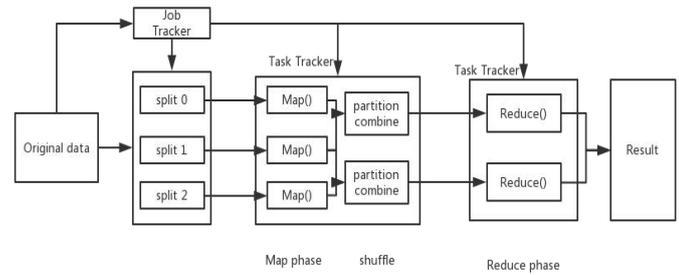


Figure 1. General working principle of Map Reduce

## V. DESIGN & IMPLEMENTATION

Based on the theoretical foundation of the two encryption algorithms that we will apply in our work and the technical details in the Map Reduce platform on which our work is based, we will discuss the concrete process of the secure cloud computing models we designed in this section. Specifically, we will illustrate how we combine the encryption algorithm with the existing distributed computing platform. Moreover, we will display how the data are shown when they are calculated in the cloud servers.

We have assumed that none of the nodes in the cloud cluster can be trusted, and the client who owns the data is the only one that is convincing. Accordingly, to achieve our goal of privacy protection, we make full use of the properties of Paillier cryptosystem and OPE (refer to homomorphism and order preservation) and make little change to the details of Map Reduce. Generally, the core of our work is dealing with input data, focusing on making it more secure when calculated in the servers.

### A. DRP Design

For the DRP type, we use the OPE algorithm to encrypt the input data. HDFS (Hadoop Distributed File System) is a type of file system that is specifically designed for Hadoop, a basic framework for a distributed system. Map Reduce is a part of the Hadoop framework. Since we do not want to change the basic framework of Map Reduce when considering the efficiency of our framework, we take fully advantages of the Map Reduce programming interface. Our work inserts the encryption codes and decryption codes into the main function. Since we never trust the calculation nodes in the cloud cluster, we have found that the best way to protect the private data is to not allow the JobTracker and TaskTracker to know what the real data is from the start. Accordingly, we encrypt the original input data and make the nodes read and calculate the encrypted data. We divide the main function into three stages.

#### 1) Encryption Stage

We first read the input data from the original files that the user wants to process, to the memory buffer. Then we use the OPE algorithm to encrypt the data that has been read to

the buffer. We apply the encrypt function of OPE to encrypt the data. Finally, we write the encrypted data to the output file. We need to stress the point here that since Map Reduce needs to read the data from the HDFS, that means it reads the data from a fixed position in the system, so that we need to write our encrypted data to a position where the JobTracker can find.

2) *Configuration Stage*

This stage is like most of the existing Map Reduce programming examples. We configure some basic information of the job we will make the nodes do, such as setting the name of a job, setting the Mapper class, setting the Reduce class, and setting the input and output path. We have to note that the input path must be the path to which we write our encryption data.

3) *Decryption Stage*

Following the configuration stage is the decryption stage in the main function. We read the calculated encrypted data from the file according to the output path to the memory buffer, and use the key to decrypt the data. The OPE algorithm we use is a symmetric encryption algorithm so that the decrypt key is the same as the encrypt key. Finally, we write the decrypted data to another file that we have created before we process the job.

B. *DNP Design*

As for the DNP type, though the process is similar to the process of DRP, there are surely some differences.

- The DRP type uses the OPE algorithm to keep the original order of the input type, otherwise the cloud computing servers could not finish the job successfully. However, as for the DNP type, all we need to do is finish a type of calculation operation, such as adding, dividing, so that we find the Paillier cryptosystem to satisfy our needs. Having the additive homomorphism property, the algorithm can make the cloud computing servers calculate the data without knowing the real type and whenever we decrypt the result, we can get the correct result.
- Since the Paillier cryptosystem is an asymmetric encryption system, the decrypt key is different from the encrypt key. DRP: Whenever the data processing of cloud computing is related to the content of the input data, or the meaning of input data is indispensable for the processing, we call it DRP.
- Whenever we do the process in Map Function for DNP, we need to invoke the add function of the Paillier algorithm to calculate the sum of the input data. However, with regard to DRP, we usually do a job such as sorting or finding, and we do not need to calculate the input data, meaning the function is not related to the encryption algorithm.

VI. ANALYSIS

In this section, we will analyze the distributed computing framework in security we created in three respects: security, capability, and expandability.

A. *Security*

For the DRP type, the cloud-computing servers deal with the encrypted data so that they do not know the real meaning of the data. However, hackers may attempt some data mining and attempt a statistical attack on the encrypted data since the size relations of the data have been preserved even though we have encrypted it. Let us look an example. Our input data is a set of students' grades for a subject as well as their names. Though we encrypt both the name and the grade, hackers can still compare this input file with another input file in order find which students performed best in which subject. However, the attackers can only obtain limited information about the input data because we have protected the users' privacy to the best degree. In the worst case scenario, the attackers may use some advanced technique in which they could find about who got the best grade from one single encrypted input file, but they can not know what score she or he has received. We should say that the users decide on the level of security. Different users have different needs in terms of the security level. For most of the distributed computing jobs, preventing the attackers from knowing the actual content of the data could meet many users' needs, although the attackers may get some peripheral information. We may also note the less security the framework has, the higher the capability and efficiency of the framework.

A specific example for the DRP type is shown in Figure 2. The first column in the left diagram shows some student IDs, and the second column shows the corresponding grades. The right diagram shows the results of the encrypted data in the left diagram.

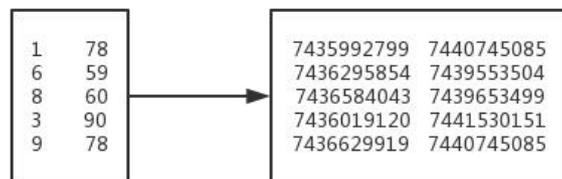


Figure 2. Test case of an example

As for DNP type, the cloud-computing servers are only responsible for calculating input data without knowledge of the data type or which data they are calculating. What's more, the additive homomorphism guarantees the correctness of the calculated result. A specific example of DNP type is shown in Figure 3. The Map Phase deals with the encrypted grades, and the Reduce Phase adds all the encrypted grades and writes the sum to the output file.

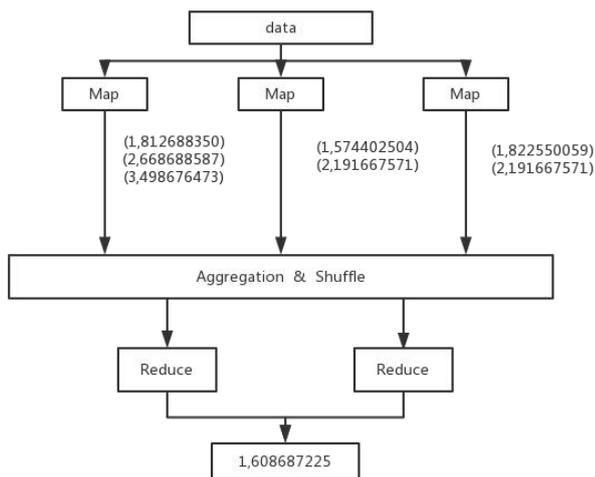


Figure 3. Test case in the working process of Map Reduce

**B. Capability**

Since the encryption stage and the decryption stage are merely doing the work of encrypting and decrypting the data of a file and writing them to another file, they have nothing to do with the process of Map Reduce. Accordingly, these two stages are not necessary for obtaining privacy protection. They can be done in an independent program. However, considering the integrity of the security framework we have to create, we insert these two stages into the main function. Under these circumstances, the capability can be analyzed from two different perspectives. The two perspectives can be summarized as follows:

- The first perspective does not consider the time cost of the encryption and decryption processes. We only test the time consumption when the data is calculated in the Map Reduce platform since the distributed computing process is the core of our work.
- The second perspective calculates all the time consumed, including the three stages we have discussed above.

We will discuss each perspective in detail in the next section and use specific examples to test the real capability of the security framework.

**C. Expandability**

With regard to the DNP type, the users can apply our model to do a host of calculation jobs. However, because of the limitation of the Paillier algorithm, since it only has the additive homomorphism, users could only use our model to jobs that only involve the add operation. However, our model provides the possibility for future researchers to use a more powerful encryption algorithm that could do all the arithmetic operations to complete a job. The fully homomorphic

algorithm is a typical example of such an algorithm. Nevertheless, the reason that we do not use this fully homomorphic algorithm is because it has a low efficiency.

As for the DRP type, our work can not only complete data sorting but also return a maximum or minimum number for all the input data. What's more, the model can also perform number counting as well. So, it is shown that the DRP type model can expand to meet other requirements of different jobs.

**VII. EVALUATIONS**

In this section, we test our work using two examples, data sorting and calculating a sum, to test the correctness and capability of our work from different perspectives. Since we concentrate more on the difference in capability between the secure model and original Map Reduce, the number of nodes in cloud computing is not important. Therefore, we tested our work on a single computer. The computer configuration has an Intel i5-3210M 4 core with 4GB RAM.

Capability Comparison Test 1: In this experiment we test the speed difference between our security model and the normal Map Reduce program. The encryption and decryption time are measured in our model in this test. The results are shown in Table 1 and Table 2.

TABLE I. DNP TYPE TEST(USING SUM EXAMPLE)

Data(MB)	Encrypted	Time consuming(s)
1	yes	352
1	no	5
2	yes	637
2	no	8
4	yes	1209
4	no	10

TABLE II. DRP TYPE TEST(USING DATA SORTING EXAMPLE)

Data(KB)	Encrypted	Time consuming(ms)
1	yes	91738
1	no	3113
2	yes	384582
2	no	2238
4	yes	722142
4	no	2105

Capability Comparison Test 2: In this experiment, we test the speed difference between our security model and the normal Map Reduce program. Note that the encryption and decryption time are measured in our model in this test. The results are shown in Table 3 and Table 4.

TABLE III. DNP TYPE TEST(USING SUM EXAMPLE)

Data(MB)	Encrypted	Time consuming(s)
1	yes	6 (without E&D)
1	no	7
2	yes	8 (without E&D)
2	no	7
4	yes	9 (without E&D)

4	no	10
---	----	----

TABLE IV. DRP TYPE TEST(USING DATA SORTING EXAMPLE)

Data(KB)	Encrypted	Time consuming(ms)
1	yes	2228 (without E&D)
1	no	3113
2	yes	3004 (without E&D)
2	no	2238
4	yes	2689 (without E&D)
4	no	2105

In Figure 4, we compare the capability of our model, which uses the Paillier cryptosystem, with the original Map Reduce framework. They have the same original input data. We can see that the time consumption of our model increases as the size of input data increases. Compared with the original Map Reduce framework, our model needs more time because of the encryption and decryption phases that guarantee the security.

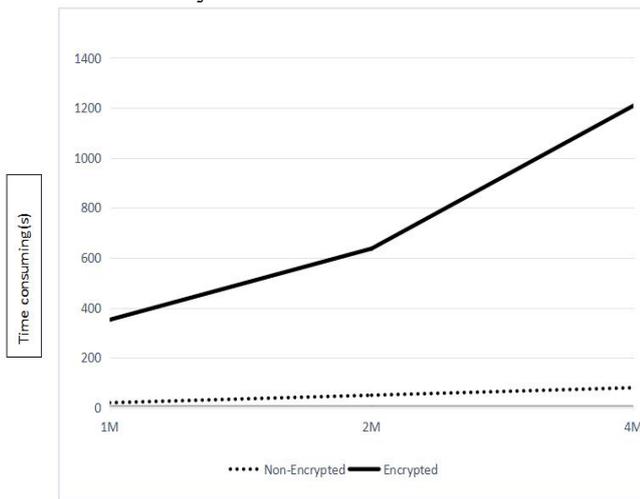


Figure 4. Comparison of capability for Paillier cryptosystem

In Figure 5, we compare the capability of our model, which uses the OPE algorithm, with the original Map Reduce framework. We can see from the figure that the time consumption has notably increased with the increasing size of the input file. Combined with the observation about the comparison of the capability of our model, which does not include the encryption and decryption phases, and normal Map Reduce work, which is similar, we find that the efficiency of the OPE algorithm is low. This is one of the disadvantages of our security model.

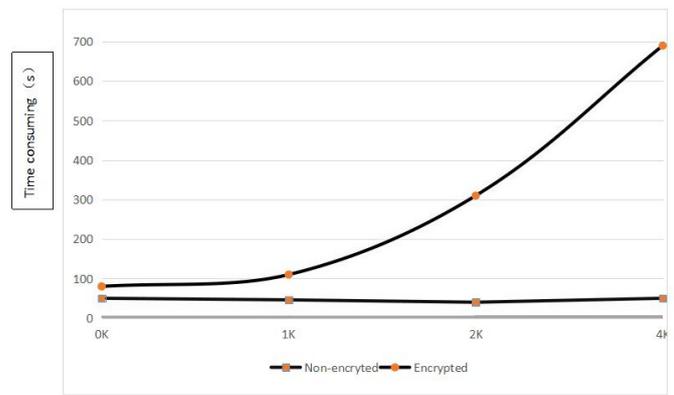


Figure 5. Comparison of capability for OPE cryptosystem

### VIII. CONCLUSION

Our work presents an extensible, applicable distributed computing framework, with two models aimed at two different types of issues. From the tests based on two simple, but compelling, examples, we can see that the efficiency of the OPE algorithm is low and may cost clients a lot of time to encrypt and decrypt the data. Time is the main factor in evaluating the efficiency and applicability of a model. Thus, we need to consider whether the time used to encrypt and decrypt the data is more than the time that is saved using Map Reduce to do calculations on the large amount of the data. From the figure above, we can see that the time consumption, including encryption and decryption, is approximately exponential. We can conclude that the OPE algorithm can only be used in processing large amounts of data, such as 5 MB. If you are eager to calculate data limited to only 1 MB, you need to bear the low efficiency of the algorithm. However, you do need to waste time to gain security. Most importantly, we need to find a balance between time consumption and security.

As for the DNP type, we found that the efficiency of Paillier is better than the OPE algorithm. With the amount of data becoming larger, the increase in time is directly proportional in our model. Thus, the additional time consumption of encryption and decryption is bearable considering the security we could gain.

Accordingly, we suggest that the users could directly use the model for the DNP type, as there is little efficiency loss. As for the DRP type, we recommend putting the encryption and decryption stages as an independent program and using the Map Reduce platform to process the encrypted data.

There is a lot of work that can be done based on our framework. Other researchers can improve our model using a more pragmatic and effective model, as our work can only deal with some particular types of operations. Another direction to take the research is to come up with a revised fully homomorphic algorithm that is highly efficient so that

it can be combined with Map Reduce and used commercially.

## IX. ACKNOWLEDGEMENT

The work described in this paper is supported by National Innovation Project of College Students. Thanks for the help and guidance from our tutor. We also wish to thank Google for developing the Map Reduce platform.

## REFERENCES

- [1] J. Ma and K. Sivakumar, Privacy-Preserving Bayesian Network Learning Using Post Randomization, (in preparation), 2005
- [2] G. Ateniese, R. Di Pietro, L. V. Mancini, and G. Tsudik. Scalable and efficient provable data possession. In Proceeding the 4th international conference on Security and privacy in communication networks (SecureComm'08), pages 1–10, New York, NY, USA, 2008. ACM.
- [3] D. Meng, K. Sivakumar and H. Kargupta. Privacy-Sensitive Bayesian Network Parameter Learning. In the Fourth IEEE International Conference on Data Mining. Brighton, UK. November 2004. Elissa, "Title of paper if known," unpublished.
- [4] J. Baek, R. Safavi Naini, and W. Susilo. On the integration of public key data encryption and public key encryption with key word search. In International Conference on Information Security (ISC'06), volume 4176 of Lecture Notes in Computer Science. Springer, 2006. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetism Japan, p. 301, 1982].
- [5] M. Chase and S. M. Chow. Improving privacy and security in multiauthority attribute based encryption. In ACM conference on Computer and communications security (CCS '09), pages 121–130, New York, NY, USA, 2009. ACM
- [6] Guo Z, Zhu X, Guo L, et al. Design of a security framework On MapReduce[C]//Intelligent Networking and Collaborative Systems (INCoS), 2013 5th International Conference on. IEEE, 2013: 139-145.
- [7] Subashini S, Kavitha V. A survey on security issues in service delivery models of cloud computing[J]. Journal of network and computer applications, 2011, 34(1): 1-11. M. Chase and S. M. Chow. Improving privacy and security in multiauthority attribute based encryption. In ACM conference on Computer and communications security (CCS '09), pages 121–130, New York, NY, USA, 2009. ACM
- [8] Zisis D, Lekkas D. Addressing cloud computing security issues[J]. Future Generation computer systems, 2012, 28(3): 583-592.
- [9] Wei W, Du J, Yu T, et al. Securemr: A service integrity assurance framework for mapreduce[C]//Computer Security Applications Conference, 2009. ACSAC'09. Annual. IEEE, 2009: 73-82.
- [10] Zhao J, Wang L, Tao J, et al. A security framework in G-Hadoop for big data computing across distributed Cloud data centres[J]. Journal of Computer and System Sciences, 2014, 80(5): 994-1007.
- [11] Popa R A, Zeldovich N, Balakrishnan H. CryptDB: A practical encrypted relational DBMS[J]. 2011.
- [12] CHEN Z, DU M, YANG Y. Homomorphic cloud computing scheme based on RSA and Paillier[J]. Computer Engineering, 2013,
- [13] Boldyreva A, Chenette N, Lee Y, et al. Order-preserving symmetric encryption[C]//Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer Berlin Heidelberg, 2009: 224-241
- [14] Paillier P. Public-key cryptosystems based on composite degree residuosity classes[C]//International Conference on the Theory and Applications of Cryptographic Techniques. Springer Berlin Heidelberg, 1999: 223-238.

## **Session 12: Infonomics and e-Technology**

Title: Seeking Academic Information on the Internet: Doctoral Students' Internet Self-efficacy and Emotions

(Author: Lee Yen-Mei)

Title: Customer Churn Analysis: A Case Study on the Telecommunication Industry of Thailand

(Author: Paweena Wanchai)

Title: The Influence of Advanced and Secure E-Commerce Environments on Customers Behaviour: The Case of Saudis in the UK

(Authors: Haya Alshehri, Farid Meziane)

# Seeking Academic Information on the Internet: Doctoral Students' Internet Self-efficacy and Emotions

Yen-Mei, Lee

School of Information Science and Learning Technology  
University of Missouri  
Columbia, Missouri, U.S.A  
pppopq596@gmail.com

**Abstract**—The main purpose of this proposed study is to understand the relationship between doctoral students' Internet self-efficacy and their affections when seeking academic information on the Internet. A structured questionnaire will be designed and conducted for 300 doctoral students. Participants' demographic information, information seeking behavior, Internet self-efficacy, and emotions will be measured in this questionnaire by using Information-Seeking Behavior Scale (ISBS), Internet Self-efficacy Scale (ISS), and the Positive and Negative Affect Schedule (PANAS). For the predicted outcomes, doctoral students' academic information seeking behaviors will be analyzed. Moreover, the interventions among doctoral students' Internet self-efficacy, emotions, and their information seeking behaviors will be analyzed and addressed. In the future, this proposed study will be conducted practically. The expected findings will be proposed for doctoral students so that they can based on the provided information to build up their suitable information seeking styles. Furthermore, the expected findings can also be presented for professionals, scholars, or even doctoral students' academic advisors who mainly do research and seek academic information online to have a better understanding of doctoral students' mental and psychical statuses on conducting academic information seeking activity on the Internet. To sum up, the results will not only reflect doctoral students' information seeking behaviors but also provide helpful suggestions for further researches.

**Keywords**—*Information seeking behavior; Internet self-efficacy; Emotions*

## I. INTRODUCTION

In the digital society, seeking academic resources such as e-books, journal articles, or conference proceedings on the Internet has become a crucial activity for scholars (including professionals, faculties, and postdoc fellows) to review previous studies and relevant research [11, 19, 23]. On the same trajectory, in terms of doctoral students, finding academic resources on the Internet is also essential work in their academic life [5, 6]. Moreover, doctoral students' information seeking behaviors are more sophisticated and complex than masters and undergraduate students [6]. Specifically, doctoral students are often faced with tasks that require them to identify

information needs, locate corresponding sources, extract and organize relevant information, and synthesize information from different kinds of sources [5]. Thus, what kind of behaviors doctoral students perform during the information seeking process and how they apply these behaviors to solve problems they encounter on the Internet are worth investigating separately from masters and undergraduate students [6].

There are many factors can influence individuals' information seeking behaviors, including prior Internet experiences, emotions and feelings (joy, pleasant, fear, anxiety, bad memory, etc.), and Internet Self-efficacy [2]. People who have more Internet experience have more knowledge of information seeking strategies used in databases or e-journals and less seeking barriers such as "face problems to retrieve records of good quality and relevant to the information need" [6, 16]. Moreover, not only does prior Internet experience affect people's information seeking behaviors, but emotions, feelings, and Internet self-efficacy can also motivate individuals to start, expand, limit, terminate, or avoid information seeking (How does it do so?) [21, 22].

There is little information about how information seeking behaviors impact doctoral students since most studies of information seeking behaviors and Internet Self-efficacy have focused on undergraduate students [10, 15, 20, 21, 25, 26, 27]. Even though some research mainly addresses graduate students' academic information seeking behaviors, they seldom demonstrate masters and doctoral students separately [7, 16, 18, 27]. Thus, doctoral students are selected as the main population in this study because this specific group are in the minority status in the previous research. More importantly, though there are studies critically putting efforts in investigating doctoral students' information seeking behaviors, they rarely address students' self-perceptions, emotions, and feelings at the same time [5].

In this vein, based on the critical thoughts mentioned above, the proposed study will investigate doctoral students' academic information seeking behaviors and how their Internet Self-efficacy and emotions correlate with the behaviors. For the next section, relevant research work and studies will be addressed.

## II. LITERATURE REVIEW

### A. *User's Information Seeking Behaviors and Internet Experiences*

With the rapid evolution of the technology, using the Internet becomes the first priority for people to begin their information seeking behavior [6]. Though the Internet is popularly used in seeking information, different prior Internet experiences can arouse different kind of behaviors. In Korobili, Malliari, and Zapounidou's research [16], they investigate 235 people's information seeking behaviors and find that when Internet experience increased, lack of knowledge of search techniques decreased. That is, individuals with higher Internet experience in retrieving information from databases or e-journals tend to have less difficulties when seeking information on the Internet. Catalano [6] also support this findings that users those who had more Internet experience had more knowledge of search strategies used in Internet resources. Moreover, in the study conducted by Aula, Jhaveri, and Käki, [1], they conclude that there was a specific information seeking behavior from 236 experienced Internet participants that individuals with higher Internet experience prefer to use multiple browser windows or tabs in parallel while seeking information on the Internet. Further-more, Thatcher [28] not only has the same conclusion with Aula, Jhaveri, and Käki [1] that experienced Internet people are likely to use the parallel information seeking style, but also demonstrates that individuals with lower Internet web experience are likely to find information by using sequential information seeking style. For example, people with lower Internet experience would not open a new website to find more information until they completely read the information they found from the previous website.

On the other hand, though people prefer to rely on the Internet to seek their initial information; in the Liyana and Noorhidawati [18] research, they recruited 217 participants to study their information seeking behaviors and noticed that even though the Internet search engine was the first information resource used by users, they tended to apply more reliable information resources such as digital libraries and online databases to evaluate the trustworthiness of the information. Liyana and Noorhidawati stated that more than 94.3 percent users (n=132) evaluated the reliability of the information they found by using other re-sources once they were acquainted with the subject matter. Liyana and Noorhidawati also described some interesting information seeking behaviors in their study. They concluded that people tended to use different information re-sources when they could not find the information that they obtained at the first time and would follow up by conducting a new seeking activity using another different combination of keywords. Thus, people's information seeking behaviors are diverse and correlated with their previous Internet experiences.

Several behaviors people have when seek information on the Internet are addressed above. However, there are other important factors which may affect people's information seeking behaviors, but researchers seldom discussed about in the past. For instance, in Liyana and Noorhidawati's [18] study, they mention that although people indicate that they can find

the information easily, but they are having problem with information excessive, as well as having difficulties to ensure the trustworthiness of the information and difficulties in understanding the information contents. People feel frustrated and insecure when they face problems during the information seeking process, but no assistants or guides can support their needs. Thus, how people's perceptions and emotions interact with the information seeking behaviors can be a crucial issue for researchers to take into consideration.

For the next section, related work about people's self-perceptions/self-efficacy when seeking information on the Internet will be discussed.

### B. *User's Internet Self-efficacy and Information Seeking Behaviors*

People's perceptions of how they believe or judge about their capabilities to complete a specific performance can be defined as an individual's Self-efficacy [3]. A core theoretical property of self-efficacy is that it is concerned not the skills an individual has; instead, it reflects what persons believe they can do with the skills they possess [13]. Based on the concept of Self-efficacy demonstrated by the social scientist Albert Bandura, Eastin and LaRose [10] and Torkzadeh and Van Dyke [26] extend the thoughts and address that the concept of the Internet Self-efficacy is individual's self-perception and self-competency in interacting with the Internet. Internet Self-efficacy does not refer to a person's skill at performing specific Internet-relevant tasks, such as writing HTML, using a browser, or transferring files. On the contrary, it assesses a person's judgment of his or her ability to apply Internet skills in a more encompassing mode, such as finding information or troubleshooting search problems. Moreover, Internet Self-efficacy may be distinguished from computer Self-efficacy as the belief that one can successfully perform a distinct set of behaviors required to establish, maintain and utilize effectively the Internet over and above basic personal computer skills [10].

Furthermore, in the practical approach, several factors, such as individuals' prior Internet experience, emotions, and anticipated performance outcomes or actual information seeking behaviors are correlated with a person's Internet Self-efficacy. First of all, previous Internet experience can be an important element interacted with people's Internet Self-efficacy [14]. Users having more prior engagement in the Internet-based context display more positive Internet self-efficacy [7, 14, 28]. Secondly, people also rely partly on their motional states in judging their capabilities. Bandura [3] mentions that positive mood enhances perceived self-efficacy, while despondent mood diminishes it. Also, a third possible factor is the anticipated performance outcomes and individual's Internet Self-efficacy. The outcomes people expect depend largely on their beliefs of how well they can perform in given situations. Bandura [4] addresses that those of high efficacy expect to gain favorable outcomes through good performance, but those who anticipate poor performances of themselves conjure up negative outcomes. Lastly, Internet Self-efficacy can influence web users' Internet behaviors [4, 21]. People who with lower confidence about their browsing skills tend to perform more verbal aggression to others on the Internet or social media [21]. People's information seeking behaviors can

be changed according to the interaction happens between the Internet Self-efficacy and the external factors. Thus, the relationship between the Internet Self-efficacy and different kind of external factors is a critical issue for researchers to discuss.

C. *User's emotions and and information seeking behaviors*

In this section, because both cognitive and affective needs are important as the factors of the information seeking behavior, people's emotions related to their behaviors will be addressed [16].

Affective factors of information seeking behavior are important for motivating people's ways of accessing, seeking and using information sources in different contexts such as academic task performance and learning. However, affective factors so far still play a residual role in the information seeking research [22].

In Savolainen's [22] research, how positive and negative emotions and feelings such as joy, anxiety and fear encourage or discourage an individual's at-attempts to seek information are discussed. Savolainen concludes that emotions and feelings motivate individuals to start, expand, limit, terminate, or avoid in-formation seeking. Savolainen mentions that a positive emotional valence is mainly related to start and expand information seeking. Furthermore, some emotions, such as anxiety, may motivate in multiple ways, including starting information seeking to information avoidance. In contrast, some positive emotions, such as joy, are typically experienced when starting or expanding information seeking.

On the other hand, emotions not only simply have interactions with people's information seeking behaviors, but also have interventions with individuals' prior Internet experiences and Internet self-efficacy during the seeking process [7]. Specifically, people who have more Internet experiences, tended to express more positive feeling, lower anxiety and independent control toward the usage of Internet [28]. Moreover, when Internet self-efficacy increases, individuals' Internet stress will decrease [10]. Thus, Internet experiences, Internet self-efficacy, and emotions are essential factors to influence people's information seeking behaviors.

III. THEORETICAL APPROACHES

In this proposed research, a specific group, doctoral students, will be investigated of their Internet Self-efficacy and emotions when seeking academic information on the Internet. To support the research purpose, the Social Cognitive Theory – Self-efficacy addressed by Albert Bandura [3] and the Affective Load Theory proposed by Diane Nahl [12] will be applied for demonstrating the theoretical foundation.

In Bandura's Self-efficacy Theory, he addressed that people not only apply their prior experiences to judge their abilities but also use mental status such as emotions, moods, and feelings to support their believes. A person's behaviors or performances will be influenced by different degrees of believing on the capabilities. On the other hand, Diane Nahl mentions that positive and negative emotions can support and motivate people's cognitive activities. Moreover, individual's affective

load is dominated by time pressures and the feeling of uncertainty such as irritation, frustration, anxiety and rag. People who with higher affective load will have more ineffective information behaviors, while those who with lower effective load when having information behaviors tend to have better performances. Based on these two theory concepts, we can infer that different-level Self-efficacy and different emotions can affect people's information behaviors. Thus, with previous related work review and theories support, three crucial elements are embedded in the proposed study, including Internet Self-efficacy, emotions when seeking on the Internet, and the third one, information seeking behaviors, see the proposed theoretical framework on the Fig. 1 Specific research questions are addressed as following:

- RQ1: What academic information seeking behaviors performed by different web-experience doctoral students?
- RQ2: What is the relationship between Information seeking behaviors and Internet Self-efficacy when doctoral students seek academic information on the Inter-net?
- RQ3: What is the relationship between Information seeking behaviors and emotions when doctoral students seek academic information on the Internet?
- RQ4: With in the same Internet experience, does doctoral students' Internet Self-efficacy correlate with their emotions when seeking academic information on the Internet?

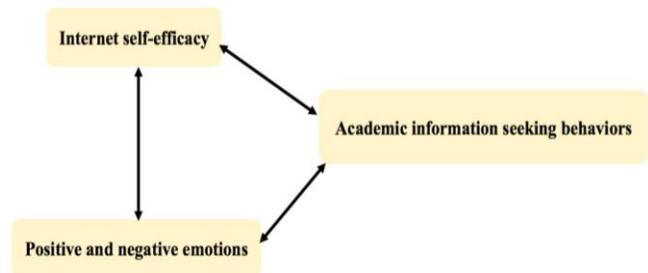


Figure 1. The proposed theoretical framework

IV. METHODOLOGY

A. *Research Design*

The practical conduction of the proposed study will follow the theoretical framework and research questions illustrated in the previous section. Detailed methodology information are provided as following.

B. *Participants*

1) *Sample size:* In this study, 300 doctoral students will be surveyed in this study. Estimate 150-200 responders answer the questionnaire completely.

2) *Sampling method:* Convenient sampling method will be used [9]. Participants will be informed to complete the survey by E-mail and social media.

C. *Data collection*

In the proposed study, a quantitative research method will be applied in the data collection process. A structured survey will be designed for doctoral students to measure their academic information seeking behaviors, Internet self-efficacy, and emotions. Different kind of scales, including the Information-Seeking Behavior Scale (ISBS), the Internet Self-efficacy Scale (ISS), and the Positive and Negative Affect Schedule (PANAS) Scale, will be used. Also, doctoral students' basic demographic information will be recorded. See more detailed information below.

1) *Measurement Instruments:* The data collection instrument will be a structured questionnaire. The questionnaire is shaped by four sections (A-D): demographics, Information-Seeking Behavior, Internet Self-efficacy, and emotions. All the items in the questionnaire were adapted from related previous studies.

a) *Section A: Demographical Information:* In the Demographical Information section, participants' fields of study, grade levels, gender, the frequency of searching academic information on the Internet such as years of Internet experience, and the place they use the Internet to search academic information are asked [14, 17].

b) *Section B: Information-Seeking Behavior Scale (ISBS):* In this section, a Information Seek-ing Behavior Scale adjusted from Timmers and Glas [25] are used. Including ten-item scale for applying search strategies (alpha= 0.68), fourteen-item scale for evaluating information (alpha= 0.74), and six-item scale for referring to in-formation (alpha= 0.81).

c) *Section C: The Internet Self-efficacy Scale (ISS):* For the C section, the Internet Self-efficacy scale (ISS) is modified from Chuang, S. C., Lin, F. M., and Tsai, C. C. [7] study. Total 36 items involved in this scale. The Cronbach's alpha of ISS was 0.92, including eight items for usage, six items for information sharing, four items for communication, eight items for verification, six items for metacognition, and four items for application.

d) *Section D: The Positive and Negative Affect Schedule (PANAS) Scale:* In the section D, The Positive and Negative Affect Schedule (PANAS) Scale is used [8].

D. *Data Analysis*

Analysis tool, a statistic measurement software, SPSS, will be applied to analysis the data. Based on the four research questions to demonstrate the results.

V. EXPECTED OUTCOMES

There are two expected outcomes being addressed in this section. First of all, the study outcome will describe doctoral students' academic information seeking behaviors, and how their Internet Self-efficacy and emotions correlate with the

behaviors. Secondly, Suggestions based on the results will provide for doctoral students and faculties (scholars or advisors) as references to help them find more efficient and meaningful way when seeking academic information on the Internet.

VI. WORK SCHEDULE AND RESEARCH PLANNING

For the timeline of conducting the proposed study, four stages are included in the plan. Firstly, a critical research questions and gaps are addressed by reviewing previous related studies. Secondly, a theoretical framework is shaped to support the study purpose. In the third stage, a concise research design involved the data collection procedure, data measuring instruments, and data analysis methods are decided in this stage. Lastly, expected outcomes and the follow-up practical conducting schedule are planned. See the proposed timeline illustrated on Fig. 2.

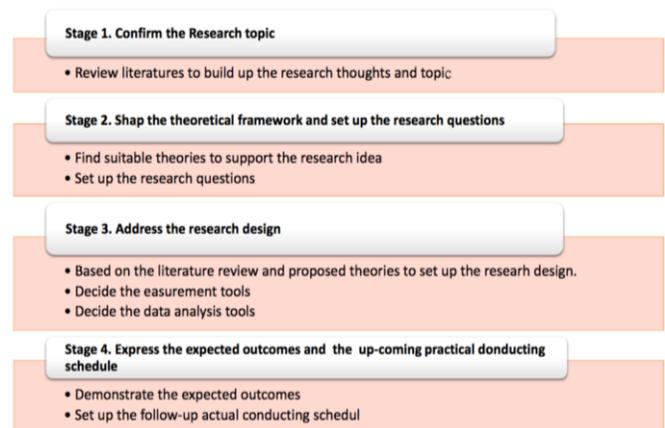


Figure 2. The proposed research timeline and schedule

ACKNOWLEDGMENT

This research proposal was supported by the doctoral seminar of Human Information Behavior in the School of Information Science and Learning Technology, University of Missouri. The authors would like to thank the instructor, Dr. Sanda Erdelez, who provided insight perspectives and feedbacks, the doctoral students who participated in the seminar, and all of the reviewers for their useful comments.

REFERENCES

- [1] Aula, A., Jhaveri, N., & Käiki, M., "Information search and re-access strategies of experienced web users," the Proceedings of the 14th international conference on World Wide Web, 2005.
- [2] Bandura, A., "Self-efficacy: toward a unifying theory of behavioral change," Psychological Review, vol. 2., 1997, pp. 191.
- [3] Bandura, A., "Self-efficacy In VS Ramachaudran (Ed.)," Encyclopedia of Human Behavior, vol. 4. New York: Academic Press, 1994, pp. 271–350.
- [4] Bandura, A., "Cultivate self-efficacy for personal and organizational effectiveness," Handbook of principles of organization behavior, 2000, pp. 2.
- [5] Brand-Gruwel, S., Wopereis, I., & Vermetten, Y., "Information problem solving by experts and novices: Analysis of a complex cognitive skill," Computers in Human Behavior, vol. 3, 2005, pp. 487-508.

- [6] Catalano, A., "Patterns of graduate students' information seeking behavior: A meta-synthesis of the literature," *Journal of Documentation*, vol. 2, 2013, pp. 243-274.
- [8] Crawford, J. R., & Henry, J. D., "The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample," *British Journal of Clinical Psychology*, vol. 3, 2004, pp. 245-265.
- [9] Creswell, J. W., *Educational research: Planning, conducting, and evaluating quantitative*: Prentice Hall Upper Saddle River, NJ., 2002.
- [10] Eastin, M. S., & LaRose, R., "Internet self-efficacy and the psychology of the digital divide," *Journal of Computer-Mediated Communication*, vol. 1, 2000, pp. 0-0.
- [11] Ellis, D., & Haugan, M., "Modelling the information seeking patterns of engineers and research scientists in an industrial environment," *Journal of Documentation*, vol. 4, 1997, pp. 384-403.
- [12] Fisher, K. E., Erdelez, S., & McKechnie, L., *Theories of Information Behavior*: American Society for Information Science and Technology, 2005.
- [13] Hsu, M.-H., & Chiu, C.-M., "Internet self-efficacy and electronic service acceptance," *Decision support systems*, vol. 3, 2004, pp. 369-381.
- [14] Joo, Y.-J., Bong, M., & Choi, H.-J., "Self-efficacy for self-regulated learning, academic self-efficacy, and Internet self-efficacy in Web-based instruction," *Educational Technology Research and Development*, vol. 48, 2000, pp. 5-17.
- [15] Kim, Y., & Glassman, M., "Beyond search and communication: Development and validation of the Internet Self-efficacy Scale (ISS)," *Computers in Human Behavior*, vol. 4, 2013, pp. 1421-1429.
- [16] Korobili, S., Malliari, A., & Zapounidou, S., "Factors that influence information-seeking behavior: The case of Greek graduate students," *The Journal of Academic Librarianship*, vol. 2, 2011, pp. 155-165.
- [17] Lin, Y.-C., Liang, J.-C., Yang, C.-J., & Tsai, C.-C., "Exploring middle-aged and older adults' sources of Internet self-efficacy: A case study," *Computers in Human Behavior*, vol. 6, 2013, pp. 2733-2743.
- [18] Liyana, S., & Noorhidawati, A., "How graduate students seek for information: Convenience or guaranteed result?" *Malaysian Journal of Library and Information Science*, vol. 2, 2017.
- [7] Chuang, S.-C., Lin, F.-M., & Tsai, C.-C., "An exploration of the relationship between Internet self-efficacy and sources of Internet self-efficacy among Taiwanese university students," *Computers in Human Behavior*, vol. 48, 2015, pp. 147-155.
- [19] Meho, L. I., & Tibbo, H. R., "Modeling the information-seeking behavior of social scientists: Ellis's study revisited," *Journal of the Association for Information Science and Technology*, vol. 6, 2003, pp. 570-587.
- [20] Nwagwu, W., & Adebayo, O., "Computer Anxiety and Computer Self-Efficacy in Computer-Based Tests in Selected Universities in South-West Nigeria," *African Journal of Library, Archives & Information Science*, vol. 1, pp. 2916.
- [21] Savage, M. W., & Tokunaga, R. S., "Moving toward a theory: Testing an integrated model of cyberbullying perpetration, aggression, social skills, and Internet self-efficacy," *Computers in Human Behavior*, vol. 71, 2017, pp. 353-361.
- [22] Savolainen, R., "Emotions as motivators for information seeking: A conceptual analysis," *Library & Information Science Research*, vol. 1, 2014, pp. 59-65.
- [23] Talja, S., "Information sharing in academic communities: Types and levels of collaboration in information seeking and use," *New Review of Information Behavior Research*, vol. 1, pp. 143-159.
- [24] Thatcher, A., "Web search strategies: The influence of Web experience and task type," *Information Processing & Management*, vol. 3, 2002, pp. 1308-1329.
- [25] Timmers, C. F., & Glas, C. A., "Developing scales for information-seeking behaviour," *Journal of Documentation*, vol. 1, 2010, pp. 46-69.
- [26] Torkzadeh, G., & Van Dyke, T. P., "Development and validation of an Internet self-efficacy scale," *Behaviour & Information Technology*, vol. 4, 2001, pp. 275-280.
- [27] Wang, J.-L., Jackson, L. A., Wang, H.-Z., & Gaskin, J., "Predicting social networking site (SNS) use: Personality, attitudes, motivation and internet self-efficacy," *Personality and Individual Differences*, vol. 80, 2015, pp. 119-124.
- [28] Wu, Y.-T., & Tsai, C.-C., "University Students' Internet Attitudes and Internet Self-Efficacy: A Study at Three Universities in Taiwan," *Cyberpsychology & behavior*, vol. 4, 2006, pp. 441-450.

# Customer Churn Analysis : A Case Study on the Telecommunication Industry of Thailand

Paweena Wanchai

Department of Computer Science, Faculty of Science  
Khon Kaen University  
Khon Kaen, Thailand  
wpaweena@kku.ac.th

**Abstract**— Customer churn creates a huge anxiety in highly competitive service sectors especially the telecommunications sector. The objective of this research was to develop a predictive churn model to predict the customers that will be to churn; this is the first step to construct a retention management plan. The dataset was extracted from the data warehouse of the mobile telecommunication company in Thailand. The system generated the customer list, to implement a retention campaign to manage the customers with tendency to leave the company. WEKA software was used to implement the followings techniques: C4.5 decision trees algorithm, the logistic regression algorithm and the neural network algorithm. The C4.5 algorithm of decision trees proved optimal among the models. The findings are unequivocally beneficial to industry and other partners.

**Keywords**-component; churn prediction; telecommunication industry; data mining; CRM; decision trees

## I. INTRODUCTION

Nowadays, the telecommunications industry is facing dramatic changes due to the influence of information technology and global competition. As the telecommunications market grows increasingly competitive, customer expectation is increasing as well. Today's customers can exercise their purchase power by choosing from many carriers or product providers to satisfy their communication needs. Accordingly, the major task facing the telecommunications industry is how to provide better services and better quality at a lower cost. There is a high competitive market; nevertheless, the companies have been working aggressively with, promotions, marketing plans or other approaches to obtain the highest and deepest marketing immersion. For that reason, it is valuable to know which customers are likely to switch to a competitor in the near future. Those customers are so called churned customers. Churn is caused by several common reasons such as dissatisfaction with the services and high bills. Moreover, customers often receive attractive offers when signing up with a new mobile service provider. Customer churn is a crucial activity in rapidly growing and competitive telecommunication sector, due to the high cost of acquiring new customers [1]. Previous research points out that there is a higher churn rate of customers in Asia than in the US and Europe.

Telecommunication companies in Thailand are operating in a highly competitive and challenging market environment.

With the introduction of the Mobile Network Portability (MNP), a number of voice and data subscribers had the leverage of churning from one subscriber to the other. Customers regularly port their mobile numbers from one telecommunication service provider to the other thereby making the companies lose large amount of revenue. Telecommunication companies in Thailand have had a herculean task in predicting expected churn customers using modern computing statistical tools.

The retention of customers is key in any business. This is because, the cost of retaining existing customers is much cheaper than the acquisition of new customers [2, 3]. The concept of identifying new customers as a marketing strategy by telecommunication companies in Thailand is much hectic than the retention of existing ones. However, it has also become quite a challenge for these companies to identify potential churn customers to respond to their needs for retention. Modeling customer lifetime value is therefore one way to detect, compute customer value and predict potential customer churn. One of the best's solutions is to use the preventive methods when customers are close to churn. The telecommunications industry produces huge amounts of data and is bedeviled with a vast array of imperfect customer information that decision makers need to deal with. High performance data warehouses and powerful business intelligence solutions enable experts to access, extract and manipulate a huge amount of data needed for churn prediction modeling. By applying data mining methods and techniques, meaningful patterns and decision supportive information can be discovered in this huge amount of data. There are many data mining techniques to predict when customers are churn, for example: support vector machines, logistic regression models, decision tree and neural network. This simulator will help the company to have the better variables for monitoring or predict customer behavior to get future customers to churn.

The research employs some algorithms of data mining, based on machine learning and statistical computing; to develop a predictive model for customer churn in the telecommunication industry in Thailand. The purpose of this research is to develop an accurate methodology to predict when the customers are going to churn in telecommunications companies (postpaid customers). These detected customers are will be using to generate to retention campaigns.

## II. LITERATURE REVIEW

A. *Theoretical Foundation of Data Mining*

Data mining represents a multifaceted range of technologies that are rooted in disciplines like statistics, mathematics, computer science and engineering among others [4, 5, 6]. Data mining comprises the process of exploring and modeling large datasets to elicit useful results and patterns [7]. Data mining is rooted in three sections: classical statistics, machine learning and artificial intelligence (AI) [8]. Data mining uses algorithms like artificial neural networks, time series analysis, association rules, clustering, regression, classification, and many others to mine relevant information for decision making and prediction [9, 10]. Classification is defined as the way to discover various characteristics in management, association is defined as rules of affinity among collected data and clustering is defined as a process of segmentation. Data mining techniques mostly used in customer relationship management (CRM) are decision trees, neural networks, association rules, sequence discovery among others. The tools and methodologies of data mining are designed primarily to discover hidden patterns to aid in decision making.

There are mainly two types of data mining techniques that are used in practice: supervised learning and unsupervised learning [11]. Supervised learning requires that the dataset should contain target variables that represent the classes of data items or the behaviors that are going to be predicted. The most important decision in customer churn management is the separation of churners from non-churners. This task is quite capably handled by supervised learning techniques. Data mining can be analyzed in two forms: predictive and descriptive. Descriptive analysis deals with classification data into sequence, patterns and trends for decision making. Predictive mining uses classified data to forecast for the future using various algorithms.

The process of building models using data mining techniques for predictive output is known as predictive modeling. It is a process for transforming data into meaningful output that can affect business decisions. This is accomplished by analyzing and modeling the stored information or data. In its stored form, information is mostly used for reporting the events occurred and observations made. Once the data is used to build a model it becomes a vital tool to recognize patterns and predicting future actions or behavior. Various steps are involved in building such a model:

- The data is collected from various sources in the business process.
- Data is then analyzed for patterns and relationships.
- A statistical technique is chosen that best suits the data and output desired.
- A model is built based on the technique chosen.
- The model is validated for usability and accuracy.

B. *Customer Churn Prediction Modeling*

Customer churn refers to when the customer cancels the contract. When the customer leaves the company prematurely, cancels one contract, two contracts or cancels the full account.

The companies can divide the churn in two categories: voluntary and involuntary. The involuntary churn is when happen involuntary situations with the customer like move to other location, not paid the bill or customer dies. When the companies use data mining for analyze the churn, the involuntary churn is excluded, because in these types of cancellation the companies not have control. The voluntary churn is when the customer cancels the contract but the reasons are based on company customer relations. This type of churn has the probability of retention. The companies have many strategies to fight it, but sometimes they use their capital and efforts for gain more customers.

Churn in the telecommunication industry has also defined as an action that leads to the cancellation of a consumer's telecommunication services. Churn can be classified into two categories such as, service provider initiate churn which is when the service provider deliberately suspends a consumer's account due to payment default, the other type of churn is, consumer initiated churn. Consumer initiated churn is often more complicated than company initiated churn. Consumer may switch or choose to cancel their subscription due to multiple reasons.

A number of researchers have used varied techniques to address churn in various fields. In the churn analysis and modeling of the telecommunication industry, common algorithms deployed include decision trees (DT), logistic regression (LR), neural networks (NN), Naïve Bayesian classifiers (NBC), support vector machine (SVM), linear discriminant analysis (LDA), self-organizing maps (SOM), among others. These techniques are applied in predicting both qualitative and quantitative data and the interpretation of the predictive models created.

Hung, Yen and Wang [12] employed decision tree and multilayer perception neural networks techniques to build a predictive churn model. In order to evaluate the techniques, the researchers used dataset from a telecommunication company in Taiwan using several groups of attributes including customer demographics, call detail records, billing information, service change log, and contract/service status. The research revealed that the attributes that are substantial to differentiate between churners and non-churners are age, gender, tenure, billing amount, number of overdue payment, in-net call duration, and number of changes in account information. The researchers also used K-means clustering technique to segment the customers according to their billing amount, tenure and usage. One decision tree is created for each cluster. A base decision tree is also created for all customers without any segmentation. The finding showed no significant difference between the performances of decision trees with and without customer segmentation. Comparison of decision tree and multilayer perception neural networks verifies that neural networks perform better on this particular dataset. Findings from this study indicated that data mining techniques were useful in accurately determining possible churners.

Owczarczuk [2] evaluated different data mining techniques for chum prediction model including decision trees, Fisher linear discriminant analysis, and logistic regression. The

dataset of this study consisted of 85,274 customer records from pre-paid customers. Overall, all techniques returned similar results for low quantiles, but for medium and high quantiles, it was shown that logistic regression outperformed the other techniques [2]. Oseman, Mond, Haris and Abu [13] used the ID3 (Iterative Dichotomiser 3) algorithm to create a churn prediction model. ID3 is an algorithm invented by Ross Quinlan. There are two steps in classification of decision trees; first step, use the training set to establish decision trees and second step, use the decision trees to make classification to input record [14]. The dataset that they used has the following variables: the length of service, area and total of more than 10 minutes customer engaged. The results suggested that area was the main factor for customer to churn, apart from two minor causes for customer to churn [13].

Jin, Meng, Fan, Peng and Chen [15] created predictive models using decision trees C4.5 and back propagation neural network (BPN). For this study, the dataset that was used were extracted from data warehouse of mobile services providers including churning/non churning, billing, call records details, customer service detail and customer care data. The research used the following variables: gender, age, area, tenure, inbound call, outbound call, domestic call, bill amount, payment, overdue payment, monthly fee, inquire, phone no changed, and bar/suspend. Neural network was used to calculate the propensity for a customer to churn while the decision trees describe the behavior of the churners

The literature shows that various data mining techniques for churn prediction in the telecommunication industry has been used such as neural network, regression analysis, decision tree, SVM, Naïve Bayes and Bayesian network, evolutionary approach and neuro-fuzzy [16, 17, 18, 19]. However, conflicts arise when studying results and comparing the conclusion of few of these published researches in the area of churn prediction. It is also observed that most of the studies have only evaluated a limited number of data mining techniques with small sample's size. Therefore, the most important problem of which classification technique could use to approach the customer churn prediction in a more appropriate fashion, still remains an open research problem.

### III. METHODOLOGY

The main research objective is to develop an accurate methodology to predict when the customers are going to churn in telecommunications companies, developing strategies to retain customers. The churn indicator is very important and powerful in the company business and if it is accurate the results campaign will be having better results and consequently a few loss of money. In order to achieve the research objectives, the work performed can be divided into the following steps.

#### A. Business and Data Understanding

This phase start with the business perspective and data understanding. The research objective is translated into a realistic data mining problem which is possibly solvable with respect to foreseen limitations. In this phase, the researcher gets to know which data are available and how to collect them.

Exploration of data gives a basic understanding. In the business context, it is essential that the researcher gain her knowledge not only from the data itself but also from the domain experts. The telecommunication company provides the author of this research full access to their data warehouse and initially a sample of available data such as customer demographics, subscribers history, billing history, account customer details, retention history, customer consume information and others. Time and effort also spent on studying the architecture of the data warehouse and how to query the data from different sources. This step also include identifying data problems like null values or fields that not have value or the values are not significant. The different fields' values or combination was identified for make the cross matching information.

#### B. Data Collection and Preparation

The pre-processing step involves transferring the raw data from the data warehouse to the analysis platform. This was the most difficult and time-consuming phase in data mining. In this phase when the data is collected, integrated and cleaned. The first step in data collection, this phase has random customer that is selected with filters. The final table should be cleaned of ambiguities or bad records. The pre-processing or data preparation phase could be basic steps like organizing data and structuring it into required format. The structured data is then to be checked for missing values in critical attributes, as presence of these will hamper the results of the modeling. This usually involves changing the format or type of the data into a specific manner to suit the usage of mining tools to be used.

The second part was prepared the calculate variables such as average billing, number of calling customer services with problems, number of data purchase, total internet charge, total day minutes and customer age. These calculated fields are necessary for predict the churn; normally it was very difficult to obtain the necessary data to prepare these calculations. Not all fields are obtained in the same source or database, they needed a computing process or for example wait to closed cycle period to obtain the information.

The dataset was extracted from the data warehouse of the mobile telecommunication company in Thailand. To protect customer privacy, the data source randomly selected based on their telephone numbers from January 2016 to December 2016. The whole dataset consists of 262,500 customer records: 106,800 churners and 155,700 non-churners. The dataset was randomly split into 3 parts: training 157,500 records (60%), validation 52,500 records (20%) and testing 52,500 records (20%). Missing data were treated using listwise deletion by deleting any case that had a missing value on any of the variables included in each test. The attributes mainly consist of the following information:

Demographic profiles: describes the demographic information, including age, gender, etc.

Account information: all information about customer accounts, e.g., account number/type, fees, payment type, account balance, call information, etc.

Call details: this group of information describes many aspects related to different types of calls (e.g., international or local calls), number of calls, call duration, costs, etc.

Table 1 has a details description of the dataset.

TABLE I. LIST OF ATTRIBUTES

Attribute name	Description
Age	Customer age
Gender	Account customer gender
Area	Customer area
Tenure	Length of time a customer has been with the organization
Internet plan	Boolean field, that has if the customer has internet plan contract
International plan	Boolean field, that has if the customer has international plan contract
Voice mail plan	Boolean field, indicate if the customer has voice mail plan
Num vmail messages	Number of daily voice mail message
Avg bill amount	Average bill amount, calculate in 12 months
Total day minutes	Total number of minutes that the customer used on the day
Total eve calls	Total number of calls that the customer made on the evening
Total eve charge	Calculation of evening charge is based in plan evening minutes charge
Total night minutes	Total number of minutes that the customer used on the night
Total night calls	Total number of calls that the customer made on the night
Total night charge	Calculation of night charge is based in plan night minutes charge
Total intl minutes	Total number of minutes that the customer used to made international calls
Total intl calls	Total number of calls that the customer made on international plan
Total intl charge	Calculation of internet charge is based in international plan minutes charge
Credit purchase	Total amount used to purchase airtime a month
Num credit purchase	Number of times, that the customer used to purchase airtime a month
Data purchase amount	Total amount used to purchase a data package a month
Num data package purchase	Number of times, that the customer used to purchase data a month
Total internet charge	Calculation of internet charge is based in internet plan data charge
Num customer service calls	Total of monthly calls that the customer made to spoke with customer service.
Num retention	Number of times, that the customer received a retention package
Churn	Represents if the customer is no longer with the company, (Churn/Not Churn)

### C. Attributes Selection

Feature selection are important steps in the knowledge discovery process, to identify those relevant variables or attributes from the large number of attributes in a dataset which are too relevant and that can reduce the computational

cost. The selection of most appropriate attributes from the dataset in hands, was carried out using feature ranking method titled as “Information Gain Attribute Evaluator”, using an WEKA toolkit. It evaluates the attributes worth through the information gain measurement procedure as per the class value. It’s diverse the selection and ranking of attributes that significantly improves the computational efficiency and classification. After feature ranking, it includes most relevant and ranked attributes in the decision table. Table 2 shows the results of selected attributes.

TABLE II. SELECTED ATTRIBUTES RESULTS

Attribute Evaluator	Results
Information Gain Ranking Filter	Internet plan Num data package purchase Num customer service calls Tenure Age Num retention Avg bill amount Total day minutes Data purchase amount Total internet charge Churn

### D. Modeling

In this step, the different techniques were tested, implemented with different options. WEKA was used as data mining software to implement different techniques and select the better technique according to the dataset provided.

#### 1) Decision Trees Analysis

Decision trees are classification techniques that partition data in a recursive manner into smaller divisions based on some algorithms. One advantage of decision trees analysis is that they are nonparametric and no assumptions are used in relation to input data [20]. Decision trees are able to handle nonlinear data, missing values, numeric and categorical data that makes it idyllic for predictive modeling for churn management since primary data from customers are used. Decision trees are formulated using key variables related to previous variables in training models to predict future outcomes, churn intentions of customers.

C4.5 is an algorithm developed by Ross Quinlan in 1993 to generate decision trees. It is an extension of an earlier version called ID3. C4.5 known in WEKA as J48 is an algorithm used to generate decision trees. After a J48 fully grown decision tree is constructed, pruning is carried out using 10 folds cross-validation on training sets. C4.5 builds decision trees from a set of training data using the concept of information entropy. C4.5 examines the normalized information gain (difference in entropy), which is also called gain ratio, that results from choosing an attribute for splitting the data. Gain ratio is a modification of the information gain that reduces its bias on high-branch attributes. Suppose  $S$  represents a dataset and  $A$  is an attribute. Values  $A$  is the set of all possible values for attribute  $A$ , and  $S_v$  is the subset of  $S$  for which attribute  $A$  has value  $v$ . Intrinsic information measures how much information we need to tell which branch an instance belongs to. It is defined as:

$$IntrinsicInfo(S, A) = - \sum_{i=1}^{|S|} \left( \frac{|S_i|}{|S|} \right) \log_2 \left( \frac{|S_i|}{|S|} \right) \quad (1)$$

The information gain is defined as:

$$Gain(S, A) = Entropy(S) - \sum_{v \in values(A)} \left( \frac{|S_v|}{|S|} \right) Entropy(S_v) \quad (2)$$

Gain Ratio normalizes information gain by:

$$GainRatio(S, A) = \frac{Gain(S, A)}{IntrinsicInfo(S, A)} \quad (3)$$

Gain ratio takes the number and size of branches into account when choosing an attribute; thus, it overcomes the bias of the information gain measure, which favors attributes with many values over those with few values. The attribute with the highest normalized information gain is the one used to make the decision. The algorithm then recurs on the smaller lists. C4.5 handles missing data by assigning a probability to each possible value of the related attribute. The assigned probability is estimated by the observed frequency of the attribute value among the instances at the tree node. C4.5 handles continuous-valued attributes by dynamically defining a best cut point of the attribute. After sorting the instances according to the attribute values, C4.5 discretizes values with midpoints as candidate thresholds; the value of the threshold that maximizes information gain must always lie at the boundaries, so the best cut point is the one that maximizes information gain.

### 2) Logistic Regression

Logistic regression is a predictive modeling technique where there is a correlation between the probability of a result and its predictor variables as seen in equation (4).

$$\log \left[ \frac{\pi_i}{1 - \pi_i} \right] = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} \quad (4)$$

where

- $\pi_i$  is the probability of the outcome,
- $\beta_1, \dots, \beta_k$  are coefficients,
- $X_1, \dots, X_{ik}$  are predictor variables.

The  $\beta$  coefficients are transformed into odds ratios with the degree of importance of predictors well known. The Hosmer-Lemeshow statistic is accepted extensively in assessing the goodness of fit of developed models in logistic regression with a dichotomous outcome [21]. In a specific manner, logistic regression details the linear function of observed attributes for endogenous variables as the fitted probability of event. The fitted probability, logit ( $\pi_i$ ), is defined as:

$$logit(\pi_i) = \log \left[ \frac{\pi_i}{1 - \pi_i} \right] \quad (5)$$

### 3) Neural Networks

Neural networks are used for descriptive and predictive data mining. Neural network is the linking of neurons (computed units) with respect to their weights [22]. Artificial neural networks computes based on input signals and importance weight just as the brain does computations. The input signals conduct a combination function with the weights and threshold value, and activated by the activation function to produce an output signal as seen in equation (6).

$$y_i = f(x, w_j) = f(P_j) = f \left( \sum_{i=0}^n x_i w_{ij} \right) \quad (6)$$

Where  $j$  is a generic neuron,  $x$  is the input signals,  $w_j$  are the weights,  $P_j$  is the potential and  $y_j$  is the output signal.

Neural networks fit non-linear functions very well in addition to recognizing patterns. The algorithm is used in a wide range of fields such as aerospace, automotive, banking, defense, electronics, entertainment, financial, insurance, manufacturing, oil and gas, robotics, telecommunications, and transportation industries [22].

## IV. EXPERIMENTAL RESULTS

### A. Classifiers Results and Performance Evaluation

In order to assess the developed model and compare it with different data mining techniques used for churn analysis, this research use the confusion matrix shown in Table 3 which is the primary source for evaluating classification models. A correctly classified instance is counted as a true positive (TP) or a true negative (TN) if its actual class is positive or negative respectively. A positive instance which is misclassified as negative is counted as a false negative (FN) and a negative instance which is misclassified as positive is counted as a false positive (FP).

The research compare 3 techniques: decision trees C4.5, logistic regression and neural networks. The best result is the decision trees C4.5 with 93.71% of correctly classifiers and 6.29% of incorrectly classifiers. This 93.71% is the result of all class classification includes when the customers churn and not churn. To compute this number are summarized the true positives values and the false positives values, then the correctly classified instance are the true positives divided between the sum of all the instances, the incorrectly classified is the sum of the false positives divided between the sum of all instances. The confusion matrix gives the correct classification and the incorrect classification. The accuracy and error rate of the predicted results are shown in Table 3. The accuracy rate and error rate are computed as shown in the following equations.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (7)$$

$$Error\ rate = \frac{FN + FP}{TP + FN + FP + TN} \quad (8)$$

TABLE III. CLASSIFIER RESULTS

Technique	Classified Instances		Confusion Matrix		Actual Class
	Accuracy	Error rate	Not Churn	Churn	
Decision Trees C4.5	93.71%	6.29%	29,180	1,794	Not Churn
			1,510	20,016	Churn
Logistic regression	90.88%	9.12%	28,629	2,344	Not Churn
			2,442	19,085	Churn
Neural networks	86.13%	13.87%	26,678	4,296	Not Churn
			2,988	18,539	Churn

### B. Churn and Not Churn Detection Rules

This step applies the C4.5 (J48) rules directly to the dataset and analysis the result to create the rules. These rules are used to detect correlated information with churn detection. The following are the example rules detection churn information:

- If (internet plan) = 1 and  $20 \geq (\text{age}) \leq 25$  and (data purchase amount > 1,000) then the tendency is to churn.
- If (internet plan) = 1 and (total internet charge) > 1,500 and (num retention) > 3 then the tendency is to churn.
- If (internet plan) = 1 and (num customer service calls) > 5 and (num data package purchase) > 7 then the tendency is to churn.
- If (tenure) < 2 and  $26 \geq (\text{age}) \leq 30$  and (avg bill amount) > 2,000 then the tendency is to churn.
- If (internet plan) = 0 and (avg bill amount) < 100 and (total day minutes) < 2 and then the tendency is to churn.

### C. Rules Analysis

The churn detection is a very complex process and is very difficult to detect the behavior of churn with one attribute, but when data mining techniques are used, the technique apply the predictions. Data mining systems are the solutions for detection and to create the better and useful campaigns, which will help the companies to decrease the churn. Taking all this information in mind, these are the example tendencies:

- If the customer has internet plan and the customer age is between 20-25 and the total amount used to purchase a data package a month are more than 1,000 Thai baht (THB) the tendency is to churn.
- If the customer has internet plan and the total internet charge are more than 1,500 THB and the total number of times that the customer received a retention package are more than 3 times the tendency is to churn.
- If the customer has internet plan and the customer calls to customer services are more than 5 times and the total number of times that the customer used to purchase a data package a month are more than 7 times the tendency is to churn.
- If the customer tenure is less than 2 years and the customer age is between 26-30 and the average bill

amount are more than 2,000 THB the tendency is to churn.

- If the customer has no internet plan and the customer's average monthly payment for the past 12 months are less than 100 THB and the total day minutes used are less than 2 minutes the tendency is to churn.

## V. CONCLUSION

In the mobile telecommunication industry, customers are able to choose among multiple service providers and actively exercise their rights of switching from one service provider to another. In this fiercely competitive market, customers demand tailored products and better services at lower prices, while service providers constantly focus on acquisitions as their business goals. In the telecommunication market, competition is very high and the products and offerings are more and more comparable. This leads to reduced customer loyalty. Losing an existing high-volume customer means losing lots of revenue. It is more expensive to gain a new customer than to retain an existing one. It is necessary to identify customers that are willing to move to a competitor before they do so. Developing solutions to model and understand churn before the customer cancels is critical to the success of the telecommunications service providers.

To model the solution, the beginning process was to prepare and analyze the company dataset. All simulations were done with WEKA software. Then, the second step was to select the most significant dataset attributes. The third step was to classify several techniques that were tested, the techniques are the following: decision trees C4.5, logistic regression and neural networks, where the better percent of classified churn was decision trees C4.5 with 93.71% of classified correctly the churn and not churn. This is a great percent and will be enough to create a good predicted system and working together with proactive retention campaigns, will be represent revenues to companies.

Data mining is a significant tool in the telecommunication industry that can utilize the large volume of data generated for pattern analysis. The recent increasing embrace of predictive algorithm of data mining has given room for companies to assess their future success, challenges and targets. The study brings to fore the relevant untapped customer data and knowledge for churn prediction and customer classification for better decision making and customer management in Thailand. Even though the telecommunication industry is applied in this research, the quality/value relationship obtained is quite suggestive of results that can be derived for more sectors, hence the model can be used by companies in the service sector for customer churn analysis with same predictor variables. Applying this technique in predicting the behavior of the most valuable asset (customers) in the telecommunication industry in Thailand and the implementation of the developed model, guarantees a higher level of customer assessment, customer management and customer profiling for continuous growth in the sector.

One important research, not investigate in this research are prepaid, combined services and others. Each service is individual and the characteristics are not the same, this means that the datasets are different. Each research needs individual time and preparation and maybe the select technique in postpaid is not the same for prepaid customers. In facts, all the classifiers tested have a good percent of accuracy and could be used to calculate customer churn, this is important to know because maybe in other research depending of the data the selected technique is different. The important point for future research is to make simulations with several techniques; because the dataset structure and information will change all the results.

#### REFERENCES

- [1] Hong, T.-P., Horng, C.-Y., Wu, C.-H., and Wang, S.-L.: 'An improved data mining approach using predictive itemsets', *Expert Systems with Applications*, 2009, 36, (1), pp. 72-80
- [2] Owczarczuk, M.: 'Churn models for prepaid customers in the cellular telecommunication industry using large data marts', *Expert Systems with Applications*, 2010, 37, (6), pp. 4710-4712
- [3] Ngai, E.W., Xiu, L., and Chau, D.C.: 'Application of data mining techniques in customer relationship management: A literature review and classification', *Expert systems with applications*, 2009, 36, (2), pp. 2592-2602
- [4] Wu, X., Zhu, X., Wu, G.-Q., and Ding, W.: 'Data mining with big data', *IEEE transactions on knowledge and data engineering*, 2014, 26, (1), pp. 97-107
- [5] Li, D., Wang, S., and Li, D.: 'Spatial Data Mining: Theory and Application' (Springer, 2016.)
- [6] Shmueli, G., and Lichtendahl Jr, K.C.: 'Data Mining for Business Analytics: Concepts, Techniques, and Applications in R' (John Wiley & Sons, 2017.)
- [7] Peng, Y., Kou, G., Shi, Y., and Chen, Z.: 'A descriptive framework for the field of data mining and knowledge discovery', *International Journal of Information Technology & Decision Making*, 2008, 7, (04), pp. 639-682
- [8] Han, J., Pei, J., and Kamber, M.: 'Data mining: concepts and techniques' (Elsevier, 2011.)
- [9] Chattamvelli, R.: 'Data mining algorithms' (Alpha science international, 2011.)
- [10] Linoff, G.S., and Berry, M.J.: 'Data mining techniques: for marketing, sales, and customer relationship management' (John Wiley & Sons, 2011.)
- [11] Coussement, K., and Van den Poel, D.: 'Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques', *Expert systems with applications*, 2008, 34, (1), pp. 313-327
- [12] Hung, S.-Y., Yen, D.C., and Wang, H.-Y.: 'Applying data mining to telecom churn management', *Expert Systems with Applications*, 2006, 31, (3), pp. 515-524
- [13] Oseman, K., Haris, N.A., and Abu Bakar, F.: 'Data Mining in Churn Analysis Model for Telecommunication Industry', *Journal of Statistical Modeling and Analytics Vol*, 2010, 1, (19-27)
- [14] Danwa, S., Ning, H., and Dandan, L.: 'Construction of Forestry Resource Classification Rule Decision Tree Based on ID3 Algorithm', in Editor: 'Book Construction of Forestry Resource Classification Rule Decision Tree Based on ID3 Algorithm' (IEEE, 2009, edn.), pp. 867-870
- [15] Jin, S., Meng, Y., Fan, C., Peng, F., and Chen, Q.: 'The Research on Applying Data Mining to Telecom Churn Management', in Editor: 'Book The Research on Applying Data Mining to Telecom Churn Management' (2012, edn.)
- [16] Shaaban, E., Helmy, Y., Khedr, A., and Nasr, M.: 'A proposed churn prediction model', *International Journal of Engineering Research and Applications*, 2012, 2, (4), pp. 693-697
- [17] Kirui, C., Hong, L., Cheruiyot, W., and Kirui, H.: 'Predicting customer churn in mobile telephony industry using probabilistic classifiers in data mining', *IJCSI International Journal of Computer Science Issues*, 2013, 10, (2), pp. 1694-0814
- [18] Farquad, M.A.H., Ravi, V., and Raju, S.B.: 'Churn prediction using comprehensible support vector machine: An analytical CRM application', *Applied Soft Computing*, 2014, 19, pp. 31-40
- [19] Abbasimehr, H., Setak, M., and Tarokh, M.: 'A neuro-fuzzy classifier for customer churn prediction', *Int J Comput Appl*, 2011, 19, (8), pp. 35-41
- [20] Nabareseh, S., and Klimek, E.A.D.a.P.: 'Security on Electronic Transactions in Developing Countries: A Cluster and Decision Tree Mining Approach', in Editor: 'Book Security on Electronic Transactions in Developing Countries: A Cluster and Decision Tree Mining Approach' (Academic Conferences Limited, 2015, edn.), pp. 85
- [21] Hosmer Jr, D.W., Lemeshow, S., and Sturdivant, R.X.: 'Applied logistic regression' (John Wiley & Sons, 2013. 2013)
- [22] Pham, V.-T., Volos, C., Jafari, S., Wang, X., and Vaidyanathan, S.: 'Hidden hyperchaotic attractor in a novel simple memristive neural network', *Optoelectronics and Advanced Materials, Rapid Communications*, 2014, 8, (11-12), pp. 1157-1163

# The Influence of Advanced and Secure E-Commerce Environments on Customers Behaviour: The Case of Saudis in the UK

Haya Alshehri  
School of Computing, Science and Engineering  
University of Salford  
Salford, UK  
h.alshehri1@edu.salford.ac.uk

Farid Meziane  
School of Computing, Science and Engineering  
University of Salford  
Salford, UK  
f.meziane@salford.ac.uk

**Abstract** - The work reported in this paper is part of a larger study comparing the online activities of Saudi citizens living in Saudi Arabia and those living in the United Kingdom. That study aims to answer the question of whether the environment plays a key role in influencing the activities of Saudis when purchasing goods and services online. This paper considers only that part of the research conducted in the United Kingdom. It attempts to understand the activities and perception of Business to Customer E-Commerce among Saudis living in the UK, and hence what impact being away from their home environment actually has on their online shopping behaviour. Quantitative data was collected from 169 Saudis living in the United Kingdom. Trust in both security and payment were tested, with the result that a high number of Saudis resident in the UK show trust in the security and payment systems associated with online transactions in the United Kingdom. These primary outcomes suggest that the environment plays an important role in changing the shopping behaviours of online customers.

**Key words** - E-Commerce; Security; Saudi citizens; Payment; Behaviour

## I. INTRODUCTION

Currently, E-Commerce is one of the most discussed themes in business, and much research has been conducted on various issues related to it. E-Commerce allows companies to access new customers, as items and services can be offered to more geographically-dispersed buyers [7]. Zhu [15] identified many risks associated with the expansion and development of E-Commerce, and emphasised the impacts of this phenomenon on economies, politics, and the law. However, having developed in the Western world, E-Commerce struggles to achieve the same success in developing countries, and an examination of the success realised in Western countries is, therefore, useful in understanding what criteria must be fulfilled for that achievement. The United Kingdom is one of the most developed countries in Western Europe in its use of Business to Commerce (B2C) E-Commerce, and consequently represents a pertinent case for exploring how success is achieved in this respect. Kamalabadi et al. [8] have observed that many companies in developing countries are in the initial phases of implementing E-Commerce. In this connection, it is known that the poor security of the online environment with respect to personal information being illegally accessed is a major challenge to E-Commerce [6], and the optimal use of security measures is essential if the trustworthiness needed to assist the growth of B2C E-Commerce is to be generated [2]. Hence, steps must be taken by online retailers to prevent security breaches and fraud,

and this means that all potential security problems should be recognised, identified, controlled, and effectively managed/prevented [5].

This study therefore, aims to investigate whether and how Saudi citizens are influenced in their B2C E-Commerce behaviour by their perspectives on security and payment when residing in an Advanced E-Commerce Environment such as the United Kingdom. We endeavour to answer the questions posed by referring to the outcomes of previous studies and the data gathered during the empirical study.

This particular data set comes from a questionnaire survey of 169 Saudi citizens living in the United Kingdom. The key variables tested are the influences of security, and the online payment system, and these were explored via nine statements which respondents were asked to indicate their levels of agreement with. The following nine statements were presented, each one labelled as Security and Payment (SP).

- I do not mind providing my payment details to the UK companies (SP-statement 1)
- I do not mind providing my payment details to overseas companies (SP- statement 2)
- I prefer companies that provide different payment methods (SP- statement 3)
- Companies must have a secure payment system (SP-statement 4)
- Companies should make the security of the payments clear on their websites (SP- statement 5)
- The technology used to protect online payments is very important (SP- statement 6)
- I do not mind my payment details being stored by the company for future transactions (SP- statement 7)
- It is important that companies' websites include guidance explaining the payment method (SP-statement 8)
- I will buy online if my bank guarantees my transaction to be safe (SP- statement 9)

The rest of the paper is structured as follows. In section 2 we provide the background to the study in the form of a literature review, and in section 3 the research methodology is described. The analysis of the data and the subsequent findings are reported in section 4, and a discussion of these follows in section 5. Finally, in section 6, a conclusion is provided.

## II. LITERATURE REVIEW

### A. Background

In this age of technological growth, E-Commerce represents one of the important developments worldwide in the field of business and commerce. Through such growth, E-Commerce is essentially changing countries' economies, and the commercial methods by which they are managed [8]. are effective in their conduct of business, promote satisfaction and loyalty amongst their customers, who return to them in the sound knowledge that their purchasing will be trouble-free [7], and who, therefore, recognise a distinct benefit of this purchasing platform. Mostafaiepour [1] stated that one benefit associated with E-Commerce is the savings which it is possible for parties to a purchase/sale transaction to realise, since transportation, paperwork, and time-wasting are all reduced. However, in many developing countries, the Information and Communication Technology (ICT) infrastructures are still in their infancy and therefore, do not support the development of E-Commerce [13]. Al Ghamdi et al. [10] note that with respect to E-Commerce, the initial prediction of its growth in Saudi Arabia has not been met, and this is contrary to what is expected of a nation with Saudi Arabia's importance in the global economy. In addition, companies in Saudi Arabia do not appear to be following the developed countries' rapid growth in E-Commerce [10]. Alshehri and Meziane [3] have reported the absence of research that demonstrates any robust understanding of how Saudi customers behave in connection with online shopping when they are removed from their home country, and live instead in a developed environment. Nor has there been any research seeking to determine how to recreate the E-Commerce environment encountered in the advanced countries, in the Saudi context – an effort that would clearly assist the growth of E-Commerce in Saudi Arabia.

### B. TRUST

Trust is an important factor in buyer behaviour, and it plays a key role in E-Commerce activities because customers and vendors do not actually see each other. This marks a fundamental difference between E-Commerce and traditional business [13], in which the buyer and seller come into closer contact, and it is, therefore, essential to generate trust for E-Commerce efforts to be successful. Indeed, it is confirmed that one of the major obstacles to the adoption and development of E-Commerce is the lack of trust [13] felt by buyers in the ability of the seller to offer appropriate protection. Specifically in the Saudi context, it is known that over 50% of the Saudi population use the Internet between one to ten hours per day, and about 70% use it more than once a day [4]. In terms of the reluctance for this large slice of the Saudi population to engage in online purchasing, AlGhamdi et al. [10] reported the anxiety about 'stolen credit cards numbers' being common, as is the worry about companies' inability to provide "a secure payment platform". Indeed, these are the main reasons why Saudis tend not to engage in E-Commerce, as indicated by two of the interviewees in that study, who said respectively that "buying from outside Saudi Arabia is safer than buying from an e-retailer inside Saudi Arabia" and "people have perceptions that Arab e-retailers are not trustworthy".

### C. Security and Payment System

Many studies have shown that payment systems are the main obstacles to the development of E-Commerce. Malhotra et al. [9] for example, found that information privacy features as one of the main obstacles to the development of E-Commerce, and this was confirmed by the study conducted by Belanger et al. [2]. Likewise, online payment systems have been acknowledged as a key trust concern by customers [12]. In this connection, it is recommended that payment methods, all associated information, and a comprehensive explanation of all the steps that customers should follow in order to make a payment should be included on the website [11]. In this respect, Zhu [15] suggests the implementation of various strategies to increase security awareness among businesses, including educational and training initiatives in respect of E-Commerce, efforts to support the development of techniques for E-Commerce security, and to build environments for the promotion of effective E-Commerce [15]. The method of payment is the foremost problem which must be solved [1]. It has been shown that security and online payment are among the key factors that impact on Saudis living overseas when they make their decisions as to whether to buy online or not [3]. This stems from the fact that they come from Saudi Arabia where they perceive the several issues that are currently causing the slow development of E-Commerce, and they are programmed to look for these in the new environment. The lack of effective payment systems, and the difficulties encountered in transacting with banks are among these issues [14].

### B. B. TRUST

## III. RESEARCH METHODOLOGY

In designing this study, a quantitative approach was used, and a questionnaire designed to establish opinion among Saudi nationals living in the UK. This questionnaire was piloted as a means of gaining further information and of ensuring the validity and reliability of the instrument. The pilot involved distributing the questionnaire to Saudi nationals in different cities of the UK (Manchester, Liverpool, London, Norwich, Cardiff, and Bangor). The instrument was constructed in a logical order, initially asking for respondents' demographic data, and progressing to establish actual Internet usage, the experiences of the respondents in engaging with E-Commerce. It then considered the independent factors - "ICT infrastructure, Culture, Payment, Security, Privacy, Integrity, Personal Information, and Fulfillment of Transactions". Respondents were requested to indicate their level of agreement with each statement provided, using a five-point Likert scale ranging from 1 - "strongly disagree" to 5 - "strongly agree". The Statistical Package for the Social Sciences (SPSS) software was used to analyse the data obtained from the questionnaires.

IV. FINDINGS

A. Demographic Information

The demographic information relating to the Saudi citizens living in the UK and who responded is as follows: 80 were males (47.3%) and 89 females (52.7%); the age group 25-34 was the largest, containing 79 respondents (46.7%), and the age group over 55 was the smallest with just one respondent (6%).

With respect to the level of education possessed by the respondent, the largest group comprised individuals with a Bachelor’s degree, and they totalled 65, thereby representing 38.5% of the sample. In terms of location, the respondents were spread throughout the UK. The middle region attracted the highest number of respondents with 67.5%, followed by the west with 27.8%. With respect to income, 31.4% of the respondents earned between 4,000RS (Saudi Riyal) and 8,000RS, and 5.3% earned over 20,000RS. In terms of occupation, 39.1% of the sample were students, 37.9% were in employment, and 17.8% were the spouses of Saudis [5].

B. Analysis of Security and Payment (SP)

The following Figures show the analysis of all statements indicated in the Introduction section. The results are reported in percentages.

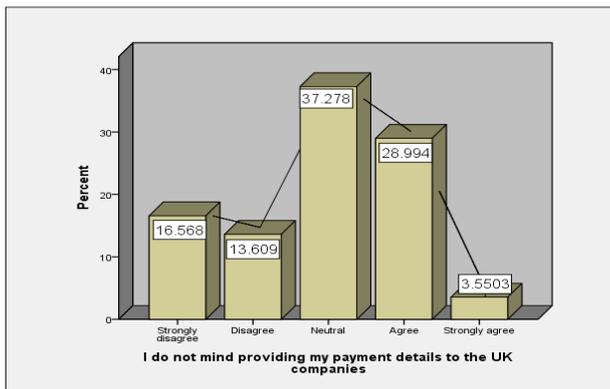


Figure 1: SP- statement 1

Fig. 1 concerning whether Saudis have any psychological problem with the idea of providing their payment details to UK companies shows that 3.5% agree strongly with the statement that this is not an issue for them, 16.5% disagree strongly, 28.9% agree, and 13.6% disagree. Hence, there was a fairly balanced response with 32.4% of respondents being in agreement and 30.1% not. A very large percentage (37.2%), however, did not give their opinion and recorded a neutral answer, possibly indicating that they had never been presented with that particular decision to make.

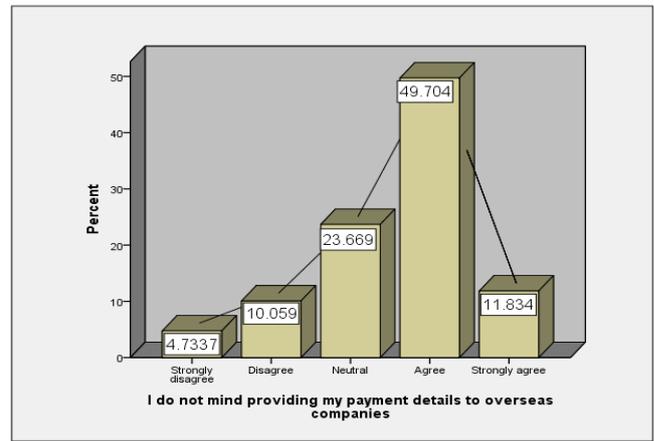


Figure 2: SP- statement 2

Fig. 2 relates to whether respondents were concerned about providing their payment details to overseas companies. The results obtained are as follows: 11.8% agree strongly with the statement that they are not bothered by this, 4.7% disagree strongly, 49.7% agree, and 10% disagree. Hence, 61.5% agree that they do not object to providing their payment details to overseas companies, while 14.7% disagree with the statement and clearly are apprehensive in this respect. However, almost one quarter (23.6%) of the population were not sure what they felt and recorded a neutral answer, again potentially because they had not engaged in E-Commerce and had never really had to think about this issue.

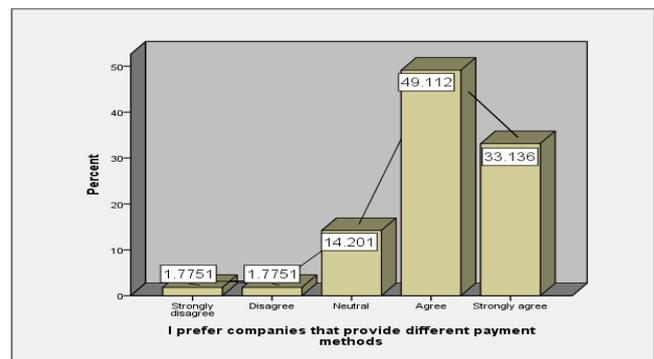


Figure 3: SP- statement 3

Fig. 3 refers to the statement that people prefer companies that provide different payment methods, and from this we can see that 33.1% agree strongly, 1.7% disagree, 49.1% agree, and 1.7% disagree. Overwhelmingly, a majority of 82.2% agree with this statement and only 3.4% disagree, so it is clear that most people prefer the opportunity to choose from different payment methods. However, 14.2% provided a neutral response, again signifying a small percentage of the research sample that probably had never given this issue much thought.

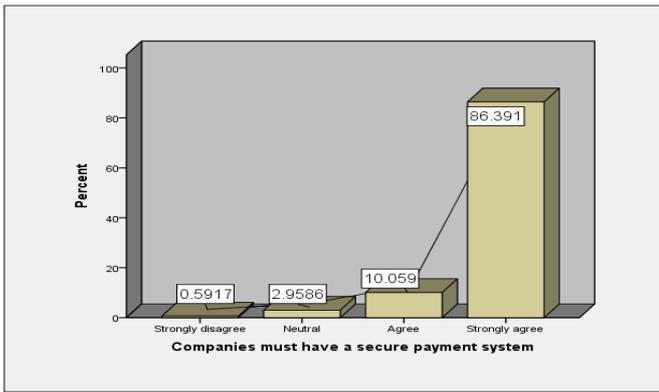


Figure 4: SP- statement 4

Fig. 4 relates to the statement that people believe companies must have a secure payment system, and in the reported answers, there was overwhelming agreement that this should in fact, be the case, as 86.3% of respondents recorded their strong agreement with the statement 10% reported their agreement, and only 0.5 % disagreed. Hence, the vast majority of the sample (96.3%) believed it to be necessary for online sellers to have a secure payment system. Again a very small percentage (2.9%) gave a neutral answer, which can be interpreted as representing that proportion of the sample who had either not engaged with online purchasing, or never thought about the issue of payment systems.

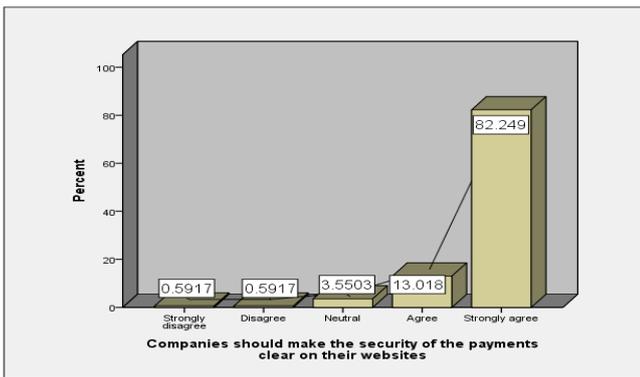


Figure 5: SP- statement 5

Fig. 5 indicates whether people believe that companies should be clear on their websites about their security arrangements in terms of customer payments, such that they confirm to customers that all payments are in fact secure, and therefore, allay their fears in this respect. In this respect 82.2% agree strongly that online companies should do this, a further 13% agreed with the statement, only 0.5% disagree strongly, and a further 0.5% disagreed. Therefore, overall, 95.2% of the population were in agreement that companies should be clear in this respect, and only 1% disagreed with this idea. Again, a small percentage (3.5%) gave a neutral answer. These results point to the definite need for companies to give proper attention to publicising their operational details in respect of payment security so that customers can make an educated judgement on whether they are prepared to make purchases online.

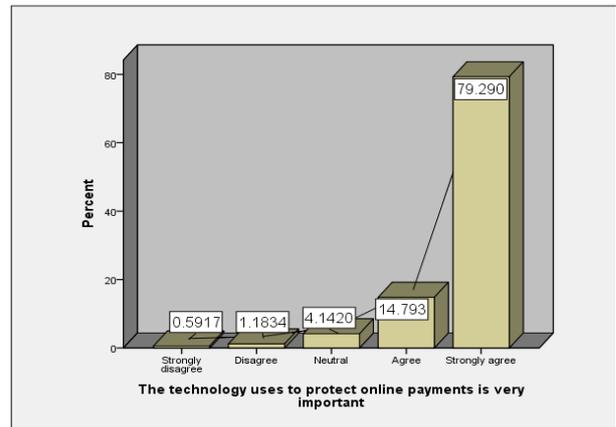


Figure 6: SP- statement 6

Fig. 6 presents the results of the statement that the technology used to protect online payment is very important, indicating whether respondents believe companies should pay due attention to their use of technology in their online activities. The responses to this statement show that: 79.2% agree strongly, with this statement, 14.7% agree, only 0.5% disagree strongly, and 1.1% disagree, thereby revealing that a huge majority of participants (93.9%) believe that companies should capitalise upon their technological ability to protect their online payments. A small percentage (4.1%) of the sample did not give an opinion, again indicating that this tiny proportion of people were not really engaged in online purchasing.

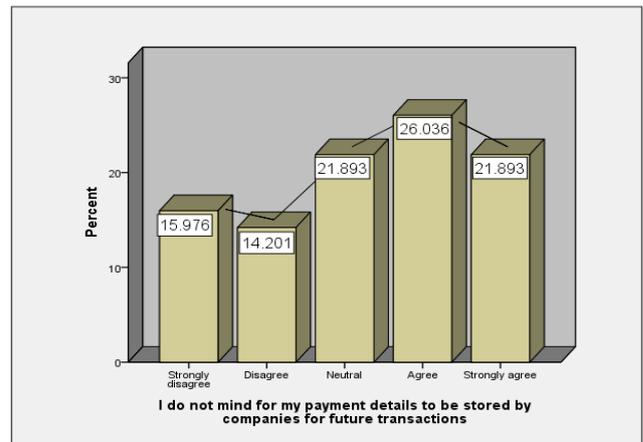


Figure 7: SP- statement 7

Fig. 7 provides the outcome of the statement that individuals are not disturbed by the fact that companies may store their details for future transactions. From the results it is seen that 21.8% were in strong agreement that they were happy for their payment details to be stored by the company for future transactions, and a further 26% agreed. However, 15.9% disagreed strongly, and 14.2% disagreed, thereby showing that whilst 47.8% do not mind that their payment details might be stored for future transactions, almost one third (30.1%) do find this an unacceptable situation. Moreover, almost one quarter of the sample (21.8%) gave a neutral answer, signifying not only that the usual small group in that category had never had to think

about this, but also that some people who did have opinions about other issues had not actually made up their minds about the benefits and/or disadvantages of this practice.

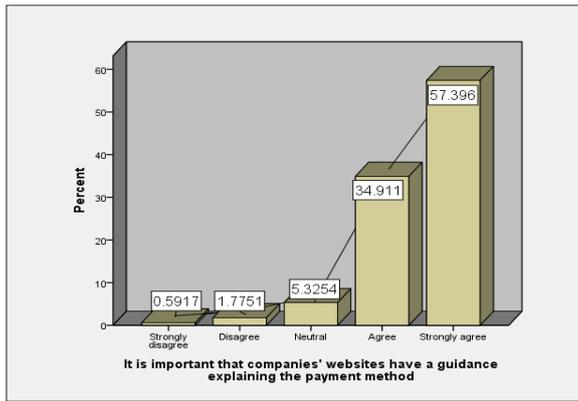


Figure 8: SP- statement 8

Fig. 8 concerns the statement that people believe it is important for companies' websites to include guidance about the payment method, and in respect of this statement, there was good agreement among respondents that this was good practice and should be the case. Specifically, 57.3% agreed strongly, 34.9% agreed, and only 0.5% disagreed strongly, and a further 1.7% disagreed. Therefore, a large majority of 92.2% are of the belief that it is important for companies to provide comprehensive instructions regarding the payment method. Clearly, apprehension and uncertainty about how to pay is a major cause of reluctance to make online purchases. Again a small number of respondents (5.3%) did not offer an opinion and simply recorded a neutral answer.

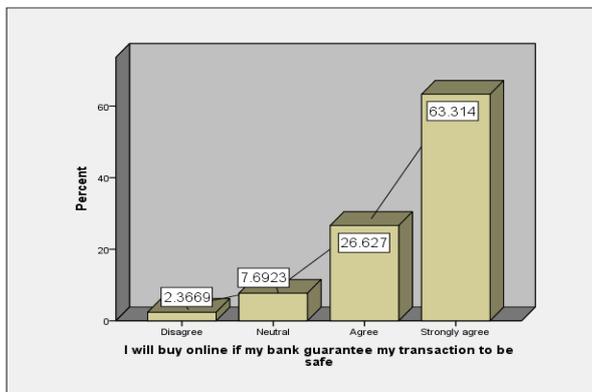


Figure 9: SP- statement 9

Fig. 9 reveals whether people will buy online if their bank guarantees their transactions to be safe. In this respect, 63.3% agreed strongly that they would do so in this situation, and a further 26.6% agreed. Only a very small percentage (2.3%) disagreed. Consequently, it can be seen that the participation of the banks is a major factor in allaying fears among the online buying community, since a majority of 89.9% of the sample signaled their willingness to buy online if they are supported in this way by their bank, and only 2.3% indicated that this support

would not be enough for them. A small percentage (7.6%) were not sure either way and gave a neutral answer.

## V. DISCUSSION

The literature has outlined that E-Commerce usage, security and payment arrangements represent the dominant issues considered by Saudis when they are deciding whether or not to engage in E-Commerce, and that online sellers must direct their attention towards satisfying the concerns potential buyers have in this direction. The results from this study are similar to, and support the study conducted by [2, 9], that found information privacy to feature as one of the main obstacles to the development of E-Commerce.

The development of trust depends on the levels of trust already existing, and in the case of the Saudi population, these are low. According to our finding, 86.4% of customers strongly agreed that companies must have a secure payment system supporting the results reported in [10], as they indicated their anxiety about 'stolen credit cards numbers' and companies' failure to provide 'a secure payment platform'.

Clearly, these worries about payment over the Internet must be addressed and dissolved if the integrity of the entire system is not to be challenged. Consequently, online vendors must devise solutions and implement certain strategies to avoid security breaches and fraud [5], and they must at the same time, publicise these to their potential buyers as an encouragement and to foster the belief that their payment over the Internet is secure, and that buyers are not being jeopardised in any way. The poor security of the online environment in respect of personal information being illegally accessed is a major challenge to E-Commerce [6], and the optimal use of security measures is essential to create the trustworthiness needed to assist the growth of E-Commerce [2]. In the Saudi context, several issues are currently causing the slow development of E-Commerce, and the lack of effective payment systems, and the difficulties encountered in transacting with banks are among these [14]. These shortcomings bolster the lack of trust which already exists in Saudi Arabia, and which represents the greatest deterrent to customers' confidence in Internet payment systems. As a result there should be access to secure online payment gateways to guarantee safety when paying for goods and services online [11]. This is particularly important where Web merchants do not have real-world equivalents of their Web stores (e.g., amazon) since there is no possibility to visit in person in the event of a problem. According to AlGhamdi, [12], access to secure online payment platforms is the best way to elevate the protection of online payments and this is also confirmed by this study. Therefore, Security and Payment are highly influential factors in the encouragement of Saudis living in the UK, to become participants in E-Commerce.

## VI. CONCLUSION

This study was undertaken to explore the environmental influences brought to bear upon Saudi citizens when they contemplate engaging in E-Commerce in the UK. Specifically, it aimed to investigate their perspectives on issues of security and payment in an Advanced E-Commerce Environment.

From the findings presented in the previous section, and from the knowledge of how Saudis residing in Saudi Arabia behave in this respect, it is seen that in general, there is a difference in perception and degrees of willingness to engage in online buying. Saudis in the UK are more open and/or have greater trust in online stores than their counterparts in the Saudi Arabia. They are more willing to provide their payment details when shopping online than individuals in Saudi Arabia are, and this is an important difference which is worth noting, since this could be the key to supporting online businesses in Saudi Arabia.

TABLE 1. Summary of the Results

*Item	SA%	A %	N %	D %	SD %
SP-1	3.6	29.0	37.3	13.6	16.6
SP-2	11.8	49.7	23.7	10.1	4.7
SP-3	33.1	49.1	14.2	1.8	1.8
SP-4	86.4	10.1	3.0	-	.6
SP-5	82.2	13.0	3.6	.6	.6
SP-6	79.3	14.8	4.1	1.2	.6
SP-7	21.9	26.0	21.9	14.2	16.0
SP-8	57.4	34.9	5.3	1.8	.6
SP-9	63.3	26.6	7.7	2.4	-

\*Strongly Agree (SA), Agree (A), Neutral (N), Disagree (D) and Strongly Disagree (SD).

## VII. REFERENCES

- [1] A. Mostafaiepour, "Contagion Aspects of Implementing E-commerce: A Case Study of B2C". *International Journal of Computer Applications*, 29(10):32-40, 2011.
- [2] F. Belanger, J.S. Hiller and W. J. Smith, "Trustworthiness in Electronic Commerce: the Role of Privacy, Security, and Site Attributes", *The Journal of Strategic Information Systems*, 11(3):245-270, 2002.
- [3] H. Alshehri and F. Meziane, "An Investigation into Saudi Online Shoppers' Behaviour Abroad," *Sixth International Conference on Developments in eSystems Engineering (DeSe 2013)*, Abu Dhabi, pp. 323-327, 2013.
- [4] H. Alshehri, and F. Meziane, "Current State of Internet Growth and Usage in Saudi Arabia and Its Ability to Support E-Commerce Development", *Journal of Advanced Management Science*, 5(2):127-132, March 2017.
- [5] J. K. Shim, A. A. Qureshi, J. G. Siegel and R. M. Siegel, "The International Handbook of Electronic Commerce", Routledge, 2013.
- [6] M. Alini, "Identifying application barriers of electronic commerce regarding agricultural products in Iran using the Delphi method", *WALIA journal*, 30(1):289-295, 2014.
- [7] M. Sobihah, M. Mohamad, N. A. M. Ali, & W. Z. W. Ismail, "E-Commerce Service Quality on Customer Satisfaction, Belief and Loyalty: A Proposal", *Mediterranean Journal of Social Sciences*, 6(2):260-266, 2015.
- [8] N. Kamalabadi, A. Bayat, P. Ahmadi, & A. Ebrahimi, "Identifying and Prioritization of Challenges and Barriers of EC Implementation in Iran", *World Applied Sciences Journal*, 5(5):590-597, 2008.
- [9] N. K. Malhotra, S. S. Kim, & J. Agarwal, "Internet users' information privacy concerns (UIPC): The construct, the scale, and a causal model". *Information Systems Research*, 15(4):336-355, 2004.
- [10] R. AlGhamdi, S. Drew and O. AlFaraj, "Issues Influencing Saudi Customers' Decision to Purchase from Online Retailers in the KSA: A Qualitative Analysis" *European Journal of Scientific Research*, 55(4):580-593, 2011 (a).
- [11] R. AlGhamdi, S. Drew, & W. Al-Ghaith, "Factors influencing E-Commerce adoption by retailers in Saudi Arabia: a qualitative analysis", *The Electronic Journal on Information Systems in Developing Countries*, 47(7):1-23, 2011(b)
- [12] R. AlGhamdi, S. J. Nguyen, A. Nguyen and S. Drew, "Factors influencing e-commerce adoption by retailers in Saudi Arabia: a qualitative analysis", *International Journal of Electronic Commerce Studies*, 3(1):83-100, 2012 (c).
- [13] S. A. A. Rajon, Abdullah-Al-Nahid, and A. S. M. Arif, "A Generic Framework for Implementing Electronic Commerce in Developing Countries", Vol. 01, pp. 42-53, 2011.
- [14] SOCG (Sacha Orloff Consulting Group). "E-Commerce in Saudi Arabia: Driving the evolution, adaption and growth of e-commerce in the retail industry", 2012, [online] <http://sacha-orloff-group.com>, last accessed in 12 March 2017
- [15] Y. Zhu, "Research on the Data and Transaction Security of Enterprise E-Commerce Countermeasure", *International Journal of Security and Its Applications*, 7(6):259-268, 2013.

## **Session 13: Infonomics and e-Technology**

Title: Effects of Lexicon Size on Solving Bananagrams  
(Authors: Paul Helling, Ahana Roy, Aspen Olmsted)

Title: Work in Progress: Nobody Knows Who You Really Are Online/On the World Wide Web – Amazon-Recognition-Service based 'Virtualnet' Supported Authentication: How certain are you it is indeed you in front of a screen?  
(Authors: Anderson Holguin Avila, Brigitte Rodríguez Mendoza, Maria Bohorquez Sotelo, Delgadillo Loaiza Juan Sebastián)

Title: Considerations for OSV over Linux-based Virtual Machines  
(Authors: Wayne Chen, Aspen Olmsted)

Title: Using Analysis of Temporal Variances within a Honeypot Dataset to better predict Attack Type Probability  
(Authors: Seamus Dowling, Michael Schukat, Hugh Melvin)

# Effects of Lexicon Size on Solving Bananagrams

Paul Helling, Ahana Roy, Aspen Olmsted  
 Department of Computer Science  
 College of Charleston  
 Charleston, SC USA

**Abstract**— In this paper, we investigate the problem of playing Bananagrams with a computer player. Bananagrams is a fast-paced board game where players compete to place and rearrange their letter tiles as quickly as possible. Despite being superficially similar Scrabble, Bananagrams is not nearly as well studied. We seek to address this knowledge gap by designing a simple program that plays Bananagrams by itself. In this paper, we investigate the effects of lexicon size on the ability of the program to play Bananagrams. We found that using a larger lexicon was slower but increased the win rate.

*Keywords*-Bananagrams; algorithm performance

## I. INTRODUCTION

Bananagrams®, a registered trademark of Bananagrams and hereafter referred to as Bananagrams, is a relatively young board game where players compete to arrange all their letter tiles into connecting and intersecting words. It is estimated to have sold more than 5.5 million copies since its international debut in 2006 [1]. Despite being superficially similar to Scrabble™, a Hasbro trademark hereafter called Scrabble, Bananagrams is a distinct in a few ways. First, Bananagrams is a game of speed. Once a player has played all their tiles, all players must draw a tile from the shared pile until there are none left. The first player to play all their tiles wins. Second, Bananagrams players have personal playing spaces and can rearrange their tiles at will. Whereas Scrabble players are incentivized to play valuable words, occupy strategic bonus point squares, and play defensively, Bananagrams players just need to place their tiles as quickly as possible. This makes solving Bananagrams a different problem from solving Scrabble.

In this work, we will measure the performance of a Bananagram solver that uses lexicons of three difference sizes. The performance will be measured in terms of both win rate and total play time, which will serve as a proxy for the time cost of rearranging the tiles after they are played.

The organization of the paper is as follows. Section II describes the related work and the limitations of current methods. In Section III we give a motivating example where our algorithm is useful. In Section IV we describe the implementation of the algorithm. In Section V we show the preliminary data. In Section VI we review our conclusions and opportunities for future work. In section VII we provide references to the works cited.

## II. RELATED WORKS

Appel and Jacobson wrote what was at the time the world's fastest Scrabble program. It worked by storing the lexicon in a

Directed Acyclic Word Graph (DAWG). This was itself an improvement over the trie because the DAWG represented the lexicon as a graph instead of a tree, leading to significant savings in space. The authors identified four ways in which their program could be improved: a two-way DAWG, looking at fewer left parts, an adversary search, and heuristics [2].

Gordon uses the DAWG for the basis for GADDAG (not a known acronym). Whereas a DAWG can only work from right to left, a GADDAG words from right to left and left to right, which makes it twice as fast but the data structure five times larger. Given the advances in computing since Appel and Jacobson first implemented the DAWG, this is an acceptable trade-off. Gordon didn't test the GADDAG on a full lexicon with large words, but a faster algorithm makes it easier to design tournament-ready Scrabble AIs and heuristics [3].

Chris Piech wrote a Bananagrams solver as a part of a handout for a computer science course at Stanford University. His algorithm uses heuristics to determine the best possible way of placing the first 21 tiles drawn in a game of Bananagrams. However, this program does not play through a whole game [4].

## III. MOTIVATING EXAMPLE

Maven, one of the most competitive Scrabble programs, benefits from having a large corpus of words to choose from. Specifically, Maven has a dictionary with nine letter words compared to the traditional seven of most Scrabble dictionaries [5]. The creator of Maven originally used their own data structure to store the lexicon, but ended using the DAWG in future versions because it was simpler [5]. In comparison, there is little research on Bananagrams, and as of 2016, there is only one freely available Bananagrams AI.

## IV. IMPLEMENTATION

In order to have full control of the testing parameters, we chose to write our own Bananagrams solver. The solver attempts to place the 21 tiles it starts with on the board. We chose to model 21 tiles because this is the number of tiles that each player has in a game with 2-4 people. This is a reasonable constraint because players don't draw additional tiles from the pile until at least one player plays all their initial tiles and is consistent with the number of tiles used in current work in this area [4].

The program works in cycles during which it attempts to place a word on the board. Over the course of a cycle, the program attempts to place a word on the board and compares the previous board state to the current state. If they are different, the program continues to run. If they are the same, all tiles returned to the hand and the program tries playing again. The

Algorithm 1 findWord()

**INPUT:** hand, board, wordList

**OUTPUT:** longest

```

longest = ""
boardState = False
for y in range(0, len(board)):
    for x in range(0, len(board)):
        if board[y][x] != " ":
            tile = board[y][x]
            words = list of all possible words with the hand the
                variable tile
            longwords = dictionary with word value as key and
                the word as the value, derived from words and
                wordList
            length = 0
            for i in longWords:
                if i > length:
                    length = i
            longest = random word from the values with the key
                of longwords equal to length
            attachments = []
            for i in range(0, length):
                if longest[i] == tile:
                    append i to attachments
            for i in attachments:
                board2 = copyBoard(board)
                sub1 = longest[:i]
                sub2 = longest[i+1:]
                canCheckBoard = True
                for j in range(length):
                    difference = j - i
                    if board2[y+difference][x] == "":
                        board2[y+difference][x] = longest[j]
                        elif i == j:
                            pass
                        else:
                            canCheckBoard = False
                            break
                if canCheckBoard:
                    boardState = checkBoard(board2, dictionary)
                    if boardState:
                        return longest
                else:
                    canCheckBoard = True
                    board2 = copyBoard(board)
                    for j in range(length):
                        difference = j - i
                        if board2[y][x+difference] == "":
                            board2[y][x+difference] = longest[j]
                            elif i == j:
                                pass
                                else:
                                    canCheckBoard = False
                                    break
                    if canCheckBoard:
                        boardState = checkBoard(board2, dictionary)
                    if boardState:
                        return longest
longest = ""
return longest

```

simulation ends when there are no tiles left in the hand, which is a win, or it's played through 500 cycles, which is a loss.

To find and place words, the program plays the first, best word it can find. If the board is empty, it plays the highest scoring word it can make in the middle of the board. Otherwise, the program checks each tile on the board and tries to make words using that tile and the tiles in the hand. After making a

TABLE 1. AVERAGE NUMBER OF TILES

	Leftovers	Time (s)	Wins	Time-to-Win (s)
OPTED	7.81	2831	39	684
Boulter	4.82	3320	62	1049
All Words	3.8	3156	70	1064

list of words and sorting by score, the program randomly picks the highest scoring word and attempts to place it on the board. If no words are played after checking all the tiles, the simulation ends. The algorithm is shown in Algorithm 1.

Word scoring is calculated using the Scrabble values of letters. Using word length resulted in a program that would never play rare letters like X and Q. Using letter frequency resulted in words being too finely separated, so the program would create the same board with a given hand every time. Because Scrabble rewards both long words and rare letters, this results in several words having the same score, in which case the program picks from them randomly. This creates choice and allows the program to explore several paths in the treespace.

## V. PRELIMINARY DATA

To measure the effects of lexicon size, we needed different wordlists. We found two freely available lexicons on the internet, The Online Plain Text English Dictionary (OPTED) with 105,715 words after filtering and Jeff Boulter's Scrabble Dictionary with 178,691 words. We obtained a third list with all unique words, which is the union of the OPTED and Boulter's dictionary with 223,847 words.

To test each lexicon, we ran the program 100 times with a randomly generated hand each time. At each iteration, we recorded time elapsed and the number of tiles of leftover. Since having zero tiles leftover is a winning position, we were then able to calculate useful statistics such as win rate, average time, and average time to win.

As shown in TABLE 1, using the combined lexicon was the most effective in increasing the win rate of the program, with the medium length lexicon in the middle. Even though the longer lexicons have higher win rates, they take longer to run because the program has to check more words in each turn.

## VI. CONCLUSION

In conclusion, Bananagrams simulations with longer lexicons were more likely to be able to play all 21 initial tiles than shorter ones. This had a trade off with time because they took longer to run than the shortest lexicon.

There are several possible ways to improve this research. A logical next step would be to create randomized subsets of the word lists of different lengths and then run the program with them to evaluate their performance. Another way would be to improve tile rearrangement or word selection. Lastly, an algorithm that does a depth-first search through the tree space could do a better a job of revealing if a set of 21 tiles can be played than waiting for the program to timeout. Additionally, adopting a novel data structure such as the DAWG or GADDAG could also result in time savings.

REFERENCES

- [1] B. Campbell, "The Addictive Appeal Of Bananagrams: NPR," National Public Radio, Inc., 18 Augst 2011. [Online]. Available: <https://www.npr.org/2011/08/18/139725678/the-addictive-appeal-of-bananagrams>. [Accessed 20 April 2017].
- [2] A. W. Appel and G. J. Jacobson, "The World's Fastest Scrabble Program," *Communications of the ACM*, vol. 31, no. 5, pp. 572-578, 1988.
- [3] S. A. Gordon, "A Faster Scrabble Move Generation Algorithm," *Software: Practices and Experience*, vol. 24, no. 2, pp. 219-232, 1994.
- [4] C. Piech, "CS106B," Stanford University, 13 January 2016. [Online]. Available: <https://web.stanford.edu/class/archive/cs/cs106b/cs106b.1164/handouts/bananagrams.html>. [Accessed 6 February 2017].
- [5] B. Sheppard, "World-championship-caliber Scrabble," *Artificial Intelligence*, vol. 134, no. 1, pp. 241-275, 2002.

# Work in Progress: Nobody Knows Who You Really Are Online/On the World Wide Web – Amazon-Recognition-Service based ‘Virtualnet’ Supported Authentication

How certain are you it is indeed you in front of a screen?

Anderson Holguin Avila, Brigitte Rodríguez Mendoza, Maria Bohorquez Sotelo, Delgadillo Loaiza Juan Sebastián  
Investigación y tecnología

Universidad Manuela Beltran, UMB  
Cajicá, Colombia

Andersonholguin@umb.edu.co, Brigitte.rodriguez@umb.edu.co, Maria.bohorquez@umb.edu.co,  
Juansebastian.delgadillo@umb.edu.co

**Abstract**— a specific project on its development phase is recounted on this paper which looks to strengthen authenticity, verification and validation of student identity levels that perform learning processes, mediated by ‘Virtualnet’ – an e-learning platform from Universidad Manuela Beltrán. To achieve this, a special face-recognition algorithm inclusion is suggested which in first instance is designed and tested to be later implemented as a prototype in the testing (exams) module from the online courses offered on this LMS.

**Keywords-component;** Facial recognition, security e-learning, authentication, biometrics.

## INTRODUCTION

The implementation of new online tools for education has brought great learning possibilities and opportunities for universities and students alike, in a way that they have accomplished leaving rigid distance and schedule barriers behind; nevertheless, with each new implementation, come greater challenges to face, among these, safety data methods that online platforms should provide to guarantee, within what the latest technology allows, the correct user identification to avoid fraudulent alteration of identity at best.

Thanks to new worldwide, artificial intelligence developments, an advance in people identification methods has been achieved, mostly through digital images, known as *Facial Recognition*, where possibilities of using this technique as a reliable tool in the process of user identification, has opened great possibilities.

Furthermore, the intention of the present investigation in process is to use the new advantages of face recognition technology, that are offered as a service to create tools with greater reliability in user recognition – process that can aid Universidad Manuela Beltrán’s and all its campuses’ Learning Management System’ in Colombia.

## I. BACKGROUND

In the new academic learning modalities, as established in [1], the Virtual Formation Systems are acquiring more

prominence in teaching, hence, Universidad Manuela Beltrán (UMB) surfaces in the cyberspace with the Learning Management System (LMS), as a necessity to inhabit the information and knowledge society with ‘VirtualNet’, a academically-taylored platform that suggests following up on the relevance of the virtualization processes of the University. In this context, VirtualNet suggests activating basic user protection mechanisms, performing profile authentication with 2 tokens: One private and one public. The public token references the identification number which cannot be changed, although it is a unique assignation per user on the University database; the private token relates a pre-determined combination of characters that even if it does not have a system-administrator based requirement to force users to create safe passwords, it does recommend the following:

- Minimal character length.
- Combination of letters, numbers, and special characters.
- Upper/lowercase characters.

Within these characteristics, the platform demands a minimum security level, seeking to avoid certain database forceful attacks. Once the password is entered in the system, a one-way encryption process is carried out as exhibited in [2], using the SHA algorithm to proceed and store the aforementioned in the database, which allows to state the hypothesis that the only person whom will know the content of the password is indeed the user. This cryptography standard takes the algorithm reduction focus or message summary, where [3] confirms that by the usage of bits to amplify its encryption information capacity, it will manage to generate in coding output between 40 to 60 entry bits (input) to 1024 exit bits (output).

Similarly, taking into account the safety and authentication characteristics, VirtualNet uses a password change protocol that forces the user to modify their private token, trying to protect the information, leaving the access responsibility to the user.

Finally, under the concept model of VirtualNet, it is established that the aforementioned described aspects evidence

a unified diagram which defines the main relationships related to e-learning standards and the best practices for the UMB educational focus, but, from its architecture and security focus there are vulnerabilities that might allow false-identity alterations when presenting platform-based activities, where, even if they are saved in the system, the logs with the log-in records that specify IP addresses with which an activity has been made, there is no certainty that the user in the platform is who really has entered the system.

After having made the platform security protocol description and, acknowledging the central role that users have in new technologic resource design and implementation, a survey was held with the participation of 22 University teachers, and 500 University students enrolled in on-site class mode that take virtual optional courses.

The survey inquires about the perception of safety that working on the platform generates and about the importance users give to these aspects, from obtained results it can be stated that in spite of the identified problems on a technologic level, people that interact within the platform feel safe and do not perceive any threat or identity-theft risk that could affect their academic activities.

78.8% of polled people have a favorable concept of security in the platform and 89.1% feel that their information is safe and consider being cyber-attacks victims a remote possibility. It is important to highlight that for the 84.8% polled, the system vulnerability could affect their performance.

Additionally, questions are made to identify the users' preferences and opposability from the users to face recognition technology and systems and it was found that even though it is not the preferred method as can be observed in fig. 1, 87 from the 522 people that answered the poll selected the aforementioned as the system with the most acceptance. When asked about if they would agree with face recognition as an identity validation method on virtual exams on the academic UMB platform to avoid identity theft, 72,2% of polled people express agreement (Agree or Strongly Agree), whilst people with reluctance to its implementation were represented with 20.9% of people polled (Fig. 2). Results that clearly support the decision of designing and implementing this authentication system on VirtualNet.

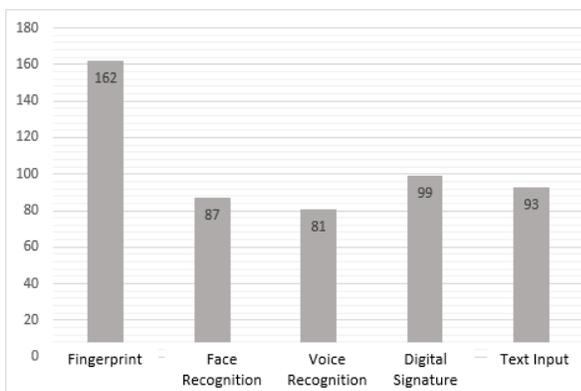


Figure 1. Authentication system preference

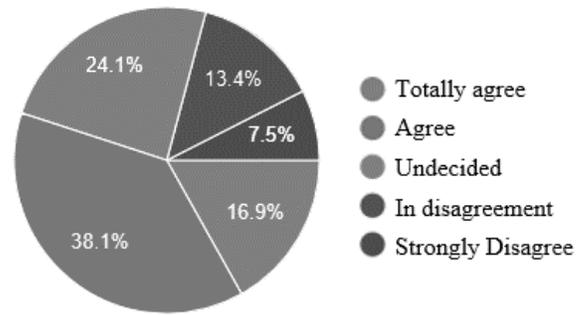


Figure 2. Face recognition acceptance as a validation method

## II. PROJECT FRAMEWORK

### A. Project and objectives Presentation

E-learning is being positioned worldwide and locally as a mode of teaching and learning, UMB as a response, not only for its formative processes on a disciplinary level, but given its commitment with a quality education for everyone, draws out virtualization rules for on-site class courses in response to the social demands of current times. Thus, the University offers extensive cross-curricular courses for students enrolled in on-site class, full-time mode, which brings the answer to the following question: How can one assure that students obtaining their certification in any course are the ones that have taken them, given that there does not exist on-site class contact? This situation evidences the vulnerability of the system to guarantee the trustworthiness in the students' authentication.

As a consequence, the project's objective is to design an algorithm to strengthen the authentication mechanisms on VirtualNet users, with the purpose of diminishing the error margin in certificating students that use the platform.

### B. Project Description

Within the nature of this work, an algorithm is designed to strengthen the authentication mechanisms on VirtualNet users; statement that in the future will be oriented towards the implementation of face recognition services that allow to validate access and presence of the students to this LMS.

This project's approach seeks to offer validation processes in user authentication to favor security aspects, for which it will be based on [4] to propose a system structuring which is more solid and that can lessen identity-theft attacks; usability, with a quality systematic focused on the user; and legally, taking into account the current legislation on cyber safety globally because of the possibilities that the cyber space offers that allows to consider ethic and legal privacy issues, data handling and integrity and respect for people that are involved in teaching and learning on virtual platforms.

Attending to the usability of virtual learning environments and in consideration of the student design-centered, results from the surveys are taken and the required times for tests and implementation of the algorithm to select the face recognition as an authentication system to validate the users' identity on VirtualNet.

The authentication processes will be implemented randomly and during the users' time of presence in the platform. The login register through user and password is kept through the said user information therefore adding the face recognition process after.

### C. Project Phases

The Project is divided in 2 phases, the first is oriented to the algorithm's initial design, where the following stages are included:

1. Current system revision in informatics safety.
2. Environment analysis and users' needs.
3. Methodological questioning on informatics security aspects on LMS.
4. Triangulation of the methodological aspects and environment analysis.
5. Establishing solution alternatives to implement methods that allow to enhance the level of trustworthiness of the face recognition implementation as a tool of identity verification.
6. Structural algorithm design for the implementation.

Subsequently, the second phase will take place, where the direct implementation of the algorithm on the LMS VirtualNet will be carried out; on this phase the initial prototype will be provided under the following environment:

1. LMS evaluation module application.
2. A specific academic UMB program will be run, to evaluate an initial sampling with the adequate characteristics for information compiling and evaluating.

### D. Environment Description

On the long run, the management learning system LMS VirtualNet, in its second phase, will count with 2 verification tools that will look to improve the user authentication on the platform, with the purpose of increasing the reliability on official certification either from government entities or students.

### E. Course Description and how are exams like?

The Evaluation/Exam module which will be the pilot module for this implementation that allows an important part of the University's development of the courses offered by the UMB on its online, non-on-site class mode, will offer identity verification tools, improving platform security in terms of identity-theft, while at the same time allowing the University to approach the certainty of ensuring the process of academic studies certification.

### F. Resources

Process summary information statistics were distributed through the same communication modules from VirtualNet, e-mails and other information sources. The strategy planning was made on the documents offered by the facial recognition services from Amazon Web Services and the previous safety study analysis made to VirtualNet.

## III. PRELIMINARY RESULTS

As results of the methodologic inquiry on LMS security information aspects, the following possible vulnerabilities under the identity-theft mode were found:

- The identity robber can use a static image of the account owner and locate it in front of the devices camera.
- The user can be in front of the device's camera, which would give a degree of approved similarity, but could be accompanied by a third party that would be answering the platform activities.
- The user can take unfamiliar objects over their face to make recognition difficult or impossible.
- Based on the aforementioned, it is determined that for this first version of the algorithm's initial design the following characteristics shall be taken into account:
- It is necessary to make different takes of the user on different random moments to obtain the best possible amount of data per try.
- The facial recognition system must show a degree of similarity between the original document's photograph and the ones capture don different times during the exam.
- The system has to offer a face recognition that is in the capacity of identifying the quantity of people on different moments throughout the exam.
- The face recognition service has to supply information on the users' facial patterns, with the intention of evaluating the different moments in session and being able to dismiss the likelihood of the usage of static photography on the camera.
- The compiled data has to generate silent alerts which will be stored for their analysis and later verification on its usefulness and truthfulness for this project.

As a result of this investigation it has been decided to implement the following processes to improve the possibilities of identifying the user that is active to avoid possible frauds. To demarcate this first phase, the specific use on exams proposal will be made on the platform, taking into account:

- 1) The system has to take the designated time for the exam and calculate the average time for it in similar evaluations. Having this timeframe, 5 to 10 moments must be taken at random for the samples to be taken on each exam for each user. This to allow to create unique simple intervals for each attempt within the exam.
- 2) Each one of these samples (images) are analyzed with the face recognition service from **Amazon Web Service (AWS)**, obtaining as an answer a series of parameters and prospects to be analyzed, taking into account the degree of similarity between comparisons of current photography taken during the exam and pictures from the students identification documents, amount of people on the pictures taken randomly, unidentified objects in the photographs and facial treats for the person taking the exam. These parameters

have to be stored to be able to have proof of the results after the analysis is made.

- 3) With the analyzed data per try, alarms of possible attempt of identity-theft will be created based on the following statements:
  - I. The actual samples have a similarity under 80% in relation to the identity document.
  - II. The user of the platform has unidentified objects on their face.
  - III. There is more than one person on the scene in more than two takes
  - IV. There are two samples that are completely the same in facial expressions which would indicate that there are static photographs in use to avoid the authentication measure.

Taking the proposed analysis into consideration, the initial design of the algorithm is done to strengthen the user authentication mechanisms on VirtualNet (Fig. 3)

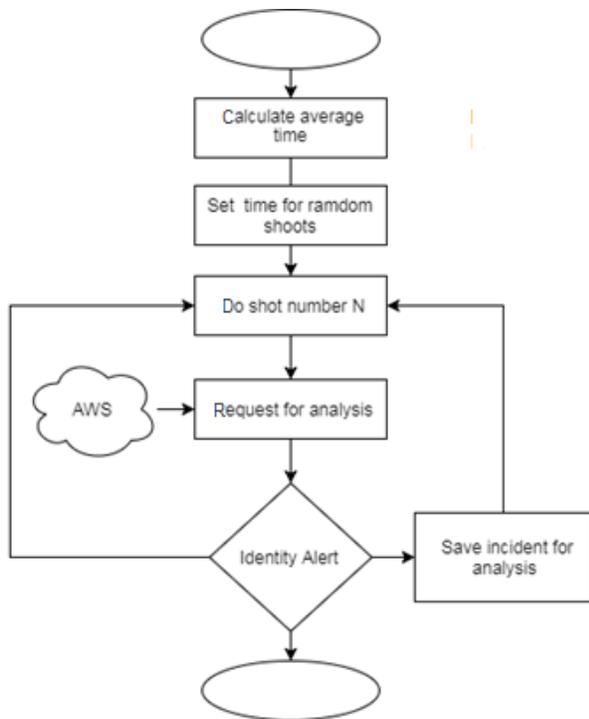


Figure 3. General flow diagram.

#### IV. ONGOING WORK / FUTURE WORK

As work in progress, on the investigation team we are implementing on the VirtualNet development environments the first version of the algorithm herein described. This algorithm will have a main controller which will be in charge of receiving and processing all requests from and to the Amazon Web Service, where the main controller has to be addressed from the main module's class which will be the pilot module.

As first implementation stage, the functionality of face recognition will be applied on the exams from the course 'Fundamentos de Investigación' which currently counts with

400 students, to be able to compile and analyze the possible cases of identity theft on exam presentation.

#### CONCLUSIONS

The work will continue until 2018's second semester, when its complete implementation is fulfilled, therefore only partial results are presented on this work, specifically on the exposed proposal on the initial phase.

It is found that VirtualNet users do not notice safety issues nor feel vulnerable to cyber-attacks, in spite of the limited levels of safety that the platform counts on. Under this fact, it is mandatory to maintain this perception of usefulness strengthening the safety system in place.

The platform users acknowledge and consider some biometric authentication systems as reliable, the one with the most acceptance is the one that is found most frequently on a local context, in this case the fingerprint method.

The majority of VirtualNet users do not consider invasive nor reject face recognition usage as a resource of authentication to work on virtual courses offered by the UMB.

To guarantee the accreditation and certification of the students that are part of the learning community mediated by the LMS, it is necessary to use algorithms and artificial intelligence resources of pattern recognition.

#### REFERENCES

- [1] J. Díaz, A. Schiavoni, A. Osorio, P. Amadeo, y E. Charnelli, "Integración de plataformas virtuales de aprendizaje, redes sociales y sistemas académicos basados en Software Libre. Una experiencia en la Facultad de Informática de la UNLP", Universidad Nacional de la Plata, 2012.
- [2] S. Bayat, M. Mozaffari and A. Reyhani-Masoleh "Efficient and Concurrent Reliable Realization of the Secure Cryptographic SHA-3 Algorithm", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 33, n.º 7, pp. 1105-1109, jul. 2014. , Member, IEEE
- [3] Y. Mena, "Algoritmos HASH y vulnerabilidad a ataques", Journal Información, Tecnología y Sociedad, p. 108, 2009.
- [4] A. Morales, J. Fierrez, R.Vera-Rodriguez, & J. Orteg "Autenticación Web de Estudiantes Mediante Reconocimiento Biométrico". In III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad. 2016.

# Considerations for OS<sup>V</sup> over Linux-based Virtual Machines

Wayne Chen, Aspen Olmsted  
 Department of Cybersecurity  
 New York University  
 wc1359@nyu.edu, olmsteda@cofc.edu

**Abstract**— The focus of this research work is to explore the operational considerations of using a lightweight operating system (OS<sup>V</sup>) over standard Linux-based operating systems for virtualization. We focus on three aspects: a review of published literature in regards to performance; a comparison of virtualization footprint for a common application (Apache Tomcat), and an analysis of the possible security advantages. The results provide a rationale as to when lightweight operating systems may be advantageous for virtualization.

**Keywords**- unikernal; lightweight operating system; virtualization; OS<sup>V</sup>; Apache Tomcat

## I. INTRODUCTION

Cloud computing, where computing resources are available at a remote site via high-speed network and consumers pay for resources as-needed, is built upon the enabling technology of operating system virtualization. The vast majority of virtualization solutions in the cloud employ hypervisors such as Xen or VMWare, and Virtual Machine (VM) operating systems are typically conventional, well-known operating systems such as Windows or Linux since most applications run natively on these systems. These virtualized, full operating systems contain a kernel hosting both primary applications as well as secondary services, allowing the virtual machine to host a number of end-user applications if so desired. This remains a key advantage of using general purpose operating systems.

The drawbacks of these virtualized operating systems stem primarily from the fact that these operating systems were not designed with virtualization in mind. An example is the external load balancers that are often necessary such that VMs can be expanded and contracted elastically based on need. Furthermore, in recent years it has become so easy and simple to deploy virtual machines that now, commonly, VMs have become single purpose, hosting a single application. Despite this trend, virtualization still overwhelmingly utilizes conventional operating systems.

There have been recent efforts to improve upon conventional architectures [1]. One area of improvement is the revisiting of the concept of the unikernel. The idea of light weight virtualization for custom applications is not new. In 2002, a group of researchers presented Mate [2], a tiny communications-based virtual machine designed specifically for sensor networks where the nodes are tens of thousands of tiny devices. Mate was written on TinyOS, a minimalist operating system because sensor network nodes (at the time)

had extremely limited hardware resources. Several years later, in 2006, a group at Sun Microsystems, is working on a similar problem of wireless sensor devices, presented Squawk [3]. Squawk was a small Java virtual machine designed to run without an operating system altogether. Squawk itself provided OS level mechanisms for these small devices. Indeed, footprinting of Squawk showed it to be the hundreds of kilobytes in size, appropriate for the intended devices.

Unikernels are specialized operating system kernels written in high-level languages for the specific purpose of virtualized appliances. Many unikernel implementations have surfaced in recent years, including Mirage, Drawbridge, HalVM, ErlangOnXen, and OS<sup>V</sup> [1]. OS<sup>V</sup> is of particular interest because it is written in Java with the specific purpose of replacing conventional operating systems for cloud computing virtualization. In this paper, we explore the performance, footprint, and security implications of using OS<sup>V</sup>.

The organization of the paper is as follows. Section II refers to published literature where OS<sup>V</sup> performance is compared against that of Linux. Section III provides a brief comparison of virtual machine footprint from empirical installations performed on a test workstation. Section IV provides analysis on security implications of using OS<sup>V</sup> for virtualization. We conclude and discuss possible future work in Section V.

## II. OS<sup>V</sup> PERFORMANCE

Morabito et al. [4] recently compared the performance of hypervisors versus lightweight virtualization. The lightweight virtualization was primarily container-based, where instead of an operating system, processes at the host OS level are isolated to avoid the overhead of an additional operating system. In addition to their primary comparison, OS<sup>V</sup> over KVM was benchmarked against Linux over KVM.

TABLE I. OS<sup>V</sup> vs Linux over KVM

Performance	OS <sup>V</sup> vs Linux over KVM
CPU	Minimal/negligible differences
Disk I/O	Test not supported by OS <sup>V</sup>
Memory	Memory throughput of OS <sup>V</sup> was roughly 50% compared to other solutions
Network I/O	OS <sup>V</sup> outperformed Linux over KVM slightly in both TCP and UDP throughput, but both underperformed container virtualization by significant margins

The results are simplified to qualitative descriptions and summarized in Table I.

Overall, the results showed OS<sup>V</sup> to have lower memory performance, slightly higher network performance, and similar CPU performance when compared to Linux virtualization.

### III. OS<sup>V</sup> FOOTPRINT MEASUREMENT

A simple empirical comparison was performed on our test system To assess the footprint impact of using OS<sup>V</sup> as a virtualization operating system.. The following hardware was used:

- Computer: Apple MacBook Pro (15 inch, late 2016)
- Processor: 2.6 GHz quad-core Intel Core i7, Turbo Boost up to 3.5 GHz, with 6MB shared L3 cache
- Memory: 16GB of 2133 MHz LPDDR3 onboard
- Disk: 256GB PCIe-based onboard SSD
- OS: macOS Sierra 10.12.5
- Hypervisor: Virtual Box 5.1.22 r115126

The following VMs were installed and configured:

- OS<sup>V</sup> instance with Apache Tomcat 7.0.47. The disk space allocated to this VM was 10GB. The virtual machine begins running Catalina on Apache Tomcat upon VM start
- CentOS 7 minimal edition. The disk space allocated to this VM was 10GB. Upon installation, yum was updated, and Java 1.7 JRE, wget, and Apache Tomcat 7.0.47 were installed. Catalina OPTS settings were configured prior to Apache Tomcat start.

For footprinting, the Activity Monitor application native to macOS Sierra was used to assess resource usage. Each VM was brought up (separately from cold to Apache Tomcat running), after which the following information was captured in Table II, which also shows the downloaded image size of the respective operating systems:

TABLE II. The downloaded image size of the respective operating systems

VM	Image Size	Memory	Threads	Ports
OS <sup>V</sup>	89 MB	300 MB	31	284
CentOS	713 MB	600 MB	36	409

OS<sup>V</sup>, even with Apache Tomcat installed, is a much lighter operating system in terms of disk image space by an order of magnitude (Apache Tomcat and additional downloads were not included in CentOS size). While Section II showed that the memory performance of OS<sup>V</sup> was lower of than that of the Linux Operating System, footprinting showed that overall memory footprint was also lower, by approximately 50%. This is unsurprising given that OS<sup>V</sup>'s focus is only on retaining the basic necessities of a host operating system. Finally, the number of running threads and mach ports for OS<sup>V</sup> were slightly less than a full Linux operating system. The memory performance hit from using OS<sup>V</sup> is balanced by OS<sup>V</sup> needing less memory, a tradeoff to be considered depending on application and user requirements.

### IV. OS<sup>V</sup> SECURITY CONSIDERATIONS

The basic premise of OSV security versus that of a virtualized Linux is that there are likely fewer security considerations because there are fewer services for which vulnerabilities can be found. If multiple applications were to be hosted on a single virtualized Linux machine, this rationale would certainly be valid, as every additional application results in the additional potential for introducing security vulnerabilities. However, as stated earlier, the trend for virtualization has been to deploy single purpose virtual machines based on conventional operating systems, as is the case of our experiment in Section III. Network scanning of both the OSV and CentOS instance showed that the CentOS 7 did not open any unexpected or additional ports when compared with the OSV. This is likely due to the minimal installation of CentOS and improved security practices in recent Linux distributions. An argument can still be made that Linux vulnerabilities are more likely to be exploited than OS<sup>V</sup>, but that argument is based on the rationale that Linux is a more well-known operating system. A search of the Rapid 7 vulnerability and exploit database shows that in recent years, relatively few operating system vulnerabilities have been found in CentOS, and those that are found are typically patched in the next release of the operating system. Therefore, any security advantage of OSV vs. single-purpose Linux virtual machines is likely to be minimal.

### V. CONCLUSIONS AND FUTURE WORK

OSV clearly offers an opportunity to reduce disk space and memory utilization. Published literature and footprinting has shown there is a trade-off between performance and resource utilization of memory to be determined by user need. In terms of security, it can be argued that OSV offers a slightly more secure alternative due to its lack of market penetration and smaller vulnerability cross-section; however, improvements in Linux operating systems have minimized much of the advantage. OSV is still in a fairly early stage of development, and future work in improving performance while maintaining lower resource utilization could make it an even more compelling choice for cloud virtualization. More performance benchmarking and security assessment are necessary.

### REFERENCES

- [1] Madhavapeddy *et al.*, Unkernels: Rise of the Virtual Library Operating System, *Distributed Computing* V11f, 2014.
- [2] Levis *et al.*, Mate: A Tiny Virtual Machine for Sensor Networks, *Proceedings of the 10<sup>th</sup> International Conference on Architecture Support for Programming Languages and Operating Systems*, 2002.
- [3] Simon *et al.*, Java™ on the bare metal of wireless sensor devices: the squawk Java virtual machine, *Proceedings of the 2<sup>nd</sup> international conference on Virtual execution environment*, 78-88, 2006.
- [4] Morabito *et al.*, "Hypervisors vs. Lightweight Virtualization," *Conference: WoC '15: First International Workshop on Container Technologies and Container Clouds. Proceedings*

*of the 2015 IEEE International Conference on  
Cloud Engineering (IC2E 2015). 9-13 March - Tempe,  
Arizona.*

# Using Analysis of Temporal Variances within a Honeypot Dataset to better predict Attack Type Probability

Seamus Dowling

Department of Business, Humanities  
and Technology, GMIT,  
Mayo Campus, Mayo, Ireland  
seamus.dowling@gmit.ie

Michael Schukat

Discipline of IT, College of  
Engineering & Informatics,  
NUI Galway, Galway, Ireland  
michael.schukat@nuigalway.ie

Hugh Melvin

Discipline of IT, College of  
Engineering & Informatics,  
NUI Galway, Galway, Ireland  
hugh.melvin@nuigalway.ie

**Abstract**— Honeypots are deployed to capture cyber attack data for analysis of attacker behavior. This paper analyses a honeypot dataset to establish attack types and corresponding temporal patterns. It calculates the probability of each attack type occurring at a particular time of day and tests these probabilities with a random sample from the honeypot dataset. Attacks can take many forms and can come from different geographical sources. Temporal patterns in attacks are often observed due to the diurnal nature of computer usage and thus attack types captured on a honeypot will also reflect these patterns. We propose that it is possible to determine the probability of differing attack types occurring at certain times of the day. Understanding attack behavior informs the implementation of more robust security measures. The paper also proposes automating this process to create dynamic and adaptive honeypots. An adaptive honeypot that can modify its security levels, can increase the attack vector at different times of the day. This will improve data collection for analysis that ultimately will lead to better cyber defenses.

**Keywords:** honeypot, temporal, predictive, probability, adaptive

## I. INTRODUCTION

A honeypot is an analytical tool first and foremost. It is deceptive and its role is to collect attack information. This information can be analysed retrospectively to determine the modus operandi of attackers. Honeypots can lure an attacker from one honeypot to another, forming a *Honeynet*. A honeynet provides the mechanism to detain an attacker for longer, gaining more insight into attack behaviour.

Capturing cyber attacks provides information on malware development, propagation methods and global activity patterns, for example by analysing source and target IP addresses [1]. A honeypot will capture different attack types depending on the configuration of the attack vector. Temporal patterns will exist for attacks due to the diurnal nature of global computer usage [2]. Identifying attack types and separating these temporal patterns can indicate the probability of different cyber attack activity occurring at particular times of the day. A dynamic honeypot can adopt alternative security levels during the day to capture more relevant attack data, which provides a better insight into cyber attack behaviour. The aim of this paper is to present a method of data analytics

that can be modelled on a honeypot to capture more relevant attack information.

The rest of this paper is arranged as follows:

Section 2 presents previous work in honeypot operations, adaptive techniques and the analysis of captured data demonstrating temporal activity.

Section 3 outlines the methodology used to capture the cyber attack dataset and presents evidence of temporal activity.

Section 4 analyses the dataset, identifies attack types and presents temporal variances for spatially diverse attack sources. It then calculates the probabilities of attack types occurring at hourly timeslots and tests these probabilities with a random sample from the honeypot dataset.

Section 5 discusses automating the probability calculation process to create a dynamic and adaptive honeypot model that reacts to attacks and adopts security levels associated with attack types.

## II. PREVIOUS CONTRIBUTIONS

### A. History

Since their introduction in the 1990s, honeypots have evolved to meet the changing landscape of cyber threats. Recently, Internet of Things (IoT) deployments attract bots and malicious code targeting IoT end devices [3]. In 1992, Bellovin [4] presented work on captured ‘crackers’ activities at the USENIX conference. Dummy machines were deployed to lure and monitor the activity of the attackers. In his seminal book, *Honeypots: Tracking Hackers*, Spitzner defined ‘honeypot’ as being a ‘security resource whose value lies in being probed, attacked or compromised’. Scientific research into honeypots and honeynets has increased since then. Provos [5] in 2003 presented *Honeyd*, an easy to deploy, low risk honeypot. It details how to deploy virtual honeypots with different IPs safely. *Honeyd* acted as a catalyst for the development of further low interaction honeypots. *Nepenthes* [6] and *Argos* [7] became very popular global honeypot tools. As honeypots became more popular, available datasets became larger. Cumulative analysis and modelling can be performed on entire datasets captured on local honeypots or from global projects [8], [9]. The popularity of honeypot deployment raised the question of their role. To address this,

Zhang [10] introduced honeypot taxonomy. This taxonomy identifies *security* as the role or class of a honeypot, and could have *prevention*, *detection*, *reaction* or *research* as values. Deflecting an attacker away from a production network, by luring them into a honeypot, has a prevention value. Unauthorized activity on a honeypot is red flagged immediately providing a detection value. Designing a honeypot to maintain an attacker's interest, by offering choices or tokens [11] has a reaction value. Finally the research value provides a crucial value of a honeypot by understanding the behaviour and motivation of an attacker. As more devices connected and threats evolved on the Internet, Seifert [12] introduced a new taxonomy. Faster networking and virtualisation technologies presented honeypot developers with a means of deploying high interaction honeypots and honeynets with low risk [7], isolating attack traffic from connected hardware and networks.

### B. Honeypot operations

High interaction honeypots provide backend databases to collect all activity such as IP addresses, timestamp, attempts, interactions, commands, downloads and executions [13]. Downloaded files can be sandboxed and analysed [14]. Sandboxing involves the reverse engineering of malware binaries, by allowing their execution in a controlled, isolated environment. This analysis is of particular interest to the antivirus industry. From this sandboxing activity, they can generate updates and fixes to add to dynamic antivirus rollouts. After an attacker has compromised a honeypot, it will attempt to interact in a structured manner. Ramsbrock models this interaction [15]. The initial engagement for an attack, post compromise, is to examine the hardware and software to determine if progression is relevant. On a live production system, this will return the underlying architecture, CPU, uptime, operating system, user privileges and further relevant. Hardware and software properties are checked, to determine if the compromised host has further potential, or if the host is a virtualised environment [16]. An attacker may then modify the host system, including passwords. On a honeypot, engaging the attack sequence at this point prolongs activity. It then attempts to download, install and run malware to complete the compromise.

### C. Botnet

The majority of honeynet attacks take the form of bots [17]. Botnets provide a mechanism for global propagation of cyber attack infection and control. They are defined as large networks of compromised machines used to carry out further attacks [18]. These botnets are under the control of a single command and control (C&C), often referred to as C2. A typical botnet attack will consist of 2 sets of IP addresses. The first set of IPs are the compromised hosts. These are everyday compromised machines that are inadvertently participating in an attack. The diurnal characteristic of end users turning on their machines in the morning, turning them off in the evening can be modelled [2]. The second set of IPs is the C&Cs. These are the sources from which the desired malware is downloaded.

There are 3 methods of communications between C&C and compromised host [19]:

- IRC (Internet Relay Chat) based model using push commands from C&C
- HTTP based model using pull commands from hosts
- P2P based model where bots use peer-to-peer communications

Botmasters obfuscate visibility by changing the C&C connection channel. They use Dynamic DNS (DDNS) for botnet communication, allowing them to shut down a C&C server on discovery and start up a new server for uninterrupted attack service. On a honeypot dataset, tracking botnets is possible by examining methods of communication and establishing patterns. Freiling et al [20] present results that C&C communications with botnets exhibit different characteristics from legitimate traffic. A similar but improved approach is taken in [21], specifically for IRC bot communication. DDNS communication [22] and fuzzy pattern algorithms [23] are used to create cluster patterns, whereby Francois et al [24] uses Google's PageRank algorithm [25] to detect stealthy P2P botnets.

### D. Adaptive Honeypots

There are often legal and ethical issues associated with operating honeypots [26]. In a desire to gather as much information as possible on attacker behaviour, a honeypot could allow the execution of malicious code [27]. The honeypot developer could be liable if their honeypot inadvertently becomes involved in further attacks. Entrapment could be a mitigating factor when it comes to the prosecution of an attacker. This highlights the need for creating honeypots that prolong attacker interaction for an optimum time period, without breaching legal or ethical responsibilities. Using a Markov Decision Process [28], Haytle et al provide a strategy for the optimal response for honeypot operators. The model proposed prolongs attack activity while minimising legal liability. Wagener [29] uses reinforced learning to extract as much information as possible about the intruder. A honeypot called Heliza was developed to use reinforcement learning when engaging an attacker. The honeypot implemented behavioural strategies such as blocking commands, returning error messages and issuing insults. Subsequent to this, he proposes the use of game theory [30] to define the reactive action of a honeypot towards attacker's behaviour. A hierarchical probabilistic automaton is presented with the purpose of making a honeypot adaptive and autonomous. Pauna [31] also presents an adaptive honeypot using similar reinforced learning algorithms as seen in Heliza. He proposes improvements in scalability, localisation and learning capabilities.

### E. Temporal and Spatial Analysis

Relevant data analytics and modelling on the captured dataset is key to understanding malware propagation behaviour. Dagon et al [2] used time zones and diurnal patterns to model six months of honeypot activity. It shows very different compromised host activity from distinct global

regions of Asia, Americas and Europe. They also demonstrate a bias for regional online populations after sandboxing the captured malware. Pouget et al [32] examined time signatures to establish similarities of multi-headed attack tools. They clarify that the signatures are the time series of a number of different sources using a given attack per day. Time signatures were also used as a framework for attack pattern discovery in a honeypot dataset [33]. Using clique-based analysis, it establishes correlation patterns based on temporal, geospatial and IP subnets. Building on the IP subnet pattern, Chen et al [1] presents IP address blocks sharing similar attack patterns. With this pattern, they establish a method to predict cyber attacks in the future. Analysis of the data can be presented in the form of graphs, charts, plots and other visualization techniques. These convey the information more efficiently. It also presents large datasets in a meaningful way for a target audience. Honeypots have backend databases for capturing attack information, depending on the interaction level. Some honeypots embed their own visualization tools to represents the data with basic graphs and charts [13].

### III. METHODOLOGY

A honeypot is deployed to attract as much malicious traffic as possible. Making a honeypot available online attracts a variety of attack activity from geo-diverse locations. Automated botnet attacks will endeavour to compromise the machine and report back to a C&C. Kippo [13] is a medium interaction SSH honeypot designed to log attacks and entire shell interaction performed by the attacker. It is available from Google Code and GitHub and can be installed on Linux or virtual platforms. A file of acceptable username and passwords can be created, permitting or denying compromise attempts. Modifying this file increases or decreases the number of successful attempts. The most beneficial property of Kippo is its ability to be modified to simulate other file systems and log all interactions with a botnet or human attacks. For these reasons, it was chosen for our research to allow the maximum level of attack traffic, through SSH. The honeypot was deployed on an unfiltered network with port forwarding enabled. The honeypot was only accessible via SSH on port 22 and located at time zone GMT. The WAN IP on this network was static, providing uninterrupted availability to the honeypot. Over a period of 3 months, the honeypot captured over 6 millions lines of data in 367 log files. For every successful compromise, it created a playback file showing the entire interaction performed by the attacker. Summary information is as follows :

- 423228 login attempts
- 413362 unsuccessful logins
- 9866 successful logins
- 5297 distinct IP sources (all attempts)
- 31328 commands
- 5368 downloads

The IP addresses involved in attacks show a very diverse spatial distribution (Fig. 1).



Figure 1: Geo-spatial distribution of attack IPs

Temporal analysis on the dataset demonstrated a temporal pattern as seen in Fig. 2

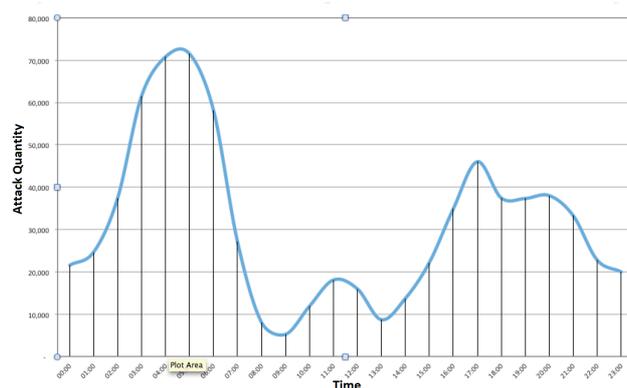


Figure 2: Temporal pattern of all attack occurrences on honeypot

Rotating the graph in Fig. 2 and overlaying it on Fig. 1, we can visualise this temporal pattern and associated volume of attacks (Fig. 3).



Figure 3: Diurnal pattern of attack sources

### IV. ANALYSIS AND RESULTS

Kippo also provides a facility to save, but not execute files downloaded to the honeypot. By sandboxing downloaded files, observing shell interactions and consulting threat advisories, it was possible to distinguish different attack types. For the deployment duration, the honeypot observed six major attack types, to and from differing IP sources. Other attack types also occurred in insignificant numbers. The six major attack types were:



Figure 4(a): Geo distribution of XOR and Recon IP sources



Figure 4(b): Geo distribution of BillGates IP sources

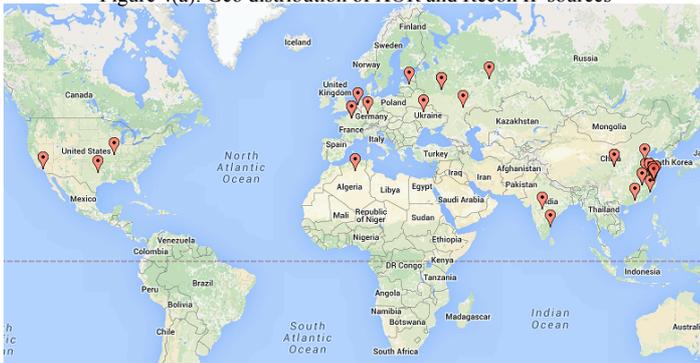


Figure 4(c): Geo distribution of Failed attack IP sources



Figure 4(d): Geo distribution of launch command from IP sources (orange) to victim IP (green)

- XOR DDoS Botnet
- Recon Scripts
- BillGates Botnet
- Failed attempts
- Launch attack command
- Dictionary Attack

- continuously over a 24-hour period. IP sources are seen in Fig. 4(b).
- Failed attacks come from well-dispersed geographic locations as seen in Fig. 4(c).
- Launch commands attempt to connect to a victim IP through various attack vectors such as http, ftp and occasionally ICMP. The honeypot is configured not to allow this connection, yet the action is still recorded. Fig. 4(d) shows the command source IPs (orange) and the victims IPs (green). All launch IP sources were from Amazon Web Services (AWS). The orange markers in Fig. 4(d) are AWS global sites.

The Dictionary attacks were responsible for nearly 94% of all honeypot activity as they continuously tried to compromise the honeypot with sequential usernames and passwords. They were recorded by Kippo but the events do not provide any further information for analysis. The pattern of attack for dictionary attacks are different from failed attempts and are therefore identifiable. They distort the legitimate ‘failed’ attack attempts and are removed from the dataset. After removing the Dictionary attacks, the IP sources of failed attacks still had a wide global distribution. The remaining types were extracted and evaluated. Each attack type was allocated a dataset. The global distribution of attack source IP addresses was mapped using Batchgeo [34] and appear in Fig. 4 (a, b, c, d). The log files provided relevant information about each attack type:

- XOR and Recon are linked. The XOR attack tool consists of batch files that attempt to compromise the honeypot. Exactly 10 minutes after this compromise the Recon attack checks for the presence of *sftp.pid*, which was created by the XOR attack. Fig. 4(a) shows both attack types, and their IP geolocation.
- BillGates botnet is an attack tool that runs constantly from China and South Korea. It is seen attacking the honeypot

It is possible to ascertain the date and time stamp of each attack type from the log files. The temporal patterns for the 5 attack types are displayed in Fig. 5 (a, b, c, d). X axis represents hourly timeslots, Y axis represents attack quantity. Patterns can be seen that match the geolocation of IP addresses in Fig. 4(a,b,c,d). Because XOR and Recon attacks originate in North America (Fig. 4(a)), the temporal pattern can be seen in Fig. 5(a). As explained, BillGates botnet continuously launch attacks from Asian sources (Fig. 4(b)), yet it still observes a temporal bias (line) in Fig. 5(b). The diverse geographic spread for Failed attacks (Fig. 4(c)) is reflected in Fig. 5(c). As the Launch command attack uses AWS global locations as IP sources, it produces a temporal pattern seen in Fig. 5(d). Both Failed and Launch attacks in Fig. 5(c) and (d) respectively, correlate with the overall temporal variance seen in Fig. 2.

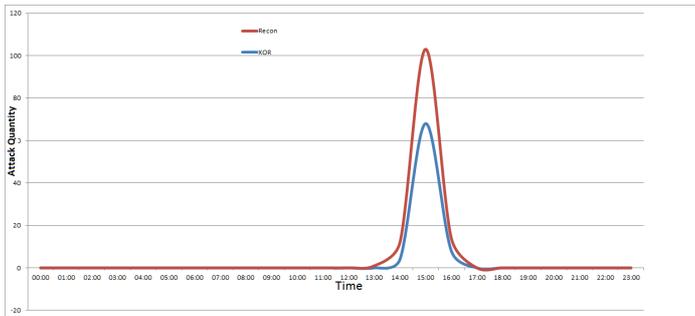


Figure 5(a): Temporal variation of XOR and Recon

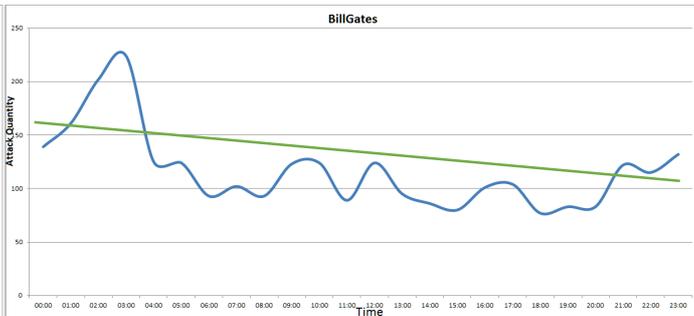


Figure 5(b): Temporal variation of BillGates botnet

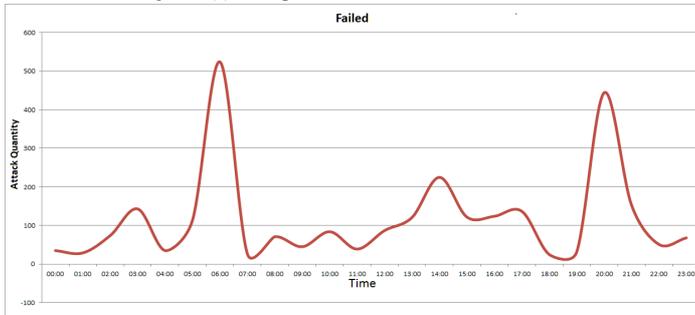


Figure 5(c): Temporal variation of failed attacks

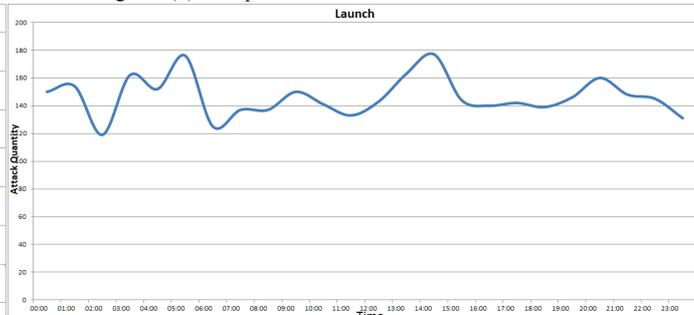


Figure 5(d): Temporal variation of Launch command

The probability of an occurrence of one attack type against others, at a particular hour, is displayed in Table 1. The quantity of each attack visible on the Y-axis, on charts in Fig. 5, is used to calculate this probability. Attacks logged between hourly slots are associated with that hour i.e. attacks between 05:00 and 05:59 are associated with 05:00. The percentages are calculated by dividing the attack type quantities per hour, by the total number of successful and failed attacks for the three month period (9866 and 2807 respectively). Fig. 6 represents these probabilities on a simple X-Y axis. It presents a diverse picture of honeypot activity during individual and groups of hourly timeslots.

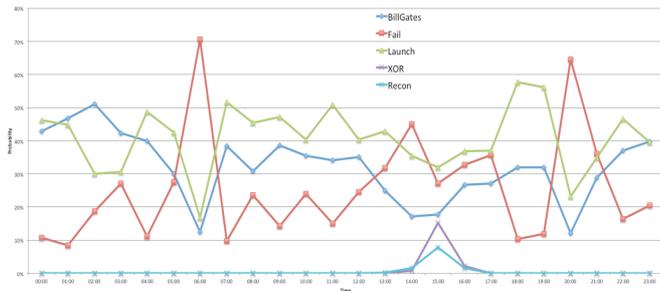


Figure 6: Probability of attack type occurrence per hour

TABLE 1: PROBABILITY OF ATTACK TYPE OCCURRENCE PER HOUR

Time	BillGates	Fail	Launch	XOR	Recon
00:00	43%	11%	46%	0%	0%
01:00	47%	8%	45%	0%	0%
02:00	51%	19%	30%	0%	0%
03:00	42%	27%	31%	0%	0%
04:00	40%	11%	49%	0%	0%
05:00	30%	28%	43%	0%	0%
06:00	13%	71%	17%	0%	0%
07:00	38%	10%	52%	0%	0%
08:00	31%	24%	46%	0%	0%
09:00	39%	14%	47%	0%	0%
10:00	36%	24%	40%	0%	0%
11:00	34%	15%	51%	0%	0%
12:00	35%	25%	40%	0%	0%
13:00	25%	32%	43%	0%	0%
14:00	17%	45%	35%	1%	2%
15:00	18%	27%	32%	15%	8%
16:00	27%	33%	37%	2%	2%
17:00	27%	36%	37%	0%	0%
18:00	32%	10%	58%	0%	0%
19:00	32%	12%	56%	0%	0%
20:00	12%	65%	23%	0%	0%
21:00	29%	36%	35%	0%	0%
22:00	37%	16%	47%	0%	0%
23:00	40%	21%	40%	0%	0%

TABLE 2: DEVIATION OF RANDOM SAMPLE FROM TABLE 1 PROBABILITIES

Time	BillGates	Fail	Launch	XOR	Recon
00:00	13%	-9%	-4%	0%	0%
01:00	7%	-12%	5%	0%	0%
02:00	11%	-1%	-10%	0%	0%
03:00	12%	7%	-19%	0%	0%
04:00	0%	1%	-1%	0%	0%
05:00	-10%	-2%	13%	0%	0%
06:00	-7%	11%	-3%	0%	0%
07:00	8%	0%	-8%	0%	0%
08:00	-9%	4%	6%	0%	0%
09:00	-11%	4%	7%	0%	0%
10:00	6%	-6%	0%	0%	0%
11:00	14%	-5%	-9%	0%	0%
12:00	15%	-25%	10%	0%	0%
13:00	5%	2%	-7%	0%	0%
14:00	-13%	-5%	15%	20%	10%
15:00	-22%	-3%	2%	-6%	-7%
16:00	-3%	-7%	7%	-10%	0%
17:00	17%	-14%	-3%	0%	0%
18:00	2%	-10%	8%	0%	0%
19:00	12%	2%	-14%	0%	0%
20:00	2%	15%	-17%	0%	0%
21:00	-1%	-4%	5%	0%	0%
22:00	7%	-14%	7%	0%	0%
23:00	0%	-9%	10%	0%	0%

Testing these probabilities, 240 random attacks were taken from the entire dataset, representing 10 samples for each of the 24 hours. These 240 samples were examined for their attack type and collated to create a new dataset of attack type probability. This dataset is compared with the probabilities in Table 1 and the deviation is presented in Table 2. The table demonstrates a favourable correlation of the random sample with the overall honeypot probability model and validates the use of temporal patterns to establish attack type probability.

## V. CONCLUSION

This paper demonstrates how temporal analysis of a honeypot dataset can be used to determine attack type probability. The probability model is tested by randomly sampling attack types from the entire dataset and comparing this sample with the original attack probabilities. The actual attack types encountered are not particularly relevant as different honeypots can be configured for specific attack vectors. Prior to the calculation of probabilities, it is necessary to perform analysis of the current honeypot dataset to determine attack types and model behaviour probability. Thereafter, automating this process and scheduling the process daily can dynamically keep the honeypot relevant. The model can help predict attack activity and be utilised to identify activity outside of the norm, such as zero day attacks. A honeypot can use this process to adopt various security levels for certain times of the day to collect more relevant data without compromising itself or accidentally contributing to further attacks.

## REFERENCES

- [1] Chen, Y.Z., Huang, Z.G., Xu, S. and Lai, Y.C., 2015, "Spatiotemporal patterns and predictability of cyberattacks", *PLoS one*, 10(5), p.e0124472.
- [2] Dagon D, Zou C, Lee W., 2006, "Modeling botnet propagation using time zones", In Proceedings of the 13th Network and Distributed System Security Symposium NDSS
- [3] Yin Minn Pa, 2015, "IoT POT: Analysing the Rise of IoT Compromises", 9th USENIX Workshop on Offensive Technologies
- [4] BELLOVIN, S.M., 1993. "Packets found on an internet", *SIGCOMM Comput. Commun. Rev.*, 23 (3), pp. 26-31.
- [5] PROVOS, N. "A virtual honeypot framework.", In *SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium*,
- [6] P. Baecher, M. Koetter, T. Holz, M. Dornseif, and F. Freiling, "The Nepenthes Platform: An Efficient Approach to Collect Malware," 9th International Symposium on Recent Advances in Intrusion Detection
- [7] Georgios Portokalidis, Asia Slowinska, and Herbert Bos, "Argos: an Emulator for Fingerprinting Zero-Day Attacks", *ACM SIGOPS Operating Systems Review*, vol. 40, no. 4, October 2006
- [8] Pouget, F., Dacier, M. and Pham, V.H., 2005. on the Advantages of Deploying a Large Scale Distributed Honeypot Platform. In Proceedings of the E-Crime and Computer Evidence Conference
- [9] D. Watson and J. Riden, 2008, "The Honeynet Project: Data Collection Tools, Infrastructure, Archives and Analysis," *Information Security Threats Data Collection and Sharing, WISTDCS '08*.
- [10] Zhang, Feng, 2003, "HoneyPot: a supplemented active defense system for network security", *Parallel and Distributed Computing, Applications and Technologies, PDCAT'2003*
- [11] Spitzner, L., "Honeytokens: The other HoneyPot", *Symantec Connect*, 2003. [Online] Available: <http://www.securityfocus.com/infocus/1713> (Access Date: 27-06-2016)
- [12] Christian Seifert, Ian Welch, Peter Komisarczuk, 2006, "Taxonomy of honeypots", *Technical Report CS-TR-06/12*, Victoria University of Wellington
- [13] Valli C., Rabadia P., Woodward A., 2013, "Patterns and Patter - An Investigation into SSH Activity Using Kippo honeypots", *Australian Digital Forensics Conference*.
- [14] Provataki A., 2013, "Differential malware forensics," *Digital Investigation*. 10, 4 (December 2013), 311-322
- [15] D. Ramsbrock, 2007, "Profiling Attacker Behavior Following SSH Compromises", *37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'07)*, Edinburgh
- [16] T. Holz, 2005, "Detecting honeypots and other suspicious environments", *Proceedings from the Sixth Annual IEEE SMC Information Assurance Workshop*.
- [17] Suchacka, G., 2014, "Analysis of aggregated bot and human traffic on e-commerce site". *Polish Information Processing Society*
- [18] Dagon D., 2007, "A taxonomy of botnet structures", *Computer Security Applications Conference*, DOI:10.1109/ACSAC.2007.44
- [19] Fedynyshyn, 2011, "Detection and classification of different botnet C&C channels", In *Proceedings of the 8th international conference on Autonomic and trusted computing (ATC'11)*
- [20] Freiling, F.C., Holz, T. and Wicherski, G., 2005, September. "Botnet tracking: Exploring a root-cause methodology to prevent distributed denial-of-service attacks", In *European Symposium on Research in Computer Security* (pp. 319-335). Springer Berlin Heidelberg.
- [21] Awadi, A.H.R.A. and Belaton, B., 2015. Multi-phase IRC botnet and botnet behavior detection model. *arXiv preprint arXiv:1501.03241*.
- [22] Choi, H. and Lee, H., 2012. Identifying botnets by capturing group activities in DNS traffic. *Computer Networks*, 56(1), pp.20-33.
- [23] Wang, K., Huang, C.Y., Lin, S.J. and Lin, Y.D., 2011. A fuzzy pattern-based filtering algorithm for botnet detection. *Computer Networks*, 55(15), pp.3275-3286.
- [24] J. Francois, S. Wang, R. State, and T. Engel, 2011, "Bottrack: Tracking botnets using netflow and pagerank," in *Proceedings of IFIP/TC6 Networking*
- [25] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," 1998, *Stanford InfoLab*,
- [26] Rubin, B., 2006, *Computer Security Education and Research: Handle with Care*, *IEEE Security and Privacy Magazine*, 4 (6):56-59
- [27] McCarty, B., 2003. "The honeynet arms race", *IEEE Security & Privacy Magazine*, 1 (6), pp. 79-82.
- [28] Hayatle, O., Otrok, H. and Youssef, A., 2013. A Markov Decision Process Model for High Interaction Honeypots. *Information Security Journal: A Global Perspective*, 22(4), pp.159-170.
- [29] Wagener, G., Dulaunoy, A. and Engel, T., 2011. Heliza: talking dirty to the attackers. *Journal in computer virology*, 7(3), pp.221-232.
- [30] Wagener, G., 2011. *Self-Adaptive Honeypots Coercing and Assessing Attacker Behaviour* (Doctoral dissertation, INRIA Nancy).
- [31] Pauna, A., 2014, *RASSH - Reinforced adaptive SSH honeypot*, 10th International Conference on Communications (COMM)
- [32] Pouget, F., Urvoy-Keller, G. and Dacier, M., 2006, June. Time signatures to detect multi-headed stealthy attack tools. In *Proceedings of the 18th annual first conference*, Baltimore, MD, June (pp. 25-30)
- [33] Thonnard, O. and Dacier, M., 2008. A framework for attack patterns' discovery in honeynet data. *Digital Investigation*, 5, pp.S128-S139.
- [34] "Create Google Maps with your data" [Online] Available: [www.batchgeo.com](http://www.batchgeo.com) (Access Date: 05-08-2016)

## **Session 14: Infonomics and e-Technology**

Title: Client Side Calculation of 'Bacon Number'  
(Authors: Thomas Briggs, Aspen Olmsted)

Title: intel-LEACH: An Optimal Framework for Node Selection using Dynamic Clustering for Wireless Sensor Networks  
(Authors: Prathap Siddavaatam, Reza Sedaghat, Aakriti Tarun Sharma)

Title: Enhancing Security in the Cloud: when Traceability meets Access Control  
(Authors: Clara Bertolissi, Omar Boucelma, Worachet Uttha)

Title: Database Multi-factor Authentication via Pluggable Authentication Modules  
(Authors: Cameron Hamilton, Aspen Olmstead)

# Client Side Calculation of ‘Bacon Number’

Thomas Briggs, Aspen Olmsted  
College of Charleston, Department of Computer Science  
Charleston, SC, United States  
tabriggs@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—This paper presents an architecture for answering the popular ‘bacon number’ game utilizing the IMDB dataset. The unique contribution to this well-known game is this paper presents implementing a breadth traversal algorithm and rendering of the data completely in the client. This paper outlines the architecture and presents evidence of comparable speed in the complete client-side rendering of the computation of large data objects as compared to traditional server-side compiled languages.

**Keywords**-component; client-side; JavaScript; Node.js; Bacon Number

## I. INTRODUCTION

Web applications are the dominant way with which information is shared across the internet. With the rise of dependable open source technology and increasingly powerful browsers, the traditional web application framework is changing. The traditional, monolithic applications that have dominated web applications for the last 20 years are giving way to a new trend in web service architecture. With technologies like Node.js and ever more powerful browsers, architects and developers can leverage these technologies to turn the traditional application stack on its side, scaling horizontally instead of vertically. Doing so results in a significantly more flexible application stack as well as limiting the amount of hardware needed to deploy and maintain a web application.

Until the last 5 years, Web application architecture has leverage the well-established LAMP stack. LAMP is an acronym representing the traditional tools for Web application development: Linux, Apache, MySQL, and PHP. Data generated from and retrieved by these applications reside in a relational database. These development stacks have proven themselves to be reliable. However, times are changing Louridas [1]

In the last two years, mobile applications have come to dominate how people consume and create data. This fact presents a challenge to the traditional web application stack because users must be able to obtain data from a wide range of end-user platforms. This need has given rise to client side rendering and single page applications. These applications, are powered by a more flexible application stack than the traditional LAMP stack. In the single page architecture, the client is not a dumb consumer of server-side content. Instead, the client is where the heavy lifting is done, a role traditionally reserved to the server.

## II. RELATED WORK

Abidi presents an argument and methodology for not only creating a single page application but creating a single page application with the end goal being an analysis of data completely in the browser [2]. Similar to this paper's idea, Abidi introduces the client as an area for heavy computation and decision making. Traditionally, heavy computation has been reserved to the server. Abidi presents the browser as the universal operating system, with JavaScript being the language of this universal operating system. While the proposed paradigm is applied to medical data leveraging the RegidsDG framework, the authors are keen to the fact that with the increased power of the browser, memory, and CPU intensive routines can be executed by the client. The authors took the RegidsDG framework and replaced the middleware with client-side JavaScript, served by Node.js. Their conclusions support this paper's premise that JavaScript is powerful for data visualization and algorithm processing by the client.

Ioannis did work comparing different servers and their speed at serving content [3]. In his work, he compared Nginx, Node.js, and Apache. The purpose of the work was to analyze the strengths and weakness of each server when subjected to different stresses. Ioannis made an analysis of each server's ability to perform IO operations. Node.js was unquestionably the most efficient utilizer of CPU and dominated speed of IO operations. The motivation for this work in connection with this project is to present other work that supports the idea that Node.js is a fully mature server. Ionnis found that by tracking memory consumption, IO statistics and CPU usage, Apache was the least performing in terms of the three categories mentioned. Nginx was a great solution for serving static content. However, most web services and applications do not have static content anymore. The findings of Ionnis confirm that Node.js is one of the most viable options for a back-end server infrastructure for serving content. Node.js outperformed each server in terms of memory consumption, IO and CPU utilization when a serving non-static content. This work confirms that Node.js, a pure JavaScript solution can be used as a replacement for traditional servers. This paper utilizes Node.js as its language and framework of choice for the back end server.

A very interesting related work to doing big data within a pure JavaScript/client-side framework is the work presented by Lin [4]. Lin introduces a modified JavaScript engine called ‘Afterburner’ for doing big data analytics completely within the Firefox browser. The authors sought to take the existing JavaScript Engine and create stronger typing of the JavaScript language to improve the runtime execution. To do this the

author's leveraged `asm.js`, a low-level JavaScript library that allows variable typing. The authors also removed JavaScript's garbage collection and just-in-time compilation features to strip the language down. This allowed the authors to have data sent to the browser and processed live. This work illustrates an experimental idea that some changes to the JavaScript engine and the language interpreter can make javascript significantly faster and a viable option for large data processing client side. Lin's work connects to this paper because, in this paper, we present a similar idea to processing large volumes of data live in the client.

### III. MOTIVATING EXAMPLE

As web browsers become more powerful than previous versions, more calculations can be performed in the browser. This paper seeks to present an example that demonstrates that the browser is a sufficient area to perform the computation. The browser is not just a technology for the client-side display. While the traditional application stack will always have its place, this paper seeks to show that data can be fed directly into the client and computations performed live with no performance loss client side as compared to the traditional application stack. The paper presents an architecture for a client side rendition of the oracle of bacon [5].

### IV. IMPLEMENTATION

This project sought to take a well-known data set, the IMDB dataset and implement the n-degrees of Kevin Bacon game porting the breadth search algorithm to JavaScript Client-side [6]. The difference being that the actual computation of the 'bacon' number for a given actor would be calculated client side. This paper presents a unique architecture to solve for the bacon number. Also, this paper outlines a series of language and architectural benchmarks in order to demonstrate that JavaScript is a mature enough technology to do serious work client side.

The data was parsed by a Perl program server side due to the strength and ease of Perl Regex. The parser is set up to accept an arbitrary number of ARGV from the IMDB dataset residing on disk. The result of this parse was two 150 MB JSON objects written to disk. Each JSON object consists of multiple associative arrays. The first JSON object, *actorsToMovies*, is a mapping of each actor as a key, and their movies as an array. The second JSON object called *moviesToActors* is the same structure, but the keys and values are reversed; each key is a movie, and the value is an array of every actor appearing in that movie. This process was timed with actors list only and took 3 minutes to compute. The bottleneck was CPU. The parser was a single process running on a single thread on an Intel i5-320M. In the architecture of the web application, the parsing process would only need to be executed daily, since the data is pulled from the IMDB and only updates daily.

Next, the two JSON objects are read by the Node.js server using the Express framework to expose the JSON objects as a REST endpoint. There are two REST endpoints, one for each respective JSON object. The JSON objects are read as chunked streams to the client for efficiency.

### Algorithm 1 Synchronous Data Integration

**INPUT:** Actor

**OUTPUT:** Bacon Number

```

push actor onto queue
while actors queue !empty:
do:
  lookup actor
  pull actors movies
  cache seen movies
do:
  lookup actors from movie
  push actors onto queue
  if kevin bacon
    done
  else
    store seen actors
end;
increment bacon number
end;

```

What is unique to this paper, is instead of computing the bacon number server side by first ingesting input from the client, sending this data to the server, executing the traversal of the JSON objects server side, then sending the result to the client; the client on page load, makes two Ajax calls to the Node.js server retrieving the JSON objects. This only takes 2-3 milliseconds. From here, the smart client takes over. The JSON objects are stored within the client, and the JavaScript then takes in the actor as a param from the user client side and computes the bacon number. The breadth search algorithm is implemented completely on the client side. After the initial page load, the client no longer needs to retrieve data from the server. The algorithm is shown in Algorithm 1.

This implementation is advantageous because there is no guarantee which platform the end user will be using. The architecture is chosen so as to be completely agnostic as to the device the user chooses to view the web page with. Also, the burden of computation is no longer on the server; it lays with the client. The browser is capable requesting more RAM and CPU from the client, freeing the server of resources. There is also no database in this architecture. The solution was designed to be implemented entirely in RAM and to read the data 'hot' from the server. The data is cached by the client, allowing the client to play the game as many times as the client would like, constantly inputting in new actors to discover their Bacon number. Fig. 1 shows the outline of the application flow.

### V. CONCLUSION AND FUTURE WORK

Due to time constraints, the full bacon number calculation was not completed. However, the infrastructure for the application was built and successful. This paper demonstrates that the client can be leveraged and mined for its' CPU and RAM. This paradigm is the future of application development. With the chosen architecture, the Bacon number can be calculated on an OS independent basis.

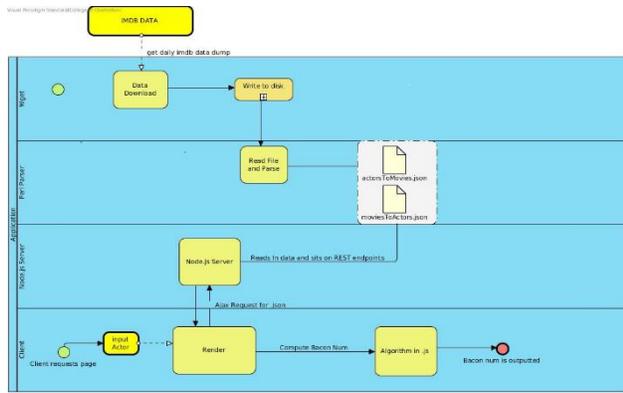


Figure 1.Application BPMN

Further work needs to be done to implement the calculation of the Bacon number completely. An area of optimization on the server can be the caching of the JSON objects on the Node.js server-to-server multiple clients and save on IO. Currently, each request requires an IO operation server side. Since the data served is static, the data should be cached and limited to 1 IO operation per day by the server. This will save on performance. The idea and architecture presented in this paper are new, and not widely if ever, developed and deployed.

## VI. REFERENCES

[1] P. Louridas, "Component Stacks for Enterprise Applications," *IEEE Software*, no. 33.2, pp. 93-98, 2016.

[2] L. Abidi, "Towards an Environment for doing Data Science that runs in Browsers," in *2015 IEEE International Conference on*, 2015.

[3] I. K. Chaniotis, K. D. Kyriakos-Ioannis and N. D. Tselikas, "Is Node.js a viable option for building modern web applications? A performance evaluation study.," *Computing*, vol. 97.10, pp. 1023-1044, 2015.

[4] K. E. Gebaly and J. Lin, "Afterburner: The Case for In-Browser Analytics.," *arXiv preprint*, vol. 1605.04035, 2016.

[5] P. Reynolds, "The Oracle of Bacon," [Online]. Available: <https://oracleofbacon.org/how.php>. [Accessed 20 June 2017].

[6] P. Reynolds, "Patrick Reynolds," [Online]. Available: <http://piki.org/patrick/>. [Accessed 20 June 2017].

# intel-LEACH: An Optimal Framework for Node Selection using Dynamic Clustering for Wireless Sensor Networks

Prathap Siddavaatam<sup>\*</sup>, Reza Sedaghat<sup>†</sup>, Aakriti Tarun Sharma<sup>‡</sup>

OPR-AL Labs,

Department of Electrical and Computer Engineering,

Ryerson University,

Toronto, ON M5B 2K3

Email: <sup>\*</sup>prathap.siddavaatam@ryerson.ca, <sup>†</sup>rasedagha@ee.ryerson.ca

<sup>‡</sup>aakrititarun.sharma@ryerson.ca

**Abstract**—Over the last decade, the field of Wireless Sensor Networks (WSNs) has made expansive strides in the area of Radio Communication Systems. A vast majority of these systems are deployed in failure-prone environments which results in chronic communication losses due to unreliable wireless connections, malicious attacks and resource-constrained features. Hence WSNs necessitate adaptive protocol frameworks applicable for differing network densities either as sparse or denser deployments. One of the major challenges in WSNs is to achieve an optimal trade-off between data precision and energy efficiency of a sensor node from the perspective of network longevity. In this article, we propose an optimized protocol *intel-LEACH* for selecting high yield nodes among randomly deployed sensor nodes based on dynamic optimization strategy. Our proposed model selects a node that can guarantee higher data precision and has maximal residual energy to serve as a cluster head (CH). Meanwhile, the others sensor nodes apart from CHs must also exhibit capability in terms of data transmission accuracy to join a given cluster irrespective of the distance between the sensor nodes and CHs. Thus our model minimizes loss of high yield nodes for every round and overall WSN performance is boosted by its network lifetime extension.

**Keywords**—Wireless Sensor Network, Energy Efficiency, Clustering; Data Accuracy; Community Networks; LEACH-C; Network Lifetime

## I. INTRODUCTION

Wireless sensor networks (WSNs) are one of the most rapidly developing information technologies destined to have a variety of applications in Next Generation Networks (NGNs). It has a wide range of applications in both military and civilian domains[1]. Research has been undertaken in a plethora of fields related to WSNs such as routing, MAC and collaborative data gathering mechanisms [2, 3]. WSNs are composed of massive, small and low-cost sensor nodes deployed to perform tasks such as environmental monitoring, vehicle route tracking, home automation, health and medical automation among others, often forming a self-organized network system through wireless communication. The sensor nodes are connected to either one or more nodes and the base station also known as sink. The task of WSNs is to cooperatively sense, collect and

process information about objects in a given coverage area, and subsequently communicate it to the observer for processing and analyzing as depicted in Fig. 1.

Two important design objectives are paramount in research related to WSNs: (1) Energy efficiency and (2) Data accuracy. The former is a key design aspect because the nodes are severely energy constrained, and battery replenishment is often not practical. The latter is important as unreliable data from key nodes may decrease the quality of service (QoS) of the entire WSN [4]. The design of poor network topologies will lead to a drastic decrease in the network longevity since a large percentile of the energy is expended in the process of communication. Thus it is imperative to conserve energy so as to increase the stability and lifetime of the network. Consequently, designing intelligent self-aware protocols aid in alleviating early network failures and enhance the stability of the overall system.

### A. Related Work

We consider Clustered sensor networks in this article since clustering allows for scalability of MAC and routing. Also, Cluster heads serve as fusion points for data aggregation to ensure that the actual amount of data transmitted to the base station is minimized. Clustered sensor networks can be classified into two broad types: (1) Homogeneous and (2) Heterogeneous – sensor networks[5]. In a heterogeneous WSN, two or more different types of nodes with different battery energy and functionality are deployed. Our discussion in this article involves only homogeneous networks. In homogeneous networks all the sensor nodes are identical in terms of battery energy and data throughput. With purely static clustering (the idea being that cluster heads are permanently elected to serve for complete lifetime of the network) in a homogeneous network, it is evident that the cluster head nodes will be over-loaded with the long range transmissions to the remote base station, and the extra processing necessary for data aggregation and protocol co-ordination. As a result the cluster head nodes expire before other nodes. However the ideal

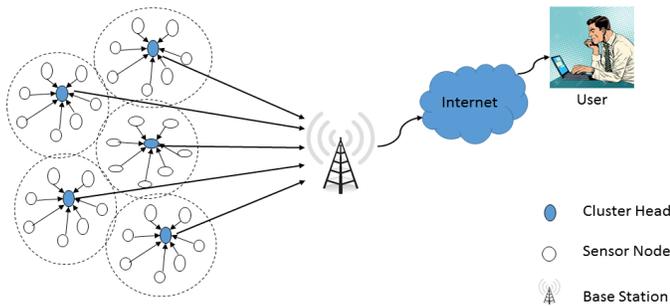


Figure 1: Illustration of the 2-tier Clustering Topology in WSN

situation is to ensure that all the nodes run out of their battery at about the same time, so that very little residual energy is wasted when the system expires. There are many algorithms proposed to achieve this[6–8] but an efficient way is to rotate the role of cluster head randomly and periodically throughout all the available nodes as proposed in LEACH[9]. LEACH-Centralized (LEACH-C)[3] is an enhanced algorithm proposed by same authors of original LEACH with a single exception in initial set-up phase of the algorithm. In LEACH-C, during the set-up phase all the sensor nodes are programmed to send the coordinate and energy level information to the sink. The base station then performs the same process of cluster head selection as in LEACH (where it is performed locally in nodes). The algorithm used in clustering process of LEACH-C is Simulated Annealing (SA)[10].

### B. Scope of Our Contribution

It is well known fact that SA inherently has several drawbacks[11, 12] within the context of clustering in networks. The problem pertains to the basic assumption of applying SA technique in determining the location of sensors in WSNs. Although the problem of localization is solved in some cases, many involving low density networks have sensor nodes prone to flip ambiguity problem[11, 12]. To address this issue, our model incorporates a clustering algorithm [13] which is different from SA algorithm used in LEACH-C protocol. This method uses the Community Detection Toolbox[14] based in MATLAB to perform computations and form communities or clusters for sensor nodes sans the metaheuristic problems[11, 12]. In this article, our proposed model known as *intel-LEACH* achieves two goals: (1) the residual energy of the nodes last over a longer period of time resulting in increased network lifetime and stability (2) providing precision based results in terms of accuracy to maintain QoS of the overall network. By taking these two goals as constraints to optimize the network, *intel-LEACH* formulates this problem as a bi-dimensional knapsack problem[15]. Our model works in two phases to address the issue of trade-off optimization. In the first phase, upper bounds are found to fix the most possible variables in order to reduce the size of the problem. Branch-and-Bound algorithm is then used in the second phase to carry out these upper bounds[16]. The solution to this problem

results as a best set of clusters that can be used to achieve our goals in a WSN.

### C. Organization

Section II details benchmark protocol and contemporary models for clustering, data accuracy estimation and optimization strategies. Section III details our proposed design. Section IV describes simulation model and outlines performance analysis of *intel-LEACH* compared to benchmark protocols. Section V concludes the superior performance of the proposed protocol.

## II. PRELIMINARIES

### A. LEACH Protocol

The LEACH algorithm consists of two phases – the set-up phase and the steady phase. A stochastic algorithm is used by the nodes to determine the clusters. In this algorithm, a node which wants to participate in CH election generates a random number between 0 and 1 and if the random number is below probability based threshold value  $T(n_i)$ , calculated according to (1), the node  $n_i$  is elected as cluster head.

$$T(n_i) = \begin{cases} \frac{P}{1 - P * (r \bmod (1/P))}, & n_i \in N \\ 0, & otherwise \end{cases} \quad (1)$$

Where  $P$  is desired percentage of cluster heads,  $r$  is current round,  $N$  is set of nodes that have not become cluster head nodes in last  $1/P$  rounds. LEACH is mostly considered to be a benchmark routing protocol in WSN researchers' community. As mentioned in the Section II, LEACH-C[3] only differs from LEACH with respect to its initial set-up phase which it is done locally in sensor node  $n_i$  for LEACH.

### B. Energy Model

The energy model in this article follows the standard radio model discussed by [3]. In this model as depicted in Fig. 2, both the free space and multi-path fading channels are used depending on the distance  $d$  between the transmitter and receiver. When this distance  $d$  is less than a threshold value  $d_0$ , then the free space (fs) model is used, otherwise, the multipath (mp) model is used. Let  $E_{elec}$ ,  $\mathcal{E}_{fs}$  and  $\mathcal{E}_{amp}$  be the energy required by the electronics circuit and by the amplifier in free space and multipath respectively. Then the energy required by the radio to transmit an  $k$ -bit message over a distance  $d$  is given as follows:

$$E_{TX}(k, d) = \begin{cases} H \times E_{elec} + k \times \mathcal{E}_{amp} \times d^2 & \text{if } d < d_0 \\ H \times E_{elec} + k \times \mathcal{E}_{amp} \times d^4 & \text{if } d \geq d_0 \end{cases} \quad (2)$$

To receive data, nodes consume power in each received bit. This used power is described as the listening power and it can be calculated according to (3). This power model has been used to design our simulation.

$$E_{RX}(k, d) = k \times E_{elec} \quad (3)$$

The  $E_{elec}$  depends on several factors such as digital coding, modulation, filtering, and spreading of the signal, whereas

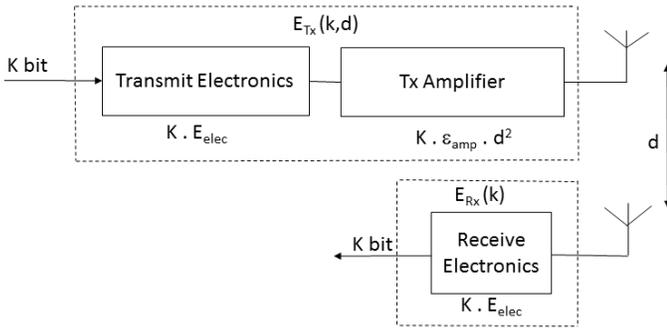


Figure 2: First Order Radio Model for Communication Systems.

the amplifier energy,  $\mathcal{E}_{amp} \times d^2 / \mathcal{E}_{amp} \times d^4$ , depends on the distance between the transmitter and the receiver and also on the acceptable bit-error rate.

### C. Cluster Formation

For our clustering purposes, we assume that WSN is a network where all the sensor nodes are deployed randomly along with a few gateways and once they are deployed, they become stationary. A sensor node can be assigned to any gateway if it is within the communication range of the sensor node. By tuning the modularity resolution parameter to scale factor of 1[17], we model the clustering of WSN as defined in [13] given by:

$$Q = \sum_{s=1}^m \frac{w_{ss}}{w} - \left(\frac{w_s}{2w}\right)^2 \quad (4)$$

where modularity  $Q$  is the quality measure of the cluster partitioned which allows quantification of the modular structure. The parameters used in forming the clusters are:  $Q$  is sum over  $m$  modules of the partitioned WSN,  $w_{ss}$  is the internal strength of module  $s$  and  $w_s$  is the total strength of the module  $s$ .

### D. Data Precision Measurement

For estimating precision of the data transmitted by any given node we use data accuracy model defined in [18] for randomly deployed sensor nodes. The main reason for choosing this model is its consideration of real environment conditions of a WSN under when some of the sensor nodes may become fault-prone and involved in unreliable transmission of data due to physical conditions. Using [18], we get the normalized estimated data accuracy model (EDAM) for  $V$  sensor nodes in the network as:

$$D_A(V) = \frac{1}{\beta} \left( 2 \sum_{i=1}^V e^{-(d_{s,i}/\theta)} \right) - \frac{1}{\beta^2} \left( \sum_{i=1}^V \sum_{j=1}^V e^{-(d_{i,j}/\theta)} \right) - \left( \frac{\left( \sum_{i=1}^V \sigma_{N_i}^2 \right)}{\beta^2 \sigma_S^2} \right) \quad (5)$$

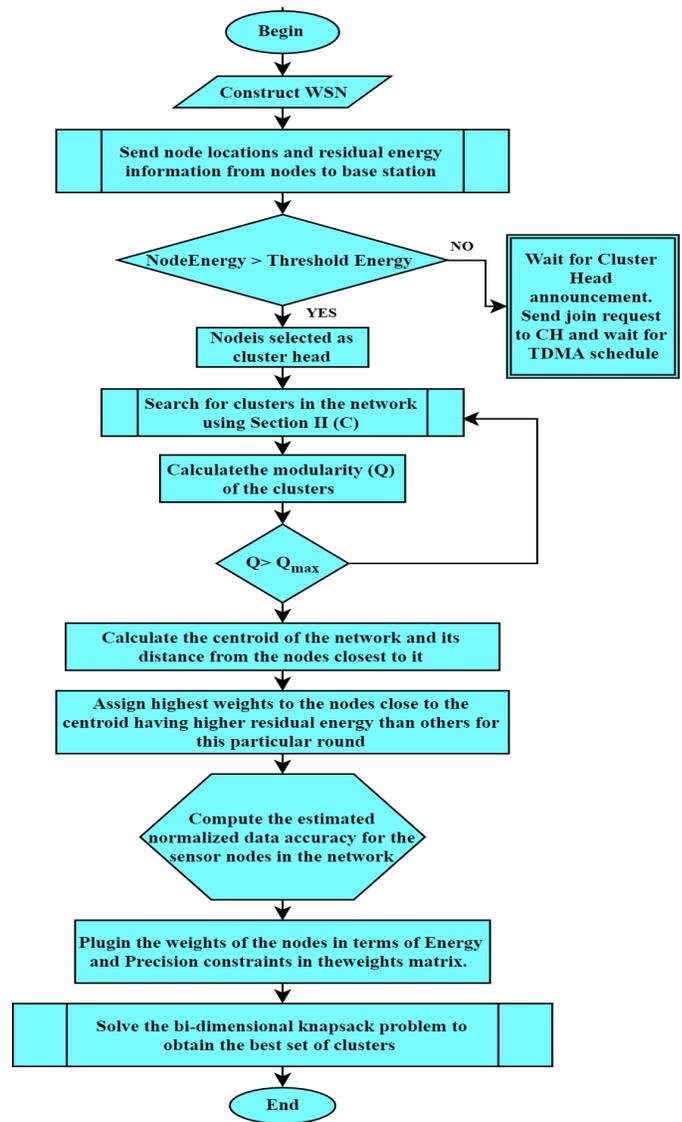
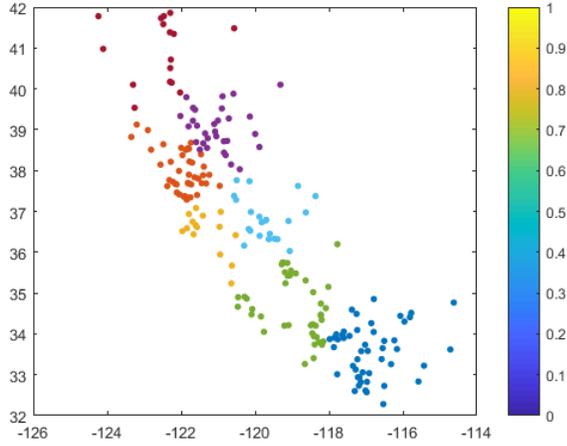


Figure 3: Flow diagram for steps in intel-LEACH Algorithm

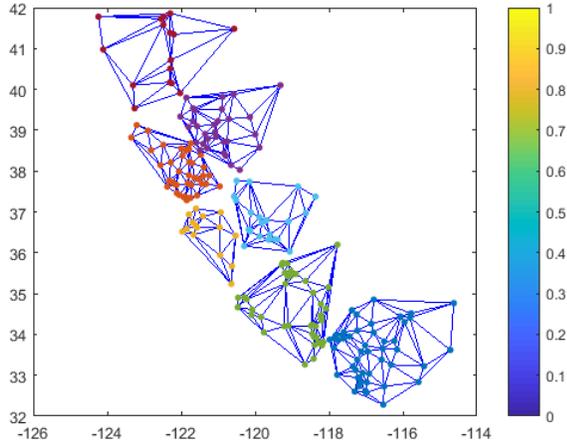
where  $d_{i,j} = \|S_i - S_j\|$  represents the Euclidian distance between node  $i$  and  $j$  of given WSN,  $K_V^{P,E}(d_{i,j}) = e^{-(d_{i,j}/\theta)}$  is the power exponential model[19], for  $\theta > 0$ , where  $\theta$  is the Range Parameter which is set for clustered networks as defined in [18] and  $\beta = V + \sigma_N^2 / \sigma_S^2$  where  $\sigma$  denotes variance of observed and noise signals at the sink.

### E. Optimization Strategy

We model the problem of optimization trade-off between energy efficiency of WSN and data signal reliability as a multidimensional knapsack problem specifically for two dimensions. A multidimensional knapsack problem(MKP)[15] is a generalization of the dynamic optimization process using knapsack problem(KP). The MKP can be formulated as



(a) Data points indicating 216 readings over 1 day period from different weather stations in California State and their scaled down coordinates used to construct the test bed network.



(b) The resultant clustering graph generated using [17]. For simulation purposes we setup the optimal number of neighbors as 7.

follows:

$$(MKP) \begin{cases} \text{maximize} & \sum_{j=1}^n p_j x_j \\ \text{such that} & \sum_{j=1}^n w_{ij} x_j \leq c_i, \quad i = 1, \dots, m \\ & x_j \in \{0, 1\} \quad j = 1, \dots, n \end{cases} \quad (6)$$

where  $n = |N|$  is the number of variables (objects),  $p_j, w_j, \forall j \in N, \forall i \in M$  and  $c$  denotes capacity in positive integers.  $(p_j, w_j)$  refer to the profit and weight associated with object  $j$  (in  $N$ ) and  $c$  is capacity of the knapsack. Without the loss of generality, we assume that the following constraint as in (7) is satisfied.

$$\text{maximize} \quad \{w_{ij} : j \in N \leq c_i \leq \sum_{j \in N} w_{ij}, \forall i \in M\} \quad (7)$$

### III. DESIGN MODEL

Our proposed protocol *intel-LEACH* for WSN has a process flow described in flow diagram Fig. 3. Using description of

the flow diagram Fig. 3, we deduce the following.

- There will be instances when more than one sensor is eligible for CH selection in the vicinity of the centroid of a cluster (refer Section II.(C))
- The proximity distance  $\psi$  of eligible sensor nodes in vicinity of centroid is system user defined.
- The set of eligible cluster heads during initial rounds of our CHs computation are assigned highest weights in weighting matrix of our model's Bi-dimensional knapsack[16] algorithm (refer Section II.(E)).
- The sensor nodes within  $\psi$  of centroid are also scanned for data accuracy levels using precision metric defined in Section II.(D)
- The candidate sensor nodes from previous step are also pooled and assigned highest weights in the next row of weighting matrix defined in Section II.(E).
- Each element  $w_{ij}$  of weighting matrix  $W$  of our model will be assigned values from the set  $\{0, 1, 2\}$  where 0, 1 and 2 signify dead, normal and task-ready eligible nodes respectively.

As the sensor nodes are selected described in Section II.(A) and subjected to steps above, our protocol then follows the trade-off model based on (6) formulated below:

$$(BKP) \begin{cases} \text{maximize} & \sum_{j=1}^n p_j x_j \\ \text{such that} & \sum_{j=1}^n w_{ij} x_j \leq c_i, \quad \forall i \in \{1, 2\} \\ & x_j \in \{0, 1\} \quad \forall j \in \{1, \dots, n\} \end{cases} \quad (8)$$

In the context of Integer Linear Programming (ILP) formulation, we use one decision matrix  $x_i : X$  which indicates if sensor  $n_i$  is assigned to be part of the cluster ( $x_i = 1$ ) or unassigned ( $x_i = 0$ ). In (8)  $c_1$  indicates the maximum capacity of the energy pool knapsack and similarly  $c_2$  indicates maximum threshold for data precision values. We observe from (8) that sensor nodes are the objects to be placed in the knapsack. The objective is to optimize the sum of their utilities under constrained energy budget  $c_i : C \forall j \in \{1, \dots, n\}$  to carry out task execution. The utilization of sensor node  $i$  in given cluster is a combination of its residual energy  $\zeta_i$  and data relevance index  $\chi_i$ . We rewrite the profit element  $p_j : P \forall j \in \{1, \dots, n\}$  in terms of coefficients  $\zeta_i$  and  $\chi_i$ . Therefore the updated profit function  $\hat{P}$  can be defined as:

$$\hat{P}(\zeta_i, w_i) = \gamma - \nu \phi(\zeta_i, w_i, \chi_i) \quad (9)$$

$$\text{where } \phi(\zeta_i, w_i, \chi_i) = \left[ \frac{w_i - \zeta_i}{\chi_i} \right], \chi_i \neq 0 \quad (10)$$

$\gamma$  and  $\nu$  are coefficients used to balance the priorities given for each term of the equation, and they depend on the application data accuracy requirements. The utilization of sensor node  $i$  in given cluster is a combination of its residual energy  $\zeta_i$  and data relevance index  $\chi_i$ . The objective of (8) is to minimize energy-precision coefficient  $\phi(\zeta_i, w_i, \chi_i)$  of selected nodes. The bi-

TABLE I: Simulation Parameters for *intel-LEACH*

Parameter	Value
Number of Nodes	216
Transmitter Electronics $E_{TX}$	50 nJ/bit
Receiver Electronics $E_{RX}$	50 nJ/bit
$E_{elec}$	50 nJ/bit
$\mathcal{E}_{amp}$	0.0013 pJ/bit/m <sup>2</sup>
$\mathcal{E}_{fs}$	100 pJ/bit/m <sup>2</sup>
Packet Transmitted from the Cluster Head to the Base Station $L_p$	6400 bits
Packet Transmitted from the Node to the Cluster Head $L_{ctr}$	200 bits
Initial Energy of the Node	0.5 J
Threshold Distance $d_0$	75 m
Range Parameter $\theta$	70

dimensional knapsack problem in (8) can be written in terms of  $\zeta_i$  and  $\chi_i$  as:

$$(BKP) \begin{cases} \text{maximize} & \sum_{j=1}^n \chi_i \hat{p}_j x_j \\ \text{such that} & \sum_{j=1}^n w_{ij} x_j \leq c_i, \quad \forall i \in \{1, 2\} \\ & x_j \in \{0, 1\} \quad \forall j \in \{1, \dots, n\} \end{cases} \quad (11)$$

#### IV. SIMULATION RESULTS

In this section, we begin by investigating the performance of clustering mechanism detailed in Section II.c by setting parameter values specified in Table. I. We obtained real world from California Irrigation Management Information System (CIMIS)[20] to setup 216 sensor nodes for WSN simulation. The entire area of WSN is scaled down to 200 x 200 sq m<sup>2</sup> so that simulations can be undertaken for testing the performance of *intel-LEACH* protocol. The dataset is generated by weather stations across the state of California, which are equipped with sensors that measure solar radiation, temperature and wind speed, among other variables. We run our simulations on temperature reading data across sensor nodes and multiple coordinate points of California state which has been to scaled down version of 200 x 200 sq m<sup>2</sup> window coordinate test bed. These coordinates of sensor nodes are illustrated in Fig. 4a We obtained temperature data over the period of 1 day which contains 216 readings from different weather stations and their scaled down coordinates to construct a network as outlined in Section II.c. Furthermore we initiate clustering algorithm of Section II.c by setting the number of neighbors for the graph as k = 7 and obtain 7 independent clusters as depicted in Fig. 4b. Next, we determined the centroid points for these m = 7 clusters. These centroid points have new coordinates which typically does not coincide with any position of the sensor nodes in WSN. Even in the case of exact match between centroid point and sensor node positions, our algorithm continues to forage for more candidate sensor nodes in the vicinity  $\psi$  to expand the pool of available candidate sensor nodes. Thus, for each cluster m, the coordinates of the centroid point  $(x_{centroid}^m, y_{centroid}^m)$  are given by:

$$x_{centroid}^m = \frac{\sum_{h=1}^{|S_m|} x_i}{|S_m|}, \quad y_{centroid}^m = \frac{\sum_{h=1}^{|S_m|} y_i}{|S_m|} \quad (12)$$

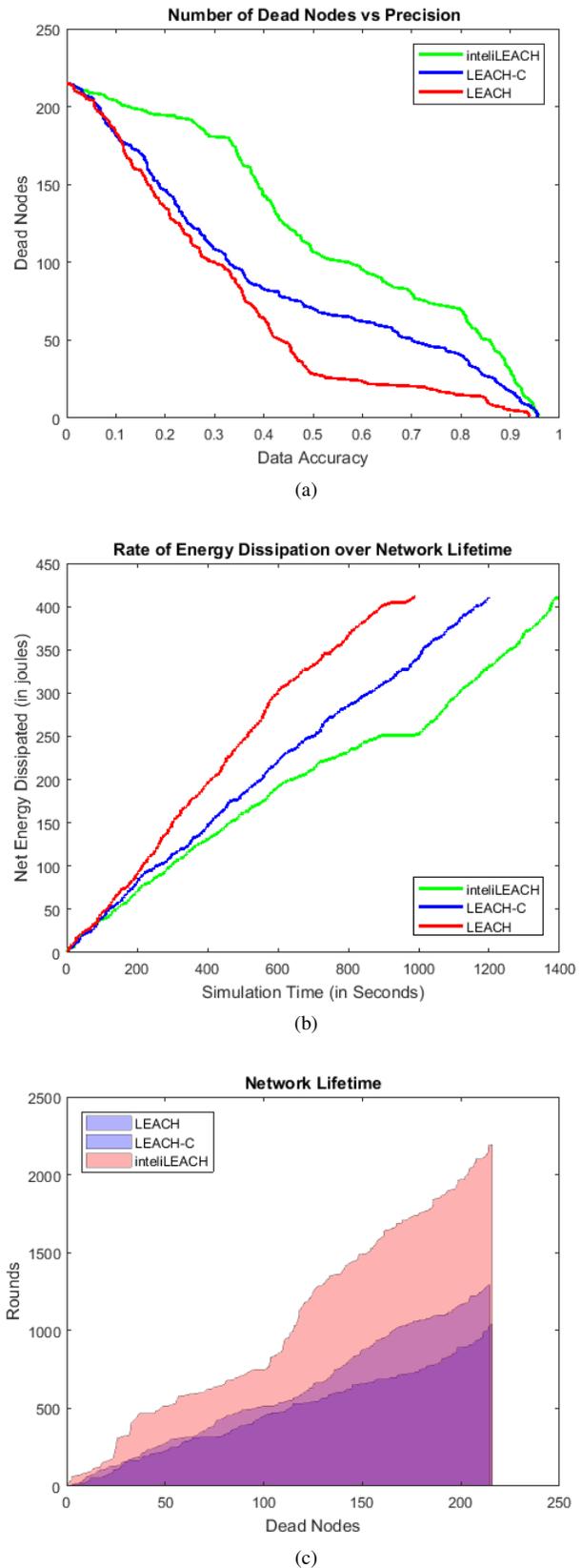


Figure 4: Performance Analysis of *intel-LEACH*

where  $S_m$  is the set of nodes in cluster  $m$ ,  $x_i$  and  $y_i$  indicate the respective coordinates of a sensor node  $n_i$ . Thereafter, we select a subset of candidate CHs from  $S_m$  which are within  $\psi$  distance and satisfied data accuracy range parameter for further processing. The rest of the procedure is executed by following steps defined in flow diagram Fig. 3 which generated simulation results as illustrated in Figures. 4a, 4b and 4c. The plot Figures. 4a depicts data precision metric for *intel-LEACH* vs LEACH and LEACH-C. Fig. 4a shows that our proposed model outperforms LEACH and LEACH-C in terms of data precision metric. Fig. 4b is the result of comparison of our algorithm with other benchmarks for predicting the rate of energy dissipation over specified time. This clearly demonstrates superior energy conservation of the given WSN by following our framework. The final plot in Fig. 4c illustrates improved longevity of the network when compared to the aforementioned benchmark protocols. Therefore these plots conclusively affirms the advantages of employing *intel-LEACH* for WSN in terms of energy conservation, network lifetime enhancement and data precision for reliable communication.

#### V. CONCLUSION

We proposed an optimized protocol *intel-LEACH* for selecting high performing nodes as CHs among randomly deployed sensor nodes based on bi-dimensional knapsack strategy. Our proposed model selects a node that can guarantee superior data precision and has maximal residual energy to serve as a CH which results in improved longevity of the WSN.

#### REFERENCES

- [1] P. Rawat, K. Deep Singh, H. Chaouchi, and J.-M. Bonnin, "Wireless sensor networks: A survey on recent developments and potential synergies," vol. 68, 04 2013.
- [2] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networking," *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 2–16, Feb. 2003. [Online]. Available: <http://dx.doi.org/10.1109/TNET.2002.808417>
- [3] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *Trans. Wireless. Comm.*, vol. 1, no. 4, pp. 660–670, Oct. 2002. [Online]. Available: <http://dx.doi.org/10.1109/TWC.2002.804190>
- [4] A. Munir, J. Antoon, and A. Gordon-Ross, "Modeling and analysis of fault detection and fault tolerance in wireless sensor networks," *ACM Trans. Embed. Comput. Syst.*, vol. 14, no. 1, pp. 3:1–3:43, Jan. 2015. [Online]. Available: <http://doi.acm.org/10.1145/2680538>
- [5] V. Mhatre and C. Rosenberg, "Homogeneous vs. heterogeneous clustered sensor networks: A comparative study." In Proceedings of 2004 IEEE International Conference on Communications (ICC 2004, 2004, pp. 3646–3651.
- [6] A. Zeb, A. K. M. M. Islam, M. Zareei, I. A. Mamoon, N. Mansoor, S. Baharun, Y. Katayama, and S. Komaki, "Clustering analysis in wireless sensor networks: The ambit of performance metrics and schemes taxonomy," *International Journal of Distributed Sensor Networks*, vol. 12, no. 7, pp. 497–499, 2016.
- [7] V. Pal, G. Singh, R. P. Yadav, and P. Pal, "Energy efficient clustering scheme for wireless sensor networks: A survey," *Journal of Wireless Networking and Communications*, vol. 2, no. 6, pp. 168–174, 2013.
- [8] K. Ramesh and K. Somasundaram, "A comparative study of clusterhead selection algorithms in wireless sensor networks," *CoRR*, vol. abs/1205.1673, 2012.
- [9] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks." IEEE, 2000, pp. 3005–3014.
- [10] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [11] M. Shahrokhzadeh, A. T. Haghighat, F. Mahmoudi, and B. Shahrokhzadeh, "A heuristic method for wireless sensor network localization," *Proceeding in Computer Science*, vol. 5, no. Supplement C, pp. 812 – 819, 2011, the 2nd International Conference on Ambient Systems, Networks and Technologies (ANT-2011) / The 8th International Conference on Mobile Web Information Systems (MobiWIS 2011).
- [12] F. Busetti, "Simulated annealing overview," 2003.
- [13] A. Arenas, A. Fernandez, and S. Gomez, "Analysis of the structure of complex networks at different resolution levels," 2008.
- [14] —, "Community detection toolbox," 2014.
- [15] C. Wilbaut, S. Hanafi, and S. Salhi, "A survey of effective heuristics and their application to a variety of knapsack problems," vol. 19, pp. 227–244, 03 2007.
- [16] A. Fréville and G. Plateau, "The 0-1 bidimensional knapsack problem: Toward an efficient high-level primitive tool," *Journal of Heuristics*, vol. 2, no. 2, pp. 147–167, Sep 1996.
- [17] S. Gómez, P. Jensen, and A. Arenas, "Analysis of community structure in networks of correlated data," *Phys. Rev. E*, vol. 80, p. 016114, Jul 2009. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevE.80.016114>
- [18] J. Karjee and H. S. Jamadagni, *Optimal Node Selection Using Estimated Data Accuracy Model in Wireless Sensor Networks*. New York, NY: Springer New York, 2013, pp. 195–205.
- [19] V. D. Oliveira, B. Kedem, and D. A. Short, "Bayesian prediction of transformed gaussian random fields," *Journal of the American Statistical Association*, vol. 92, no. 440, pp. 1422–1433, 1997.
- [20] C. W. S. Data, "California irrigation management information system (cimis)," 2017. [Online]. Available: <https://data.eol.ucar.edu/dataset/92.027>

# Enhancing Security in the Cloud: when Traceability meets Access Control

Clara Bertolissi, Omar Boucelma  
Aix-Marseille Univ, CNRS,  
Marseille, France,

Worachet Uttha  
Nakhon Pathom Rajabhat University,  
Thailand

**Abstract**—Cloud Computing technology is gaining momentum, however security concerns remain one of the top barrier to cloud projects. We propose a framework that ensures data control and privacy in the cloud by using traceability (aka Provenance) combined with expressive access control policies based on user categorization

**Keywords**-*provenance; data security; access control; cloud.*

## I. INTRODUCTION

In a cloud environment, provenance may bring an added value to cloud providers. Simply stated, provenance consists in recording entities and activities involved in producing or transforming an object. Provenance may help answering questions such as: Who created this data? What was the process used to create it? When was it modified and by whom? Behind these questions, security issues such as data integrity, privacy, access control arise. In a distributed context, components/activities used during each step can locate on different sites each applying specific management policies. Because access control models and policies are the most known approaches to enforce protection on data and resources in a system, combining provenance with access controls may lead to an efficient system for enforcing trust in the cloud.

## II. FRAMEWORK DESCRIPTION

We propose a solution combining access control features and systems' provenance data. For defining provenance, we adopt the PROV Data Model (PROV-DM [7]), a W3C Recommendation for provenance expression. It allows to represent objects and their dependencies as a directed acyclic graph composed of three vertices or object types (entity, activity, agent) and several types of edges, representing dependencies such as an activity used an entity, an entity was generated by an activity, etc (see Fig.1 for an example). For defining access policies, in the aim of defining a framework as general as possible, we have chosen to adopt the CBAC metamodel[2] which has been shown to be expressive enough to accommodate a range of different access control models. We provide a set of rules that are checked each time a process (e.g.; data alteration) is invoked in a given system. These rules use provenance data and restrictions/authorization of users depending on the category they belong to. A category is a notion of grouping based on the attributes a user owns. We also consider dynamic categorization: the memberships of a user to a certain category (and thus her/his privileges), may be affected and change dynamically as a

consequence of her/his actions. For instance, privileges can be automatically revoked depending on the number of refusals of (a specific) activity execution a user has received.

More precisely, in our framework, entities are denoted by constants in a many sorted domain including: a set agents, a set of named atomic activities, a set of entity identifiers, a set of categories. The core axiom of the model is as follows:

$$\text{Belongs\_to}(\text{agent}, \text{categ}, \text{entity}) \wedge \text{permission}(\text{categ}, \text{activity}, \text{entity}) \Leftrightarrow \text{allow}(\text{agent}, \text{activity}, \text{entity}).$$

The idea is that the first relation, *belongs to*, specifies the "qualification" of the agent with respect to the attributes she/he has and the past (relevant) actions she/he has accomplished. The relation *permissions* specifies whether the requested action can be performed against the object according to a certain level of qualification, i.e. a category. Notice that this relation is not dependent from the agent. It is used to model action validations by the system and it does not directly make use of user access privileges. This separation in an agent-dependent relation and a system-dependent relation eases the updates and maintenance operations of the policy. If the *agent* has a sufficient qualification level for belonging to a *category* to which the *activity* is permitted on the requested *entity*, then the access request *allow(agent,activity,entity)* is granted.

## III. RELATED RESEARCH

Several research works attempted to mix traceability with access control. Park et al [9] proposed PBAC, a model where the notion of object dependency lists, derived from process execution traces, are used for access requests evaluation. In [10], provenance-aware access control policies are discussed. An abstract provenance model TPM (Type Provenance Model) is proposed. TPM allows the expression of complex dependencies using regular expressions in a similar way as it is done in PBAC with object dependency lists. In [1], a provenance model (cProv) and a policy language are proposed. Their rules are generic and some of them are activity-oriented, very close to the an RBAC model, which can be expressed in CBAC. In [6] authors propose a formal model that assigns a trust degree (evaluated from provenance data) to each entity. The approach we propose provides strict authorization rules, instead of trust values granted to a particular type of activity per user. From our perspective, one the main ideas of our work is to come up with a tight integration of PROV and CBAC. The concept of category of users (or agents) has no direct correspondence in the PROV-DM

model, even with extended structures such as collections, which apply only to entities. The membership of an agent to a group can only be expressed in the past since a PROV document reflects only past and final facts. Even in PROV-O [5], the PROV Ontology, no relationships for the belonging of an agent to a group exist.

IV. FRAMEWORK VALIDATION ON AN EXAMPLE

We choose Datalog programming for evaluation and testing of our model, since provenance data in PROV-N notation [8] can be easily expressed in Datalog and declarative policy rules are very similar to Datalog rules (the full Datalog specification is available at <http://home.npru.ac.th/wuttha/research/PROV-C>).

In our example, we consider an academic course. The teacher asks students to submit 2 tasks via the University web application. Each task will be integrated incrementally by the web application on one student-specific file so that the teacher can view all one’s work in one file. We assume that students can submit a task only after they have been graded on the previous one. Students cannot re-submit a task once the teacher has rated it. As all subject belonging to the same category will have the same privileges, in order to distinguish students having already uploaded some tasks, we may refine the category student by adding two categories: *uploadedT1*, *uploadedT2*. These categories are used for grouping students who have already uploaded the task 1 or 2, respectively, and may be defined as:

```
belong_to(Agent,uploadedT1,Entity) :-
    attribute(Agent, registered_student),
    wasGeneratedBy(Entity, uploadTask1),
    wasAssociatedWith(uploadTask1, Agent).
```

This rule means that agents qualified as registered students (as recorded in the University database) and having executed the action *uploadTask1* on the resource entity (as recorded in the dependency path of the corresponding provenance graph) are assigned to the category *uploadedT1*. In order to have information such as "how many times an entity was used", we may need to perform a computation on the provenance graph for calculating how many dependencies of each type an entity has. These will produce new predicates in the Datalog program.

```
permission(student, uploadTask1, Entity) :-
    wasUsedBy_count(Entity, rateTask1, 0).
```

The rule above means that a student is allowed to re-upload the same assignment if it has not been rated by the teacher. Let us consider the scenario reported in Fig.1 with two students, Bob and Mike. By playing access queries, we can simulate the system and verify the correctness of the access response. The termination of Datalog queries is guaranteed, see [3]. If we want to know whether Bob is allowed to upload his task 1 or not, we can ask to the Datalog engine the query

```
?- allow(mike, Activity, taskMike)
false.
```

We can also ask more general questions by including variables in the query:

```
%% The only activity Mike can perform is upload task2
?- allow(mike, Activity, taskMike).
Activity = uploadTask2.
```

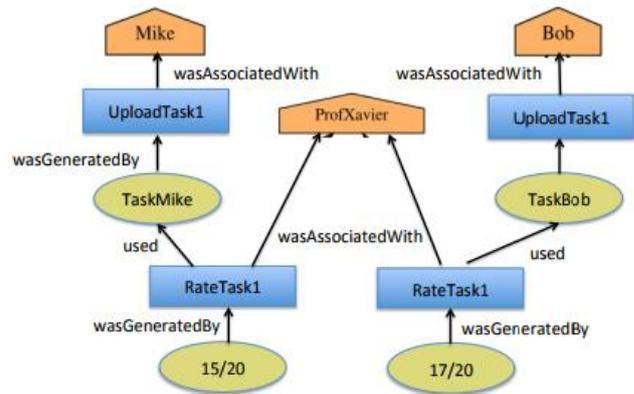


Figure 1. Example: Excerpt of the Provenance Graph

```
%% Mike is the only one who can upload his task2.
?- allow(Agent, uploadTask2, taskMike).
Agent = mike.

%% Mike is allowed to upload only his task.
?- allow(mike, uploadTask2, Entity).
Entity = taskMike.
```

This kind of simulation can help the policy designer or administrator to detect possible inconsistencies in the policy definition. Assuming the translation to Datalog is correct, the engine will compute a positive answer only if Bob is allowed to perform the access according to the specified policy.

V. CONCLUSIONE

We have described a framework that ensures trust in a distributed computing environment by means of data security enforcement. The framework combines Provenance and Category-based Access Control policies. We already have an implementation that accommodates RBAC policies [4] and we are planning to extend it with CBAC.

REFERENCES

- [1] M. Ali and L. Moreau. A Provenance-Aware Policy Language (cProv) and a Data Traceability Model (cProv) for the Cloud. In In Proc. Of GreenCom’13, p. 479–486, IEEE, 2013.
- [2] Steve Barker. The Next 700 Access Control Models or a Unifying Meta-model? In Proc. of SACMAT ’09, p. 187–196, 2009. ACM.
- [3] W. Chen and D. S. Warren. Tabled Evaluation with Delaying for General Logic Programs. J. ACM, 43(1):20–74, 1996.
- [4] J. Lacroix and O. Boucelma. Design and Implementation of a Trust Service for the Cloud. In Proc. of OTM ’15, p. 620–638. LNCS, 2015.
- [5] T. Lebo, S. Sahoo, D. McGuinness. PROV-O: The PROV Ontology, 2013.
- [6] G. Lin, Y. Bie, and M. Lei. Trust Based Access Control Policy in Multi-domain of Cloud Computing. J. of Computers, 8(5), 2013.
- [7] L. Moreau and P. Missier. PROV-DM: The PROV Data Model, 2013.
- [8] L. Moreau and P. Missier. PROV-N: The Provenance Notation, 2013.
- [9] J. Park, D. Nguyen, and R. Sandhu. A provenance-based access control model. In Proc. of PST’12, p. 137–144, IEEE, 2012.
- [10] L. Sun, J. Park, and R. Sandhu. Engineering Access Control Policies for Provenance-aware Systems. In Proc. of CODASPY ’13, p. 285–292, 2013. ACM.

# Database Multi-factor Authentication via Pluggable Authentication Modules

Cameron Hamilton  
Master's Degree Candidate  
The Citadel  
Charleston, SC  
chamilt4@citadel.edu

Aspen Olmstead  
Assistant Professor and Graduate Program Director  
College of Charleston  
Charleston, SC  
olmsteda@cofc.edu

**Abstract**— Authentication schemes containing single factors are becoming increasingly inadequate for many applications. This paper describes the use of Pluggable Authentication Modules as a means of creating a multi-factor authentication scheme for a MySQL database. The resulting scheme is one with a significantly higher degree of security.

## I. INTRODUCTION

A factor of authentication is a piece of information which proves the authenticity of a user. Combining multiple authentication factors together can have a profound effect on the security of a system. There are three principal classes of factors: knowledge factors, such as a password; possession factors, such as a smart card; and inheritance factors, such as a fingerprint [1]. These factors can be used in combination to produce multi-factor authentication.

The focus of this paper is to implement a secure multi-factor authentication mechanism for MySQL Server. This will be accomplished with the Pluggable Authentication Module (PAM) plugin recently available in newer versions of MySQL. PAM enables the creation of custom authentication schemes by allowing the mixing and matching of various authentication mechanisms. This scheme will consist of both knowledge and possession factors. Ideally, inheritance factors would also be considered. However, due to the relative scarcity of biometric-related Pluggable Authentication Modules, this category will be disregarded.

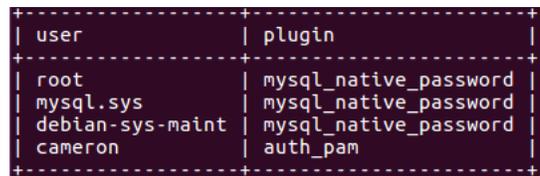
## II. RELATED WORK

Since Pluggable Authentication Modules provide a common authentication scheme, they are quite versatile and are capable of supporting authentication for an array of applications. Samar, V. demonstrated the implementation of user mapping in PAM for SSH. This was similar to the knowledge factor of our authentication scheme. Cristofaro, E et al. provided a usability study of two-factor authentication and cited Google Authenticator as a popular option. Archlinux demonstrated an implementation of Google Authenticator with PAM to support an SSH service. Our authentication scheme combines similar approaches used by the aforementioned sources to enable multi-factor authentication for a MySQL database.

## III. IMPLEMENTATION & RESULTS

The PAM plugin was installed using an open source implementation available from Percona. This plugin allows for external PAM authentication within MySQL Server [2]. One of the files resulting from this installation was a shared object library file named “auth\_pam.so.” This plugin file was loaded into MySQL by adding a reference to it within the mysqld configuration, and then MySQL Server was restarted.

Next, a user named “cameron” was created within the database with “auth\_pam” as the specified authentication plugin instead of the default “mysql\_native\_password,” as shown in Fig. 1. At this point, no password was specified since all authentication is handled externally.



user	plugin
root	mysql_native_password
mysql.sys	mysql_native_password
debian-sys-maint	mysql_native_password
cameron	auth_pam

Figure 1. MySQL users and authentication plugins

Next, a PAM configuration was created within the /etc/pam.d directory named “mysqld.” This is the file that the PAM plugin within MySQL Server will reference when performing authentication on the “cameron” user [3]. This file contains a series of directives which control how authentication is handled, which accounts are allowed access, and how sessions are managed. This file is what allows a high degree of freedom in designing our authentication scheme.

Within the “mysqld” file, directives were added to enable user mapping via the “pam\_unix” module, as shown in Fig. 2. Now, MySQL will verify that an OS user exists under the same name as the database user and that the password supplied matches the OS user. This type of authentication is similar to OS authentication within Oracle Database. With this directive in place, the first authentication factor of our scheme was implemented, satisfying the knowledge component of the scheme.

```
auth    required    pam_warn.so
auth    required    pam_unix.so audit
account required    pam_unix.so audit
```

Figure 2. “mysqld” PAM configuration file with user mapping directives added

Valid and invalid logins were issued to MySQL using the Linux password for the “cameron” user. We used both valid and invalid logins to ensure the authentication worked properly. Everything worked as expected.

In order to implement the possession factor of our scheme, Google Authenticator was chosen. Google Authenticator utilizes a time-based one-time password algorithm for authenticating users from a mobile device [4]. This password is in the form of a six-digit number generated algorithmically every thirty seconds. One advantage to using this mechanism is that no Internet connection is required since the code is time-based and is generated algorithmically on both the server and client.

The Google Authenticator package was installed on the Linux server which hosts MySQL Server. After the installation, the “google-authenticator” program was executed to create an authentication file after a series of prompts. During the execution of this program, a QR code was generated. This code was used by scanning it with the Google Authenticator mobile app which was installed on an Android mobile device. After syncing, the app displayed the six digit password along with a thirty-second expiration timer.

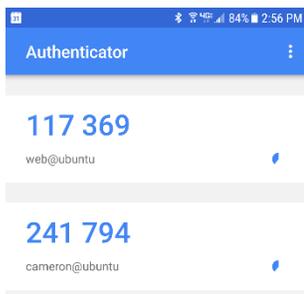


Figure 3. Google Authenticator Android app

With Google Authenticator installed, various parameters were specified to increase the security of the scheme. For instance, multiple users of the same authentication token were disabled to restrict logins to one every thirty seconds. This can help increase the chance of noticing and preventing man-in-the-middle attacks. Also, rate-limiting was enabled to restrict the number of login attempts to three every thirty seconds [4].

Next, the “pam\_google\_authenticator.so” entry was added to the mysqld PAM configuration, as shown in Fig. 4.

```
auth    required    pam_google_authenticator.so
auth    required    pam_warn.so
auth    required    pam_unix.so audit
account required    pam_unix.so audit
```

Figure 4. “mysqld” PAM configuration file with the Google Authenticator directive added

With this new entry added, MySQL Server will now attempt to authenticate first with “pam\_google\_authenticator” module, and then with the “pam\_unix” user mapping module. When attempting to log into MySQL, two password prompts are displayed. This is shown in Fig. 5.

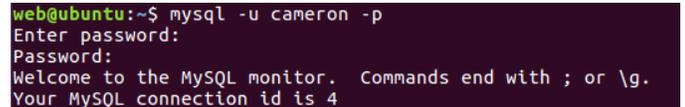


Figure 5. MySQL login attempt using both a Google Authenticator token and a user mapping password

As shown in Fig. 5, after providing a valid Google token and user mapped password, the authentication was successful.

#### IV. CONCLUSION

Pluggable Authentication Modules used within the context of MySQL Server offer a very flexible platform for implementing a custom database multi-factor authentication scheme. In our particular implementation, we were able to demonstrate both knowledge and possession factors via the use of user mapping and Google Authenticator. Combining the two factors enhanced the security of our system.

#### REFERENCES

- [1] M. Rouse, "Multifactor Authentication (MFA)," SearchSecurity, 1 03 2015. [Online]. Available: <http://searchsecurity.techtarget.com/definition/multifactor-authentication-MFA>. [Accessed 13 06 2017].
- [2] "The MySQL Plugin API," MySQL, [Online]. Available: <https://dev.mysql.com/doc/refman/5.7/en/plugin-api.html>. [Accessed 10 06 2017].
- [3] Samar, V. (1996). Unified login with pluggable authentication modules (PAM). [online] ACM Digital Library. Available at: <http://dl.acm.org/citation.cfm?id=238177> [Accessed 13 Jun. 2017].
- [4] archlinux, "Google Authenticator," 26 04 2017. [Online]. Available: [https://wiki.archlinux.org/index.php/Google\\_Authenticator](https://wiki.archlinux.org/index.php/Google_Authenticator). [Accessed 13 06 2017].
- [5] Cristofaro, E., Du, H., Freudiger, J. and Norcie, G. (2013). A Comparative Usability Study of Two-Factor Authentication. arXiv. [online] Available at: <https://arxiv.org/abs/1309.5344> [Accessed 14 Jun. 2017].

## **Session 15: Internet Applications and Technology**

Title: Improve CRUD performance on hierarchical data Nested Interval model vs. nested set model  
(Authors: Blake Badders, Aspen Olmsted)

Title: Efficient Hardware Implementation of Itubee For Lightweight Application  
(Authors: Juhua Liu, Wei Li, Guoqiang Bai)

Title: Weighing the shopping benefits of a smarter refrigerator  
(Authors: Stephen Goeddel, Pasha Sadeghian, Aspen Olmsted)

Title: Optimizing Synchronization of Cloud Storage Services: Combining Benchmark Monitoring and Learning-Based Framework  
(Authors: Preston T. Owens, Aspen Olmsted)

# Improve CRUD performance on hierarchical data

## Nested Interval model vs. nested set model

Blake Badders, Aspen Olmsted

Department of Computer Science

College of Charleston

Charleston, SC 29401

blake.badders@gmail.com, olmsteda@cofc.edu

**Abstract**— Hierarchical data models are commonly used in databases to represent data across a multitude of disciplines. In this paper, I will examine two models that can be used to represent hierarchical data in a relational database, the nested interval model, and the nested set model. These models were analyzed on a WAMP server using various queries and CRUD operations. The purpose of this paper is to examine the advantages of using a nested interval model as compared to a nested set model in a database which needs to search quickly, create, update, and destroy large amounts of hierarchical data.

**Keywords**- *hierarchical data; nested set model; nested interval model.*

### I. INTRODUCTION

Storing data using a hierarchical structure is a common use case for developers working with relational databases. Organizations, forums, sub-types, bill of materials, and content management categories all lend themselves to being stored in a hierarchy. However, the tables of a relational database can provide difficulties to producing efficient hierarchical data structures.

In recent years, non-relational databases have been developed and increased in popularity. Due to the flexible and semi-structured nature of these databases, they are increasingly being used to represent hierarchical data instead of relational databases. As a result, the need for operation optimization in relational databases when searching, creating, and deleting hierarchical data has sharply increased.

The organization of the paper is as follows. Section II describes the related work and the implementations of current methods. In Section III we give a motivating example where our application is useful. Section IV details the underlying framework, as well as the advantages of the design and the feature set, provided compared to a user running the same processes through an R shell. Section V provides a walkthrough of the application using an example dataset. We conclude and discuss future work in Section VI.

### II. RELATED WORK

A common method to increase query speed is the nested set model shown in Fig. 1. Each node has two attributes, left and right. The process starts with 1 at the left of the root node. Values are assigned to children nodes depth first. The result of this numbering is that all left and right values of children of a node fall within the left and right of the node itself.

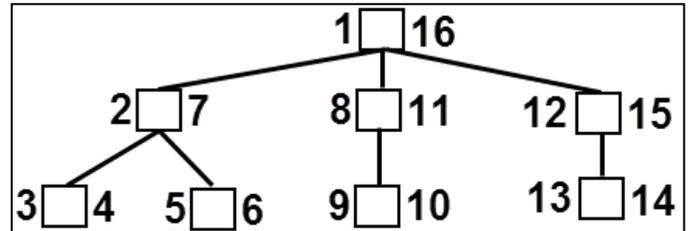


Figure 1. Nested Set Model

Gyorodi et al. [1] published a performance analysis between the nested set model and another hierarchical data structure, the adjacency list model, explained in detail in their paper. Their results showed that the performance of queries retrieving children from the parent node was three orders of magnitude faster for the nested model. This allowed me to focus specifically on the nested set model and nested interval model instead of including the adjacency list model.

Kriegel, Pötke, and Seidl [2] proposed an implementation for a nested interval model. Instead of using left and right integer attributes on nodes to aid query performance, ratios determine the left and right values. The distance between a child nodes left, and the right attribute is one-half the interval of the parent that does not have children. These ratios are created using Farey fractions. While the algorithm defined in this paper was not utilized for the performance analysis, the idea of splitting parent intervals in half to nest intervals guided the creation of the interval creation algorithm.

### III. HYPOTHESIS

My hypothesis is that due to the independence of the left and right attributes of each node in the nested interval model, create, update, and destroy methods performed on nodes in the nested interval model will execute faster than those on nodes of the nested set model. In order to test this hypothesis, I will build two data structures; one using the nested set model, the other using the nested interval model. Create and destroy operations will be performed on each model and the time required to complete each operation will be compared. The expected result is that the CRUD operations on the nested interval model will be at least one order of magnitude faster than those of the nested set model.

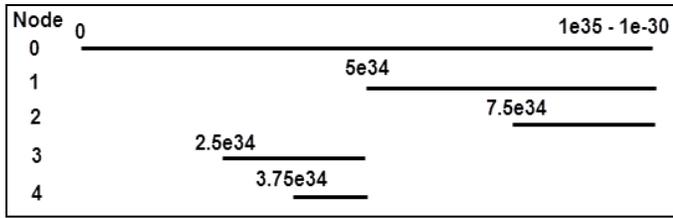


Figure 2. Nested Interval Model

#### IV. IMPLEMENTATION

To create an environment that allowed for accurate performance analysis, a WAMP server was installed on a development laptop. Two tables were created in MySQL; one using the nested set hierarchy described in the related work, the other using a nested interval hierarchy. Each tree was composed of 5,500 nodes.

In the nested interval tree nodes were decimal data types decimal with precision 65 and scale 30 as the left and right indexes. The root node was given a left value of 0 and right value containing the maximum value for the specified decimal data type. When a node is added to a parent, the interval for that node is half the interval of the parent that does not already have a child. A visualization can be seen in Fig. 2 and the code is in Algorithm 1.

For insertion and deletion on each table, stored procedures were implemented to ensure that the right and left attributes on nodes remained correct. When any node is inserted or deleted in the nested set model, all nodes located to the right of the node need to be updated as a result. In the nested interval model procedures, the left and right attributes of the parent node and left value of most recent child must be retrieved to calculate the interval. There is no need to modify left and right

##### Algorithm 1 Create Nested Interval Node

```

DECLARE parent_left_val, parent_right_val, child_left_val,
new_right_val, new_left_val decimal(65,30);
DECLARE small_num decimal(65,30) default 1E-30;
DECLARE two decimal(65,30) default 2;

SELECT leftVal, rightVal FROM nested_interval WHERE id =
parentId INTO parent_left_val, parent_right_val;
SELECT min(leftVal) FROM nested_interval WHERE leftVal >
parent_left_val INTO child_left_val;

IF child_left_val is null THEN
    SET new_right_val = parent_right_val - small_num;
    SET new_left_val = parent_left_val + (parent_right_val -
parent_left_val) / two;
ELSE
    SET new_right_val = child_left_val - small_num;
    SET new_left_val = parent_left_val + (child_left_val -
parent_left_val) / two;
END IF;

INSERT INTO nested_interval(id,leftVal,rightVal)
values(nodeId,new_left_val,new_right_val);
    
```

TABLE I. OPERATION RUN TIME IN SECONDS

	Nested Set			Nested Interval
	Worst Case	Average	Best Case	
Insert	.1145	.0696	.0025	.0020
Delete	.1135	.0643	.0043	.0009
Query	-	.0086	-	.0129

attributes on any other node as a result of the new node.

Since location within the tree of the created and deleted node would determine the speed of the operation in the nested set model, three locations were used. The furthest left leaf node was used as the worst case since all other nodes in the tree would need to be modified. The middle leaf node was used as the average case, and the furthest right leaf node was used as the best case. For the nested interval model, the location of the node would not affect the speed of the operations, so the furthest left leaf node was used for analysis.

MySQL profiling was used to track the run time for each operation. The duration of all queries related to the creation or deletion of the node were summed to calculate the total operation time.

#### V. RESULTS

In TABLE I we can see the average run times for creation and deletion for the previously defined cases. We notice that even the best case deletion and insertion for the nested set model cannot match the execution time of the nested interval model. The nested interval performance is an order of magnitude faster than the average case for the nested set in both insertion and deletion.

However, this comes at a minor cost in query performance. When listing all children of the root node (all nodes in the tree), the nested set model is able to perform the operation twice as fast as the nested interval model.

#### VI. CONCLUSIONS AND FUTURE WORK

The nested interval model brings significant CRUD optimization to hierarchical data structures in relational databases. As big data grows and needs to meet the needs of businesses that expect real-time data updates, the ability to efficiently update nodes in a tree is critical.

I hope in future research to compare the algorithm used to create this nested interval model with the relational interval tree model mentioned by Kriegel, Pötke, and Seidl [2] and the dyadic model mentioned by Tropashko [3].

#### REFERENCES

- [1] C. Gyrod, R.-R. Moldovan-Duse, R. Gyrod and G. Pecherle, "Improve Query Performance On Hierarchical Data. Adjacency List Model Vs. Nested Set Model," *International Journal of Advanced Computer Science and Applications*, pp. 1-7, 2016.
- [2] H.-P. Kriegel, M. Pötke and T. Seidl, "Managing Intervals Efficiently in Object-Relational Databases," in *VLDB '00 Proceedings of the 26th International Conference on Very Large Data Bases*, San Francisco, 2000.
- [3] V. Tropashko, "Nested intervals tree encoding in SQL," *ACM SIGMOD Record*, vol. 34, no. 2, pp. 47-52, 2005.

# Efficient Hardware Implementation of Itubee For Lightweight Application

Juhua Liu

Institute of Microelectronics, Tsinghua University,  
Beijing, China  
liujh14@mails.tsinghua.edu.cn

Wei Li

Institute of Microelectronics, Tsinghua University,  
Beijing, China  
l-w16@mails.tsinghua.edu.cn

Guoqiang Bai

Institute of Microelectronics, Tsinghua University,  
Beijing, China  
baigq@tsinghua.edu.cn

**Abstract**—Recently, a new lightweight block cryptography algorithm, ITUbee, has been proposed by Ferhat Karakoc in Lightsec 2013. An efficient hardware implementation of ITUbee is presented in this paper. Firstly, we reuse certain module, which takes a big share of hardware resource, to achieve better resource utilization. Secondly, we apply composite field to implement 8-bit S-box instead of the traditional looking up tables(LUTs) to save area requirements. In the end, we conclude that the hardware implementation of ITUbee requires about 6448 GE on 0.18 um technology. The area consumption of ITUbee is roughly 31.2% less than the round-based implementation. And it costs 365.6 GE to implement 8-bit S-box by using composite field, 32.7% less than by using LUTs.

**Keywords**—lightweight; Itubee; hardware implementation; reuse; composite field; area-optimization.

## I. INTRODUCTION

Nowadays, the increasing applications such like RFID tags and intelligent devices spur us to develop an efficient cryptography algorithm, which meets the security and privacy requirements and can be applied to resource constrained devices at the same time. For that reason, designing lightweight primitives is getting prominent. Block ciphers play an essential role in cryptography applications so that a considerable number of lightweight block ciphers have been proposed. DESXL[1], PRINCE[2], SEA[3] and KATAN[4], for example.

Ferhat Karakoc et al. proposed a new software oriented lightweight block cipher, ITUbee, for resource constrained devices that include a microcontroller and have a limited battery power such as sensor nodes in wireless sensor networks[5].ITUbee is designed based on a Feistle structure while having no key schedule, which may make ITUbee subjected to related key attacks as observed in GOST cipher[6].The author came up with a new approach that the round key was injected between two nonlinear operations to mend this weakness.

To evaluate the performance of ITUbee we have implemented the algorithm on hardware and gave the result of Design Compiler. Especially, to reduce the energy

consumption of the cipher we applied the S-box based on composite field to our design. Note that there are two F functions in each round of encryption, which accounts for more than 90% of the total area. Fortunately, this proportion can be reduced dramatically by reusing the F functions, improving its efficiency in terms of energy consumption.

The rest of the paper is organized as follows. In Section 2, we give the compact algorithm description of ITUbee. In Section 3, some details of design rationale of S-Box based on composite field is shown. We give the hardware architecture of our implementation in Section 4. In Section 5, we give the simulation details and results of implementation. We conclude the paper with Section 6.

## II. DESCRIPTION OF ITUBEEE

### A. Notation

Before we start the describing, giving the uniform notations throughout this paper makes reading much easier.

$\parallel$  : Concatenation operator.

$K_R$  : The right half of the master key.

$K_L$  : The left half of the master key.

$P_R$  : The right half of the plaintext.

$P_L$  : The left half of the plaintext.

$P$  : 80-bit plaintext.

$C_R$  : The right half of the ciphertext.

$C_L$  : The left half of the ciphertext.

$C$  : 80-bit ciphertext.

$RC_i$  : The round constant in the  $i$ -th round.

### B. Algorithm Description

ITUbee algorithm accepts the inputs  $P_L, P_R, K_L, K_R, (RC_1, RC_2, \dots, RC_{20})$ , and outputs the ciphertext  $C_L, C_R$ . ITUbee algorithm is designed with a Feistle structure with 80-bit key length and block size, consisting of 20 rounds overall and having key whitening layers at the first and the last round as illustrated in Fig.1

In each round, there are one  $L$  function, two  $F$  functions and XOR operators, the execution order of these operators is shown in figure 1. The definitions of these functions are:  $F(X) = S(L(S(X)))$ ,  $S(a || b || c || d || e) = s[a] || s[b] || s[c] || s[d] || s[e]$ ,  $L(a || b || c || d || e) = (e \oplus a \oplus b) || (a \oplus b \oplus c) || (b \oplus c \oplus d) || (c \oplus d \oplus e) || (d \oplus e \oplus a)$ , where  $a, b, c, d, e$  are 8-bit values and  $s$  is the S-box used in advanced encryption standard (AES)[5]. The constant  $RC_i$  in each round is given in Table I. Note that 16-bit round constant  $RC_i$  is XORed with the rightmost 16 bits in each round.

Table I. Round constants used in ITUbee algorithm

i	$RC_i$	i	$RC_i$	i	$RC_i$	i	$RC_i$
1	1428	6	0f23	11	0a1e	16	0519
2	1327	7	0e22	12	091d	17	0418
3	1226	8	0d21	13	081c	18	0317
4	1125	9	0c20	14	071b	19	0216
5	1024	10	0b1f	15	061a	20	0115

$(K_L || K_R)$  and  $(K_R || K_L)$  are used as whitening keys at the first and the last round of the encryption algorithm respectively and for even rounds  $K_L$  is used while for odd rounds  $K_R$  is used. Both of the round keys and whitening keys are derived from the master key directly. The decryption process of ITUbee is the same as the encryption process, while the only difference is that the round keys and constants are used in reversed order[5].

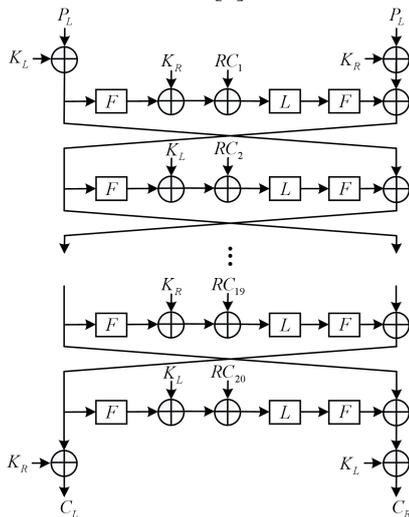


Figure 1. ITUbee encryption algorithm

### III. IMPLEMENTATION OF S-BOX USING NORMAL BASIS IN COMPOSITE FIELD

To the best of our knowledge, the efficiency of the ITUbee depends on the implementation of S-box involved in F function in each round. The operation of nonlinear multiplication inversion makes S-box the most computational

intensive in ITUbee algorithm. There are two mainly approaches in public literature to implement the S-box: using a look up table(LUT) or using a Composite Filed algorithm. Compared with the approach using LUT, the implementation of S-box using composite filed can save area consumption dramatically by performing the 8-bit Galois field inversion of the S-box using subfield of 4 bits and of 2 bits.

Generally, the S-box function with input  $a$  is defined by two steps: the *multiplicative inverse* in  $GF(2^8)$  (see Eq.(1)) and *affine transformation*(see Eq.(2)):

$$c = a^{-1}, \quad (\text{if } a = 0, \text{ then } c = 0). \quad (1)$$

$$s = Mc \oplus b, \quad (2)$$

Where  $M$  is a constant bit matrix and  $b$  is a byte vector shown below:

$$\begin{pmatrix} s_7 \\ s_6 \\ s_5 \\ s_4 \\ s_3 \\ s_2 \\ s_1 \\ s_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_7 \\ c_6 \\ c_5 \\ c_4 \\ c_3 \\ c_2 \\ c_1 \\ c_0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

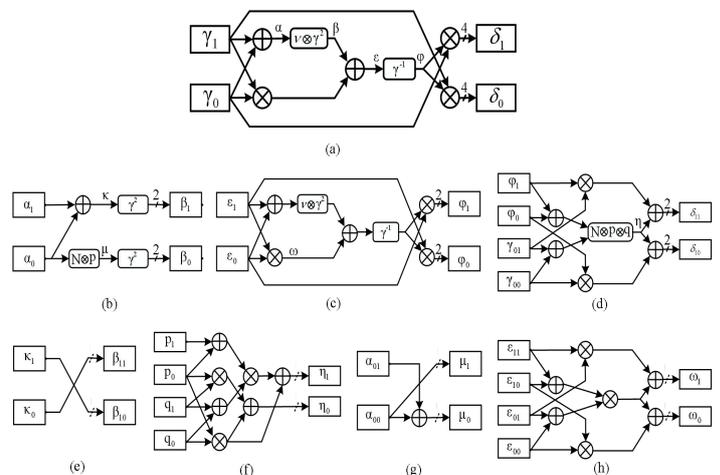


Figure 2. Hierarchical structure (a)Normal inverter in  $GF(2^8)$ :  $(\gamma_1 Y^{16} + \gamma_0 Y)^{-1} = (\delta_1 Y^{16} + \delta_0 Y)$ , the coefficients pair  $[\gamma_1, \gamma_0]$  and  $[\delta_1, \delta_0]$  are of the same bit width, shown at the output of the figure and the same as the below; (b),(c),(d) give the structures of combined operation of squaring then scaling (multiplying), normal inverter and multiplication in  $GF(2^4)$  respectively; (e),(f),(g),(h) then give the structures of squaring (same as inverter), combining the multiplication with scaling by  $N$ , scaling by  $N$  and multiplication in  $GF(2^2)$  respectively.

Compared with the first substep (multiplicative inverse in  $GF(2^8)$ ), the second substep (affine transformation) is easier

to implement on hardware. Therefore, we will focus on the key issue of finding the inverse in  $GF(2^8)$ . As we know that S-box used in AES algorithm, which is same as the S-box used in ITUbee algorithm, is designed based on the particular Galois field of 8-bit bytes where the bits are coefficients of a polynomial and multiplication is modulo the irreducible polynomial  $q(x) = x^8 + x^4 + x^3 + x + 1$ , with addition of coefficients modulo 2[7].

Direct calculation of the inverse of a seven-degree polynomial (modulo an eight-degree polynomial using extended Euclidean algorithm) is not quite easy. But calculation of the inversion of a one-degree polynomial, modulo a two-degree polynomial, is pretty easy. Hence, we compact the finding inverse in  $GF(2^8)$  into subfield  $GF(2^4)$  then into subfield  $GF(2^2)$  and finally into subfield  $GF(2)$  using a multi-level hierarchical structure, as depicts in Fig.2. Next, we give some details of this hierarchical structure[7].

Now we represent a general element  $E$  of  $GF(2^8)$  as a linear polynomial over  $GF(2^4)$ , as  $g = \lambda_1 y + \lambda_0$ , with multiplication modulo an irreducible polynomial  $r(y)$  (see Eq.(3)), whose coefficients  $[\lambda_1, \lambda_0]$  are in the 4-bit subfield  $GF(2^4)$ . Although both of normal basis and polynomial basis can decompose the multiplicative inverse in  $GF(2^8)$  into its isomorphic subfields, the most compact case uses normal basis for all subfields. Considering this, we choose the normal basis  $[Y^{16}, Y]$  to represent the element in  $GF(2^8)$  again,  $g = \gamma_1 Y^{16} + \gamma_0 Y$ , where the  $[\gamma_1, \gamma_0] = [g_{7:4}, g_{3:0}]$ . Similarly, we get all the irreducible polynomials  $s(x), t(w)$  and their normal basis  $[X^4, X]$  and  $[W^2, W]$  respectively (see Eq.(4) and Eq.(5)). Here we have:

$$r(y) = y^2 + \tau y + \nu = (y + Y^{16})(y + Y) \tag{3}$$

$$s(x) = x^2 + Tx + N = (x + X^4)(x + X) \tag{4}$$

$$t(w) = w^2 + w + 1 = (w + W^{16})(w + W) \tag{5}$$

In Eq.(3) we define the trace  $Y$   $Y$  16 and the norm  $Y$   $Y$  16 (correspondingly  $T$   $X$   $X$  4,  $N = X \cdot X^4$  for Eq.(4) and  $W + W^2 = 1, W \cdot W^2 = 1$  for Eq.(5). The most efficient choice of trace and norm is to let the trace be unity, here we let  $\tau = T = 1$ .

In  $GF(2^8)$  with a normal basis  $[Y^{16}, Y]$ , the inverse of  $g = (\gamma_1 Y^{16} + \gamma_0 Y)$  modulo  $y^2 + \tau y + \nu$  is given by:

$$g^{-1} = (\gamma_1 Y^{16} + \gamma_0 Y)^{-1} = \delta_1 Y^{16} + \delta_0 Y = [\theta^{-1} \gamma_0] Y^{16} + [\theta^{-1} \gamma_1] Y \tag{6}$$

then we have:

$$\delta_1 = \theta^{-1} \gamma_0 = [\gamma_1 \gamma_0 \tau^2 + (\gamma_1^2 + \gamma_0^2) \nu]^{-1} \gamma_0 \tag{7}$$

$$\delta_0 = \theta^{-1} \gamma_1 = [\gamma_1 \gamma_0 \tau + (\gamma_1^2 + \gamma_0^2) \nu]^{-1} \gamma_1 \tag{8}$$

So finding the inverse of  $g$  in  $GF(2^8)$  reduces to an inverse operation and several additions and multiplications in  $GF(2^4)$ , as shown in Figure 2 (a). It is easy to handle the addition in  $GF(2^4)$  by bitwise XOR. While for the inverse

and multiplication in  $GF(2^4)$  we give their definitions and algorithms using the similar approach above.

In  $GF(2^4)$  with a normal basis  $[X^4, X]$ , the inverse of  $\gamma = (\varepsilon_1 X^4 + \varepsilon_0 X)$  modulo  $x^2 + Tx + N$  has a same formulation as Eq.(6) except the 2-bit width coefficients.

$$\gamma^{-1} = (\varepsilon_1 X^4 + \varepsilon_0 X)^{-1} = \varphi_1 X^4 + \varphi_0 X = [\Delta^{-1} \varepsilon_0] X^4 + [\Delta^{-1} \varepsilon_1] X$$

then we have:

$$\varphi_1 = \Delta^{-1} \varepsilon_0 = [\varepsilon_1 \varepsilon_0 T^2 + (\varepsilon_1^2 + \varepsilon_0^2) N]^{-1} \varepsilon_0 \tag{10}$$

$$\varphi_0 = \Delta^{-1} \varepsilon_1 = [\varepsilon_1 \varepsilon_0 T + (\varepsilon_1^2 + \varepsilon_0^2) N]^{-1} \varepsilon_1 \tag{11}$$

We can see that finding the inverse of means an inverse and several additions and multiplications in  $GF(2^2)$ , shown in Figure 2 (c). As to the multiplication in  $GF(2^4)$ ,  $\varphi \cdot \gamma_0$  is defined by (shown Figure 2 (d)):

$$(\varphi_1 X^4 + \varphi_0 X) \cdot (\gamma_{01} X^4 + \gamma_{00} X) = \delta_{11} X^4 + \delta_{10} X \tag{12}$$

Where

$$\delta_{11} = \eta \oplus (\varphi_1 \otimes \gamma_{01}) = (N \otimes (\varphi_1 \oplus \varphi_0)(\gamma_{01} \oplus \gamma_{00})) \oplus (\varphi_1 \otimes \gamma_{01}) \tag{13}$$

$$\delta_{10} = \eta \oplus (\varphi_0 \otimes \gamma_{00}) = (N \otimes (\varphi_1 \oplus \varphi_0)(\gamma_{01} \oplus \gamma_{00})) \oplus (\varphi_0 \otimes \gamma_{00}) \tag{14}$$

To compact the logic and simplify the circuit, we combine the operations of squaring and scaling by the norm  $\nu$  ( $\nu = N^2 X$ ) (shown in Figure 2 (a)):

$$\nu \otimes (\alpha_1 X^4 + \alpha_0 X)^2 = \beta_1 X^4 + \beta_0 X = [(\alpha_1 + \alpha_0)^2] X^4 + (\alpha_0 \otimes N)^2 X \tag{15}$$

Obviously, we can write the coefficients:

$$\beta_1 = (\alpha_1 + \alpha_0)^2, \beta_0 = (\alpha_0 \otimes N)^2 \tag{16}$$

Note we continue decomposing the operation in  $GF(2^4)$  into  $GF(2^2)$ , note that into  $GF(2^2)$  the inverse is the same as squaring, which is free with a normal basis (shown in figure 2 (e)):

$$(k_1 W^2 \oplus k_0 W)^{-1} = (k_1 W^2 \oplus k_0 W)^2 = k_0 W^2 \oplus k_1 W \tag{17}$$

The multiplication in  $GF(2^2)$  has a same structure as in  $GF(2^4)$  except the scaling norm is 1 (show in Figure 2(h)).

Up to now the remaining operation need in subfield  $GF(2^2)$  is scaling by  $N = W^2$  and combined operation of multiplication with scaling by  $N$  (shown in Figure 2 (g),(f)).

$$N \otimes (\alpha_{01} W^2 \oplus \alpha_{00} W) = [\alpha_{00}] W^2 + [\alpha_{00} + \alpha_{01}] W \tag{18}$$

$$\begin{aligned} N \otimes (p_1 W^2 + p_0 W) \otimes (q_1 W^2 + q_0 W) &= \eta_1 W^2 + \eta_0 W \\ &= [(p_0 \otimes q_0) \oplus ((p_1 \oplus p_0) \otimes (q_1 \oplus q_0))] W^2 + [(p_0 \otimes q_0) \oplus (p_1 \otimes q_1)] W \end{aligned} \tag{19}$$

Where

$$\eta_1 = (p_0 \otimes q_0) \oplus ((p_1 \oplus p_0) \otimes (q_1 \oplus q_0)) \tag{20}$$

$$\eta_0 = (p_0 \otimes q_0) \oplus (p_1 \otimes q_1) \tag{21}$$

In  $GF(2)$ ,  $\otimes$  means AND and  $\oplus$  means XOR bitwise[7].

#### IV. HARDWARE IMPLEMENTATION

In order to reduce area consumption, we proposed an efficient hardware implementation by reusing certain module and applying composite field to implement 8-bit S-box. As depicted in Fig. 1, the F module consists of two 8-bit S-boxes and is reused two times. 8-bit S-box, as non-linear layer, costs a large proportion of area resource. Therefore, we divide one round into three cycles, so that the F module can be reused.

Three cycles are used to implement one round. Firstly, we use two 80-bit registers to store the state and key respectively. Additionally, one 40-bit register is used to store initial value for later swapping in the last cycle. And one 16-bit register is used to store round constant. The output at the end of each cycle will be stored in the registers and reused as the input data at the beginning of the next cycle. Consequently, in the first cycle, one F module is used. In the second cycle, one 40-bit XOR, one 16-bit XOR and one L module are used. In the third cycle, F module can be used again. Therefore, F module is used two times in one round.

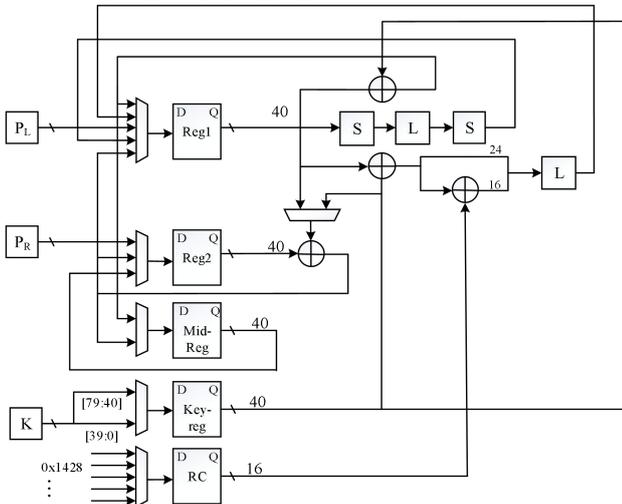


Figure 3 datapath of the hardware implementation of ITUbee

The reuse of F module saves a significant amount of area in hardware implementation. The L module can be implemented with simple XORs and bit operations. Fig.3 shows the datapath of ITUbee, which performs one round in 3 clocks and needs 61 clocks in total to implement 20 rounds. For comparison, we also implement a round-based architecture, which performs one round in one clock. In this way, one 40-bit register for storing initial value is saved, but the F module is used two in one single cycle, which increase the area extremely.

The 8-bit S-Box in ITUbee is the S-box used in AES. Except for using LUTs, which is simple 256 case statements in hardware, we can apply the mathematical formula to compute S-box. As we give in section 3, we use composite field to solve the complex inverse calculation, which turns out to be more efficient in hardware.

#### V. RESULTS

We implemented the proposed design in verilog DHL and synthesized it on 0.18 um CMOS technology to check its hardware complexity. In this area-optimized implementation, a 40-bit width datapath was used. In order to compare the area requirements independently it is common to state the area as gate equivalents (GE). One GE is equivalent to the area which is required by the two-input NAND gate with the lowest driving strength of the corresponding technology. The area in GE is derived by dividing the area in  $um^2$  by the area of a two-input NAND gate. Encrypting 80-bit plaintext with an 80-bit key occupies about 6748 GE and requires 61 clock cycles.

Specifically, in the above implementation, the area requirement is mostly occupied by S-boxes registers. 80-bit state register costs 612GE. 40-bit key register costs 306 GE. 40-bit middle register costs 306 GE. 16-bit round constant register costs 122.4 GE. one 8-bit S-box requires 356.6 GE. L module consumes 170.3 GE. 40-bit XOR costs 106.4 GE.

Table III gives the comparison between LUTs implementation and composite field implementation of 8-bit S-box, and table 4 gives the comparisons between proposed implementation and round-based implementation of ITUbee.

Table III. Area comparison between LUTs and composite field

Composite-field(GE)	LUTs(GE)
543.8	356.6

Table IV. Area comparison between proposed and round-based implementation

	Combinational area	Noncombinational
Proposed (6747.9)	5497.7	1350
round-based (9812.8)	87757	10371

Conclusively, the proposed hardware implementation of ITUbee requires about 6448 GE on 0.18 um technology. The area consumption of ITUbee is roughly 31.2% less than the round-based implementation. And it costs 365.6 GE to implement 8-bit S-box by using composite field, 32.7% less than by using LUTs.

#### VI. CONCLUSION

There is a great improvement in terms of area consumption by using the composite field to implement the S-box compared with the approach using LUTs. In our hardware architecture, we reuse the S-box in F function, which consists of ten 8-bit S-boxes and area occupancy proportion is more than 90%, by dividing each round into 3 clock cycles to reduce the area consumption further. We implemented the proposed design in and synthesized it on 0.18um CMOS technology, the results show that the area is saved by 32.7% by using composite field

S-box compared with using LUTs, and 31.2% by reusing S-box compares with not reusing.

#### ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (Grants 61472208) and National Key Basic Research Program of China(Grant 2013CB338004).

#### REFERENCES

- [1] Leader, G., Paar, C., Poschmann, A., Schramm, K., (2007): New Lightweight DES Variants. In: Biryukov, A. (ed) FSE LNCS, Vol. 4593, pp196-210. Springer, Heidelberg (2007).
- [2] Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knezevic, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçın, T. (2012): PRINCE - A Low-Latency Block Cipher for Pervasive Computing Applications - Extended Abstract. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 208–225. Springer, Heidelberg.
- [3] Standaert, F.-X., Piret, G., Gershenfeld, N., Quisquater, J.-J. (2006): SEA: A Scalable Encryption Algorithm for Small Embedded Applications. In: Domingo-Ferrer, J., Posegga, J., Schreckling, D. (eds.) CARDIS 2006. LNCS, vol. 3928, pp. 222–236. Springer, Heidelberg.
- [4] De Cannière, C., Dunkelman, O., Knežević, M. (2009) : KATAN and KTANTAN — A Family of Small and Efficient Hardware-Oriented Block Ciphers. In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 272–288. Springer, Heidelberg.
- [5] Ferhat Karakoc, Hüseyin Demirci. & A. Emre Harmanci, (2013): ITUbee: A Software Oriented Lightweight Block Cipher, Lightsec, LNCS 8162, pp16-27.
- [6] Zbotin, I.A., Glazkov, G.P., Isaeva, V.B., (1989): Cryptographic Protection for Information Processing Systems. Cryptographic Transformation Algorithm. Government Standard of the USSR, GOST 28147-89.
- [7] D.Canright, (2005), A very compact S-box for AES, Springer.

# Weighing the shopping benefits of a smarter refrigerator

Stephen Goeddel, Pasha Sadeghian, Aspen Olmsted

Department of Computer Science

College of Charleston

Charleston, SC 29423

goeddelm@g.cofc.edu, psadeghi@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—An IOT smart fridge containing scales can increase the utility of the product. By adding scales to regular image sensing technologies, a fridge can be created that makes fewer mistakes when creating a shopping list. This, combined with the fact that having a shopping list when heading to the grocery store can improve health and save money, makes for a useful smart fridge.

**Keywords**—IOT; Smart Fridge; Scales; Image Sensing;

## I. INTRODUCTION

The Internet of Things (IOT) trend has provided many advancements in everyday life as typical household devices keep getting smarter. Nutrition.gov states that creating a shopping list before heading to the grocery store can “improve your health while saving you time and money” [1]. Because of this, an IOT refrigerator contains the potential to improve the diets and ease of shopping for many people.

Research has been done on how to identify the items in a smart fridge in the past based on image recognition and RFID tag sensing, but these methods alone do not allow the fridge to track the remaining contents of a grocery item. By utilizing an image recognition system, as well as weight sensors, a smart fridge can create better shopping lists by accurately identifying the amounts of food within itself.

This paper will be organized into the following six sections. Section II describes related work as well as the limitations of the current methods. Section III provides a motivating example where using scales and image recognition is useful. Section IV details the algorithm to be used with the sensor input. Section V discusses the findings. Section VI concludes the paper and discusses possible future work.

## II. RELATED WORK

### A. *The use of image scanning to detect the items stored in a fridge*

“ProVisions: Object Recognition Applied to Smart Refrigerators” discusses the usefulness of a scanning system to accurately identify the products contained within a smart fridge [2]. Our research improves upon the content sensing of the image scanning detailed in that paper with the addition of scales.

### B. *Developing a shopping list based on user’s eating habits*

“iFridge: An Intelligent Fridge for Food Management based on RFID Technology” details how a recipe recommendation can be formed based on the items in a smart fridge [3]. Our research

adapts the ideas of recipe recommendation detailed in this paper to use scales and image processing rather than RFID.

### C. *Nutritional value can be gained by knowing what is in the fridge*

“A smart fridge with an ability to enhance health and enable better nutrition” discusses how dietary needs can be better achieved using a smart fridge [4]. We utilize the ideas within this paper and improve upon them with the introduction of scales.

## III. MOTIVATING EXAMPLE

Grocery shopping is more productive and much easier with a list but creating a list that properly utilizes the products that we currently own can be challenging. If a consumer had a refrigerator that knew what grocery products were contained within it and how much of each product was left, grocery shopping would be a breeze.

## IV. SMART SHOPPING LIST CREATION

A smart fridge could be designed with each shelf having four or more cameras and a grid of scales each small enough to have a part of a single product on them. Figure 1 details what a shelf on this fridge might look like.

This fridge would be able to use a cloud based database of products and recipes as well as a local database of what products, along with their amounts in grams, are currently stored within the fridge. Figure 2 shows a schema that this database might follow.

It was necessary to gather some sample recipes and ingredients to enable us to test out the usefulness of scales in a smart fridge. This was done by using the Food2Fork API [5]. We gathered 140 recipes using 245 total ingredients. Out of the ingredients in these recipes around 50% of them were used in at least one other recipe.

Since the ingredients came in a non-standard form, they required normalization before they were usable by an algorithm. We chose to standardize around grams as a unit of measure for the amount necessary of each ingredient.

To properly test our hypothesis that the addition of scales would be helpful in creating a shopping list, we needed to be able to model the contents of a user’s fridge. We created a script in Python to build up a model of a user’s fridge in the database.

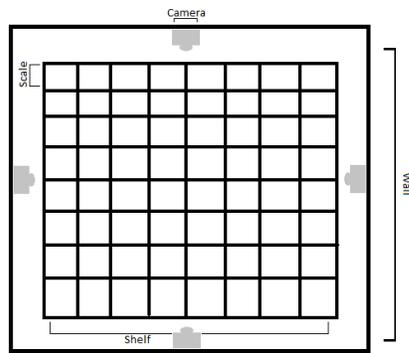


Figure 1. An example of how a shelf could look

This script was made to allow configuration of the number of shelves in the sample fridge.

We formulated two algorithms for creating shopping lists: one with the knowledge of each product’s weight factored in, and one without. The algorithm that didn’t account for weight simply iterated through each recipe while determining which ingredients were required to make that recipe. If the user owned the ingredient, it stored it in a list denoting such. Otherwise, the ingredient was stored in a second list. Finally, the shopping list was determined by choosing the recipes with the fewest ingredients left to purchase.

The second algorithm incorporated the weight information from the scales into the shopping list determination. This algorithm follows a similar pattern to the first. Initially, we iterate through the recipes determining which ingredients are needed to make them. Each recipe receives two lists: one with the ingredients that need to be purchased by the user as well as the amount, and another with the ingredients the user already owns along with the amount they will have left after making this recipe. The final step involves determining which recipes will be made as well as what ingredients need to be added to the shopping list. The algorithm again considers the recipes with the least number of unowned ingredients first. Once a recipe is chosen, and its ingredients are added to the shopping list every other recipe must be updated to remove the respective amount (in grams) of each owned ingredient that was used to make the newly added recipe. This allows the weighted algorithm to be

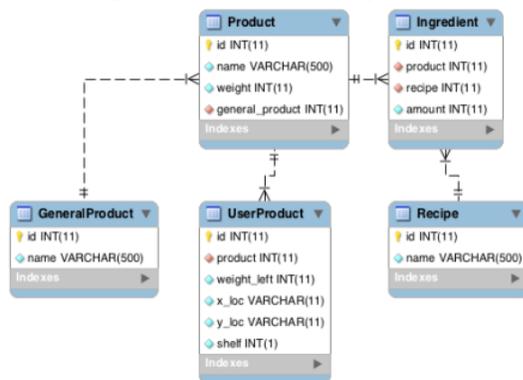


Figure 2. Possible smart fridge database schema

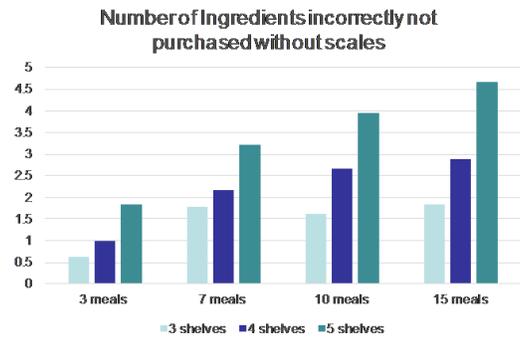


Figure 3. Ingredients incorrectly unpurchased without scales

sure that no ingredients are incorrectly unpurchased and that every recipe it suggests can be made by only purchasing the exact amount of each ingredient listed.

### V. FINDINGS

We then ran the two previously detailed algorithms fifty times each for varying numbers of shelves and number of meals to be included on the shopping list. Through this, we noticed a trend toward more mistakes made by the simple (non-weighted) algorithm when there were more shelves in the user’s fridge and/or more meals were included on the shopping list. Figure 3 details this trend.

### VI. CONCLUSION AND FUTURE WORK

By utilizing scales as well as image sensing technologies, a smart fridge was proven to have greater accuracy when creating a shopping list. A fridge that uses a similar algorithm would be able to create shopping lists that are more useful to a user and could possibly make the user healthier and provide them with a more enjoyable shopping experience. In the future, it would be useful to go through the exercise of outfitting a fridge with a grid of scales and cameras and testing out the accuracy of which it identifies the weight of each item.

### REFERENCES

- [1] C. Pester and L. Wilder, "nutrition.gov," 10 2 2017. [Online]. Available: <https://www.nutrition.gov/shopping-cooking-meal-planning/food-shopping-and-meal-planning/build-healthy-diet-smart-shopping>. [Accessed 11 2 2017].
- [2] L. Mangano, "Object Recognition Applied to Smart Refrigerators," 2013.
- [3] L. Xie, B. Sheng, Y. Yin, S. Lu and X. Lu, "iFridge: An Intelligent Fridge for Food Management based on RFID Technology," in *UbiComp '13*, Zurich, 2013.
- [4] S. Luo, J. Li and J. S. Jin, "A smart fridge with an ability to enhance health and enable better nutrition," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 4, no. 2, pp. 69-79, April 2009.
- [5] Food2Fork, "Food2Fork API," [Online]. Available: <http://food2fork.com/about/api>. [Accessed 10th March 2017].

# Optimizing Synchronization of Cloud Storage Services

## Combining Benchmark Monitoring and Learning-Based Framework

Preston T. Owens, Aspen Olmsted  
 Department of Computer Science  
 College of Charleston, Charleston, SC, US  
 owenstp@g.cofc.edu, olmsteda@cofc.edu

**Abstract**— With data storage moving further away from locally based storage and into an age of cloud storage, users are going to need their data on the go, and they're going to need it to be fast, and they're going to need it to be accurate. Eventual consistency is the theoretical guarantee that, provided no new updates to an entity are made, all reads of the entity return the last updated value. In this paper, a comparison is made on cloud-based studies to formulate an idea as to how to improve the synchronization of cloud storage systems by first benchmarking the eventual consistency and then utilizing a framework that dynamically learns the characteristics of a storage node so that eventually consistency is achieved quickly.

**Keywords**-Cloud Storage, Synchronization

### I. INTRODUCTION

Consistency among data synchronization is a major concern on the client side as well as the rate at which that synchronization occurs. This paper will address that topic and come up with a viable solution for cloud storage consistency.

### II. RELATED WORKS

Researched used: This idea comes from using two research papers that were studied. The first one [1] was a long-term study done to determine the most cost efficient way to continuously monitor consistency behavior, using the Indirect Consistency Monitoring method. Once the system has been evaluated, then a learning-based framework can be implemented that aims at scheduling data writes during user idle times such that the impact on user performance is limited within strict predefined targets while the updates are complete as fast as possible, as outlined in the second research paper [2].

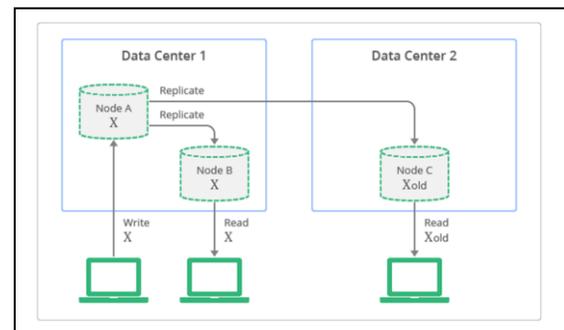
### III. CONSISTENCY

If all replicas are identical and all ordering guarantees of the corresponding consistency model are observed [3] – [4]. Eventually, consistency does not require any ordering guarantees and only offers eventual replica convergence. Consistency is formed through two dimensions, staleness and ordering. Staleness refers to how far apart the replicas are in terms of time or missing versions. Ordering refers to which

degree; incoming requests may be reordered on different replicas [1].

### IV. EVENTUAL CONSISTENCY

It is a theoretical guarantee that, provided no new updates to an entity are made, all reads of the entity will eventually return the last updated value. The following diagram, figure 1, illustrates that although replicas are always available to read some replicas may be inconsistent with the latest write on the originating node, at a particular moment in time<sup>1</sup>.



**Figure 1: Conceptual Depiction of Replication with Eventually Consistency**

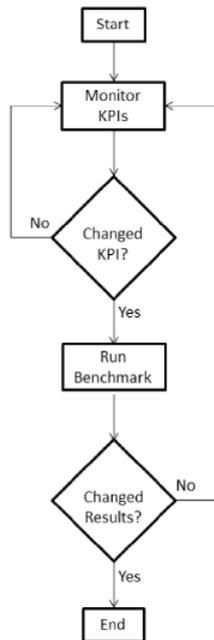
### V. INDIRECT CONSISTENCY MONITORING

The way the monitoring works is that instead of directly measuring consistency behavior KPIs (key performance indicators) are monitored instead. KPIs are both directly affected by inconsistencies and less expensive to track. Every time a KPI changes the consistency benchmark is rerun. This reduces the cost of consistency monitoring since the benchmarks are only triggered when necessary [1].

There are four steps to Indirect Consistency Monitoring before it can be used. Step 1 is identifying datastore interaction patterns. Step 2 is identifying potential conflicts between patterns. Step 3 is identifying the affected KPIs. And Step 4 is to identify and track suitable KPIs [1].

Indirect consistency monitoring is currently not directly supported by a tool. The consistency benchmark tool used by

the researchers is an open source one that can be automatically deployed and run by a system like Opscode Chef. KPIs can also be tracked by standard monitoring tools like Ganglia[1] .



**Figure 2: Flow Chart for Indirect Consistency Monitoring**

## VI. LEARNING BASED FRAMEWORK

Once we have evaluated the system to know whether its consistency is lacking, then we implement the framework. Without the framework, the user may be forced just to switch providers completely. Combining these two methods removes the cost and frustration associated with finding a new provider and transferring data [2]. The framework focuses on how to schedule the asynchronous data updates such that the performance in the sending and receiving nodes meets the predefined quality of service (QoS) goals. The framework learns the characteristics of the workload the nodes are serving and continuously updates the scheduling parameters [2].

The three methods of scheduling the researchers compared were Aggressive scheduling, Utilization-guided (Aggressive) scheduling, and Utilization-guided (Conservative) scheduling. The Aggressive scheduling schedules asynchronous updates immediately and without any consideration of foreground user traffic. Utilization-guided (Aggressive) takes the user traffic into consideration by monitoring the utilization. If the utilization is below a certain point, then it schedules asynchronous updates immediately. And Utilization-guided (Conservative) uses the system utilization as guidance and schedules the asynchronous updates only when the utilization is low [2].

The learning aspect of the scheduling policy consists of understanding the available idle times that can be used to serve

the asynchronous updates. Experiments involving simulations from data collected in real storage systems shows just how powerful the framework can be. It proves to be faster than normal utilization-based scheduling and updates like an aggressive policy that schedules the asynchronous updates as soon as the involved nodes become idle [2].

## VII. CONCLUSION

The purpose idea is that indirect inconsistency monitoring is combined with the learning based framework which will help to optimize cloud-based storage systems by first identifying systems that are underperforming and then utilizing the framework to determine how fast the newly written data can be asynchronously transmitted between nodes in a distributed storage environment. With the framework implemented making major changes to application implementation or cloud service provider becomes unnecessary.

## REFERENCES

- [1] Vogels, "Eventual Consistent", *Queue - Scalable Web Services*, New York, NY: ACM, 2008.
- [2] Yan, Hughes, Riska and Smirni, "Overcoming Limitations of Off-the-Shelf Priority Schedulers in Dynamic Environments," in *Modeling Analysis & Simulation of Computer and Telecommunications Systems (MASCOTS) 2013 IEEE 21st International Symposium on*, pp. 505-514, 2013.
- [3] Steen and Tanenbaum, *Distributed Systems - Principles and Paradigms*, 2nd ed, Upper Saddle River, NJ: Pearson Education, 2007.
- [4] Tai and Bermbach, "Informed Schema Design For Column Store-Based Database Services," in *Service-Oriented Computing and Applications (SOCA) 2015 IEEE 8th International Conference on*, pp. 163-172, 2015.

## **Session 16: Cloud Security and Digital Forensics**

Title: Preventing vendor lock-ins via an interoperable multi-cloud deployment approach  
(Authors: Roland Pellegrini, Patrick Rottmann, Georg Strieder)

Title: A Novel Multimedia-Forensic Analysis Tool (M-FAT)  
(Authors: Shahlaa Mashhadani, Hiba Al-kawaz, Nathan Clarke, Steven Furnell, Fudong Li)

Title: Best of Two Worlds: Secure Cloud Federations meet eIDAS  
(Authors: Thomas Zefferer, Dominik Ziegler, Andreas Reiter)

# *Preventing vendor lock-ins via an interoperable multi-cloud deployment approach*

Roland Pellegrini

University of Applied Sciences  
Burgenland

Campus 1, 7000 Eisenstadt

Email: 1610781022@fh-burgenland.at

Patrick Rottmann

University of Applied Sciences  
Burgenland

Campus 1, 7000 Eisenstadt

Email: 1610781002@fh-burgenland.at

Georg Strieder

University of Applied Sciences  
Burgenland

Campus 1, 7000 Eisenstadt

Email: 1610781020@fh-burgenland.at

**Abstract**—A vendor lock-in makes customers dependent of a propriety product, service or technology provided by a vendor. In terms of cloud services, it is achieved by providing and developing services that are platform-dependent with proprietary technologies, interfaces or formats. This can be a huge barrier to adopt cloud services because customers are tied to specific vendors which prevent portability through proprietary or a limited set of interfaces. As consequence, customers are hindered to alternate vendors easily without substantial migration costs. A multi-cloud strategy is one possible way to avoid vendor lock-ins. The scope of the is paper is to design and implement a solution-stack to provide web services which operate independent from proprietary cloud service provider technologies. We provide a solution stack prototype which will support key functionalities such as portability, interoperability and platform-independency by implementing modern technologies and standards.

**Keywords:** *Cloud Computing; Multi-cloud; Vendor lock-in*

## I. INTRODUCTION

Cloud computing transforms the way how individuals and organizations operate across domains. Gartner [2] forecasts the transition from a cloud-first to a cloud-only approach by 2019.

Today, many organizations have developed strategies to outsource workloads to the cloud, mostly by using a single cloud service provider. Current software stacks do allow interoperability among platforms, but APIs of different cloud service providers are often proprietary or do not rely on standards [3]. One potential solution to grant interoperability is to standardize the APIs in such a way that clients can deploy services and data across multiple cloud providers without adapters or complex migration scenarios [3] [7]. In addition, interoperability among cloud service providers would enable multi-cloud deployments; this would prevent the customer from major damage in case of a failure or outage of a single cloud service provider.

The basic characteristics of cloud computing include capabilities such resource-pooling and rapid elasticity [1]. However, the standardization of application programming interfaces (API) in the Cloud context is not part of the NIST definition [1]. The demand of data migration between different

cloud service providers and interoperability between cloud services exists [7]. Therefore, appropriate standards are necessary to ensure interoperability, portability, compliance, trust, and security.

This document addresses the following objectives:

1. Which factors must be considered to prevent vendor lock-ins through cloud service provider using a multi cloud deployment?
2. Which layer-specific mechanisms must be implemented to deploy and operate a service in an interoperable multi cloud environment?

A solution-stack prototype including a multi-cloud capable service will be designed and implemented based on the principles based on rapid prototyping-, requirements engineering- and multi cloud principles to prevent the negative effects of vendor lock-ins.

## II. RELATED WORK

The reviews of the following literature have shown an urgent need for preventing vendor lock-in. As discussed by [17], vendor lock-in already exists with traditional in-house software or technical infrastructure. Once, the software or the technology from one or a small group of vendors has been installed, it requires a great effort to switch to the technology of different vendors. Cloud computing often uses an architecture different than the one used by the traditional on-premise system. Once, a system is moved to the cloud, it isn't always a simple matter to bring it back on premise or to move it to another cloud. Cloud computing, as defined by [1], comes with three different service models (IaaS, PaaS, and SaaS) which allow user, organizations, and enterprises to consume a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort. Cloud Computing offers on-demand access to IT resources operated from a remote source. Many organizations choose Cloud services from a single Cloud Service Providers (CSP) because they value the comfort and usage of integrated services, single sign-on, harmonized accounting, standardized support processes, and a common user interface. Consequently,

customers often accept a certain dependency although CSP offers their own way on how Cloud consumers interact with the cloud [18]. As stated by [20], existing Cloud Computing solutions have not been built with interoperability in mind. Controversy, they lock customers in to a single Cloud infrastructure, platform or service preventing the portability of data or software. Furthermore, big vendors like Amazon, Google, or Salesforce, propose their own solution with proprietary formats and interfaces instead of widely accepted standards. This heterogeneity of single Cloud solutions and service interfaces increases the lock-in effect since most of the Cloud service resources bind the customer to stick with one cloud technology due to high cost in porting the application and data to a different Cloud Service Provider [7]. When a cloud consumer wants to find more suitable solutions for customers need or when the vendor is no longer able to provide the required service, a vendor lock-in can prevent a customer from switching to another cloud service provider [19].

To promote independency from one single CSP, organizations need to use more than one cloud as their service providers [15]. This infrastructure, also known as the multiple-clouds or multi-cloud, refers to the ability of two or more distributed Cloud services to collaborate and work together to be consumed in serial or simultaneously. As mentioned by [15], multiple Cloud Computing is a newly founded area and it is still suffering from lack of transparency. In multi-cloud environments, the main problem is portability as well as automatized deployment, service aggregation, and so on. According to [16], portability is the ability to move software assets among different runtime platforms without having to rewrite them partly or fully. However, the current software stacks of Cloud services are heterogeneous and the provided features that are often incompatible between different service providers. As a result, this diversity is an obstacle with respect to demands such as promoting portability and preventing vendor lock-in. The portability is requested by reasons varying from optimal selection regarding utilization, costs or profits, to technology changes, as well as legal issues. If the Cloud portability is achieved, data, application or service components can be moved and reused regardless of the operating system, storage or application programming interface (API). From the implementation perspective, [16] classifies Cloud portability in three categories:

1. Data portability - This enables the re-use of data components between Clouds.
2. Functional (or application) portability - This refers to the Cloud service agnostic definition of application functionality.
3. Service (or platform) portability - This is the ability to add, reconfigure and remove Cloud resources on the fly, independent on the Cloud provider.

Based on the works in literature, the following sections will describe an approach for preventing vendor lock-in by using a multi-cloud environment.

### III. USE CASE

As mentioned in previous chapters, the ability to prevent a vendor lock-in by a cloud service provider requires an appropriate service architecture and solution stack. Providers, stacks, libraries and frameworks are key elements to achieve multi-cloud capabilities as stated by [12]. According to [11] requirements like portability, interoperability, heterogeneity and geo-diversity are major challenges in cloud environments. Moreover, [13] states reasons like, avoiding the dependence on only one provider, react to changes of the offers by the provider or react to constraints, like new locations as a need for multi-clouds.

Key criteria for multi-cloud environments are [12] [13]:

- Portability between cloud service providers
- Interoperability between cloud service providers
- Service operation in heterogenous cloud environments
- Service management in heterogenous cloud environments
- Service deployment in heterogenous cloud environments
- Avoid dependency on only one provider
- React to changes of the offers by the provider
- Improve availability and disaster recovery

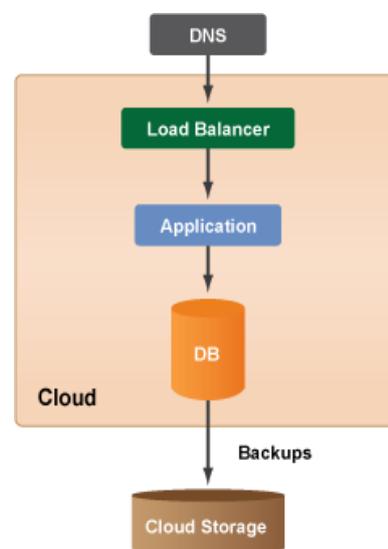


Figure 1. Single Cloud Site Architectures - Non-Redundant 3-Tier Architecture [14]

A lack of interoperability between CSP, but current software stacks do allow interoperability among platforms [3] [7] [4]. Solutions like interoperable software stacks, modern frameworks or micro service architecture enable opportunities to bypass a vendor lock-in [3] [7] [4].

The use case is based on a simple three-tier web service consisting out of a presentation tier, an application tier and a data tier (see Fig. 1). We assume that the service is running on a single CSP infrastructure in first place. The service itself is a small browser game. It writes the game scores continuously to the database during an active game session, the

current high score is displayed as leader board in each game session.

Comparing the initial situation to the relevant criteria like avoiding dependency on only one provider, it is obvious that the service does not comply with them. To prevent a vendor lock-in it is necessary to adapt the service components. Modern software stacks, service architectures or frameworks support the required steps.

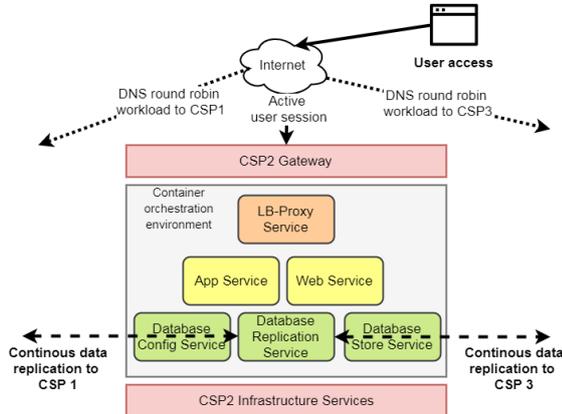


Figure 2. Multi-cloud provider Architecture - Microservice approach

The goal is to achieve interoperability through adding an abstraction layer represented through a container orchestration solution as well as a microservice approach. Therefore, major changes need to be done in terms of architecture, software and components to transform the generic service into a multi-cloud service. Mentioned criteria and requirements will support the construction process.

The solution stack to transform the initial web service into a multi-cloud solution will include a microservice architecture (see Fig. 2), a container orchestration layer, a NoSQL database layer and modern DevOps tools.

#### IV. PROTOTYPE

The use case forms the basis for the prototyping process, initially requirements are defined based on the future state of the service [21]. To bypass technical limitations of the initial scenario given by single cloud provider 3-Tier architecture adaptations need to be done. Major changes in comparison to the initial state of the service:

- Service Architecture - It is necessary to change the architecture from a 3-tier into a scalable and loosely coupled architecture, therefore a microservice is used.
- Database - To support replication, synchronization and scalability across multiple sites or cloud providers a NoSQL database replaces a typical RDBMS.
- Infrastructure & Orchestration - Virtual machines as basis for the services are replaced by container virtualization, but will still be used as host systems for

containers. A unified container orchestration is used to decouple dependencies to specific cloud providers APIs.

- Application Modelling Tool - An application modeling tool is used to build, scale and manage the container orchestration environment across multiple cloud service providers. This eliminates initial limitations and grants flexibility.

The service consists out of a basic browser game which relies on a stateless web service and a high available database service hosted on three Kubernetes clusters. Three VMs build the basis for the Kubernetes cluster, each cluster runs inside LXD containers within the VM. Each Kubernetes cluster is in different network zone; this simulates three cloud service providers. The zones are connected via routing instances simulating a global network connection. Per default, all services run in active-active mode across three Kubernetes clusters. A stateless load balancing mechanism which can be implemented via load balancer or DNS round robin distributes user requests between the three sites. The web frontend forwards data to the database router which redirects it to the primary database node (see Fig. 3).

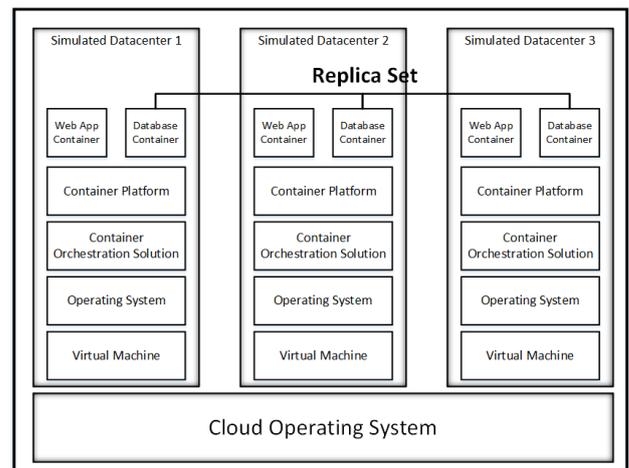


Figure 3. Software stack overview

The Solution stack components are:

- Programming Language Go
- Webservice Apache/Nginx
- Database MongoDB
- Container engine Docker
- Container engine LXD
- Container orchestration Kubernetes
- Operating system Ubuntu
- Application modelling tool Jujucharms
- Cloud Service Provider Openstack/CSPs

##### A. Service Architecture

Cloud computing is moving towards distributed and federated architectures of different services [5] [22]. Therefore, it is crucial to manage resources dynamically as response to changes across the provided service (application, cloud platform, etc.). One architectural concept for services to handle

dynamic changes is the microservice architecture [22]. This concept aims to realize software systems and services by packaging its components in small services. Each service has a specific task and can be deployed independently. All related service components run in separate processes and can run on different platforms. The services communicate through APIs without any centralized control [22]. This is a major advantage compared to the traditional 3-tier architecture.

As stated in [5], container approaches like Docker support and enable microservice architectures and improve service scalability as well as deployment processes. In order to simplify the deployment process, all service related components are packaged within one single file. That means each service and related client software are packed together as a single file and can be rolled out. The service is scaled or deployed by creating a new Docker container using Docker files and images. Based on a configuration file, Docker first retrieves a defined image from the Docker Registry, updates the dependencies, run through required configuration, and downloads the bundled server-client application from a central code repository. Depending on the stated operating system, given by the Docker file, different executables, builds or bundles are provided via a static central repository. Once the assembly is complete, the image is ready to be deployed by Kubernetes.

**B. Database**

The database layer is realized by a MongoDB cluster which supports replica sets and sharding. In terms of MongoDB, a replica set is a group of Mongo processes that maintain the same data set. Replica sets provide redundancy and high availability. In contrast, sharding is a method for distributing data across multiple MongoDB instances.

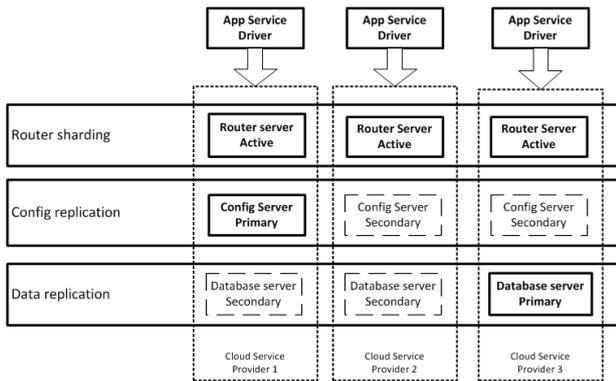


Figure 4. Database architecture

All MongoDB instances are deployed by and under to control of Kubernetes. The desired state is described in YAML files, which are processed by the deployment controller of Kubernetes for creating new services. The final setup (see Fig. 4) is derived from the MongoDB reference model [23].

**C. Infrastructure and Orchestration**

As discussed by [5] and [15], the main problem in multi-cloud environment is portability as well as automatized

deployment, service aggregation, interoperability, and so on. As a result, it is necessary to abstract the infrastructure layer of all CSPs. This means IaaS products are still in use but adapted by adding a container orchestration environment on top of basic VMs. This enables the consumer to control the unified APIs of the orchestration environment. Virtual machines now operate as container hosts (Kubernetes worker) while containers (Docker container) replace virtual machines as application hosts.

The container orchestration environment is realized by Kubernetes. It is an open source platform to deploy, scale and operate services based on container engines like Docker within fault tolerant cluster architecture. In addition, cloud federation functionalities are also supported. Currently, Kubernetes is certified for Google Cloud platform, Microsoft Azure, VMware, Amazon AWS, Joyntet, MAAS and Openstack [9].

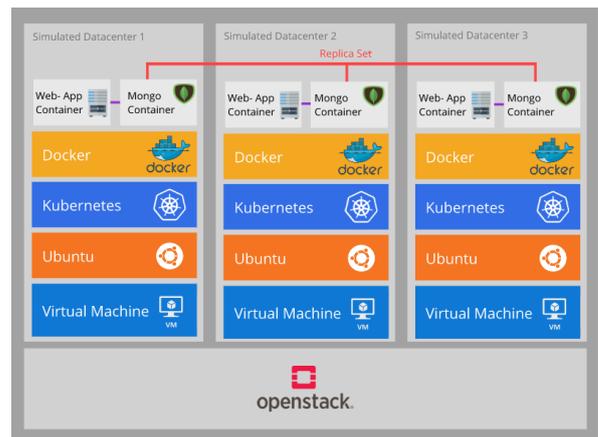


Figure 5. Prototype architecture

Therefore, Kubernetes can orchestrate multiple Kubernetes clusters distributed across multiple cloud service providers. Due to the minimal setup of the initial prototype setting, three separate Kubernetes clusters without federation have been implemented. The unified management is achieved by sharing the configuration files and images via a central repository. As a result, Kubernetes and Docker enables portability, interoperability, and automatized deployment within a multi-cloud environment.

As shown in Fig. 5, the initial prototype environment is deployed on an Openstack test environment, a free and open-source software platform for cloud computing, mostly deployed as Infrastructure-as-a-Service (IaaS). The launch of the prototype using cloud services is planned for the near future, but does not need additional testing, because the core software components are supported by major cloud service providers.

**D. Application Modelling Tool**

For the deployment of Kubernetes, the application-modelling tool Juju charms is used in combination with Conjure-up. Both tools guarantee standardized deployments of Kubernetes and can connect to APIs of major cloud service

providers and Openstack. To deploy Kubernetes via application-modelling tools, API keys with appropriate permissions are required [8] [10]. After connecting to the CSP's APIs, the application modeling tool deploys a standard deployment of Kubernetes.

#### E. Improvements

As discussed in the previous sections, portability and interoperability of data, functions, and services are the main problems in a multi cloud environment. In case of a vendor lock-in scenario, e.g. when vendor reliability becomes critical for a customer's business, cloud service may be limited or even no longer available for a customer. In contrast, each single Cloud in our multi-cloud approach remains independent, but still can interoperate with the remaining Clouds through our solution. This approach ensures data, application, and service portability and interoperability. Furthermore, our solution prevents any sort of vendor lock-in scenarios or outage scenarios.

### V. RESULT

Regarding to the first research objective which focuses on relevant factors to prevent vendor lock-ins, the following findings have been identified:

Various Cloud Service Providers offer different APIs and data formats for comparable cloud services. One opportunity of interoperability is determined at the lowest level of service model. IaaS allows the creation of any virtual machine with an operation system of your choice which supports a software stack of your choice. As stated in [3] [4] [7], a potential vendor lock-in can be bypassed by the implementation of an abstraction layer in combination with this software stack which enables you to provide unified API, portability, interoperability and provider independency.

Regarding to the second research objective which focuses on a layer-specific mechanism across multiple cloud providers, the following finding has been identified:

The prevention of a vendor lock-in cannot be guaranteed with a selection of individual solutions. As the prototype demonstrates, solutions such as Jujy, Kubernetes and Docker supports the deployment process, portability and the interoperable management. These tools enable customers to create a homogenous infrastructure on top of a heterogeneous cloud environment. In addition, a stateless application, which follows a microservice architecture paradigm, allows fine-grained interfaces, business-driven developments and lightweight container deployments with a decentralized continuous delivery. Furthermore, the implementation of decentralized data storage with appropriate replication mechanism allows data consistency and availability across different Cloud Service Providers.

An abstraction layer with a suitable software stack is the most crucial part of any multi-cloud architecture. This is also the bottom layer which must be under the direct control of a customer. The prototype provides capabilities like, dynamic provisioning, monitoring, and managing infrastructure across

multiple cloud service providers. In other words, this concept of infrastructure programming allows creating applications as well as the underlying infrastructure for the applications by using modern DevOps and orchestration tools. However, not all software stacks are applicable for all use cases and scenarios. Therefore, the selection depends strongly on the specific use case.

### VI. CONCLUSION

By using methods of prototyping, a comprehensive analysis of vendor lock-in has been discussed in this paper and the following reasons have been identified: (a) the lack of worldwide adopted standards or interfaces to leverage the dynamic landscape of cloud related offers, and (b) the absence of standards for defining parameters for cloud applications and their management. Therefore, a critical look has been addressed to the factors which must be considered to prevent the negative effects of vendor lock-in as well as implementation of layer-specific mechanism across multiple Cloud Service Providers.

Different approaches and methods for interoperability between various cloud service providers can be selected, such as standardized interfaces, middleware, or hybrid solutions. Based on the results of the prototype, an additional opportunity of interoperability through abstraction the software layer has been assessed. As a result, this approach puts us in the position to implement and run an own layer above the cloud service provider's infrastructure which helps us to achieve interoperability across any Cloud Service Providers.

We do not expect any significant changes of Cloud Service Provider's behavior to standardize their interfaces and corresponding API's. It is therefore reasonable to keep an eye of the development of software stacks that support implementation of abstraction layers for interoperability.

### REFERENCES

- [1] Peter Mell, Timothy Grace (2011), The NIST Definition of Cloud Computing. [Online]. Available: <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf> . Accessed April 29, 2017.
- [2] Gartner, Inc. (2017): Gartner Says By 2020, a Corporate "No-Cloud" Policy Will Be as Rare as a "No-Internet" Policy Is Today. [Online]. Available: [url{http://www.gartner.com/newsroom/id/3354117}](http://www.gartner.com/newsroom/id/3354117) . Accessed April 29, 2017.
- [3] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andrew Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia (2009). Above the Clouds: A Berkeley View of Cloud Computing. [Online]. Available: <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html> . Accessed April 29, 2017.
- [4] Adel Nadjaran Toosi, Rodrigo N. Calheiros, and Rajkumar Buyya. Interconnected cloud computing environments: Challenges, taxonomy, and survey. *ACM Computing Surveys (CSUR)* 47.1 (2014): 7.
- [5] Pahl, Claus, and Brian Lee. "Containers and clusters for edge cloud architectures--a technology review." *Future Internet of Things and Cloud (FiCloud)*, 2015 3rd International Conference on. IEEE, 2015.
- [6] Jothy Rosenberg, Arthur Mateos.: *The Cloud At You Service*. Manning Publications Co., 2011.

- [7] Justice Opara-Martins, Reza Sahandi, and Feng Tian. "Critical analysis of vendor lock-in and its impact on cloud computing migration: a business perspective." *Journal of Cloud Computing* 5.1 (2016): 1-18.
- [8] The Canonical Distribution Of Kubernetes. [Online]. Available: <https://jujucharms.com/canonical-kubernetes/55>. Accessed April 29, 2017.
- [9] Canonical Ltd. Kubernetes for the enterprise. [Online]. Available: <https://www.ubuntu.com/kubernetes>. Accessed April 29, 2017.
- [10] Conjure up. [Online]. Available: <https://conjure-up.io/>. Accessed April 29, 2017.
- [11] Fawaz Paraiso, Nicolas Haderer, Philippe Merle, Romain Rouvoy, and Lionel Seinturier. A federated multi-cloud PaaS infrastructure. In *Cloud Computing (CLOUD)*, 2012 IEEE 5th International Conference on Cloud Computing. IEEE, 2012.
- [12] Nicolas Ferry, Alessandro Rossini, Frank Chauvel, Brice Morin, and Arnor Solberg. Towards model-driven provisioning, deployment, monitoring, and adaptation of multi-cloud systems. In *Cloud Computing (CLOUD)*, 2013 IEEE Sixth International Conference on Cloud Computing. IEEE, 2013.
- [13] Dana Petcu. "Multi-Cloud: expectations and current approaches." *Proceedings of the 2013 international workshop on Multi-cloud applications and federated clouds*. ACM, 2013.
- [14] RightScale Docs. Cloud Computing System Architecture Diagrams. [Online]. Available: [http://docs.rightscale.com/cm/designers\\_guide/cm-cloud-computing-system-architecture-diagrams.html](http://docs.rightscale.com/cm/designers_guide/cm-cloud-computing-system-architecture-diagrams.html). Accessed November 24, 2017.
- [15] Pierre Riteau. Building dynamic computing infrastructures over distributed clouds. In: 2011 First International Symposium on Network Cloud Computing and Applications (NCCA), pp. 127–130 (2011)
- [16] Dana Petcu, Athanasios A. Vasilakos. Portability in clouds: approaches and research opportunities. *Scalable Comput.: Pract. Experience* 15(3), 251–270 (2014).
- [17] Michael Hugos, Derek Hultitzky. *Business in the Cloud: What Every Business Needs to Know About Cloud Computing*. Wiley Publishing, NJ. November 2010.
- [18] N. Pramod, Anil Kumar Muppalla, and K.G. Srinicasa. Limitations and Challenges in Cloud-Based Applications Development. In *Software Engineering Frameworks for the Cloud Computing Paradigm*, pages 55-75. Springer London, London, 2013.
- [19] Rajkumar Buyya, Christian Vecciola, and S. Thamarai Selvi. *Mastering Cloud Computing: Foundations and Applications Programming*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, First Edition, 2013.
- [20] Nikolaos Loutas, Eleni Kamateri, Filippo Bosi, and Konstantinos Tatabanis. Cloud Computing Interoperability: The State of Play. In 2011 IEEE Third International Conference on Cloud Computing Technology and Science, pages 752-757. IEEE Computer Society, 2011.
- [21] Ian Sommerville (2012): *Software Engineering*, 9th Edition, Page 96-118, Pearson Studium, Imprint Pearson Education Deutschland GmbH.
- [22] Pahl, Claus, and Pooyan Jamshidi. "Software architecture for the cloud—a roadmap towards control-theoretic, model-based cloud architecture." *European Conference on Software Architecture*. Springer, Cham, 2015.
- [23] MongoDB. Sharding. [Online]. Available: <https://docs.mongodb.com/manual/sharding/>. Accessed November 26, 2017

# A Novel Multimedia-Forensic Analysis Tool (M-FAT)

Shahlaa Mashhadani<sup>1</sup>  
Centre for Security, Communications and Network  
Research  
Plymouth University  
Plymouth, UK  
shahlaa.mashhadani @ Plymouth.ac.uk

Hiba Al-kawaz<sup>2</sup>  
Centre for Security, Communications and Network  
Research  
Plymouth University  
Plymouth, UK  
hiba.al-kawaz @ Plymouth.ac.uk

Nathan Clarke<sup>3</sup>  
Centre for Security, Communications and Network  
Research  
Plymouth University  
Plymouth, UK  
N.Clarke @ Plymouth.ac.uk

Steven Furnell<sup>3,4</sup>  
Centre for Security, Communications and Network  
Research  
Plymouth University  
Plymouth, UK S.Furnell @ Plymouth.ac.uk

Fudong Li  
Centre for Security, Communications and Network Research  
Plymouth University  
Plymouth, UK  
fudong.li @ Plymouth.ac.uk

**Abstract**— Digital forensics has become a fundamental requirement for law enforcement due to the growing volume of cyber and computer-assisted crime. Whilst existing commercial tools have traditionally focused upon string-based analyses (e.g., regular expressions, keywords), less effort has been placed towards the development of multimedia-based analyses. Within the research community, more focus has been attributed to the analysis of multimedia content; they tend to focus upon highly specialised specific scenarios such as tattoo identification, number plate recognition, suspect face recognition and manual annotation of images. Given the ever-increasing volume of multimedia content, it is essential that a holistic Multimedia-Forensic Analysis Tool (M-FAT) is developed to extract, index, analyse the recovered images and provide an investigator with an environment with which to ask more abstract and cognitively challenging questions of the data. This paper proposes such a system, focusing upon a combination of object and facial recognition to provide a robust system. This system will enable investigators to perform a variety of forensic analyses that aid in reducing the time, effort and cognitive load being placed on the investigator to identify relevant evidence.

**Keywords**- *Multimedia, Forensic Image Analysis, Annotation, Face Recognition.*

## I. INTRODUCTION

With the enormous increase in the number of images, videos, and audio recordings available, multimedia evidence has come to play a fundamental role in criminal investigations [1, 2]. The significant increase in the volume of photographs (images\*) and video context is having a direct impact on the time and cost of

investigations, with much of the current effort resulting in investigators having to manually analyse the context. Consequently, the forensic investigators require a set of forensic analyses to enable them to more efficiently identify relevant evidence [3]. As a result, Forensic Image Analysis (FIA) has emerged as a new branch of digital forensics that enables the investigators to effectively and accurately extract evidence from a huge number of images in an automatic and forensically sound manner [4]. However, there are at present many challenges that still exist. For instance, forensically, little work has been undertaken using object and face recognition to better understand the context of images. Should an investigator wish to identify all images with a particular object in, they would need to manually investigate each item. Likewise, should an investigator be interested in a particular individual (possibly the suspect) and wish to understand within the sources available, who this individual has interacted with, again, using current tools, a manual inspection and verification would be required. Whilst facial recognition could be utilised, current implementations only operate well within a very constrained set of external conditions (namely front-facial images with a consistent illumination) which often are not present within cases. Existing forensic tools such as EnCase and the Forensic Toolkit (FTK) are insufficient in areas such as automatic content image analysis, extraction of evidence, facial recognition, and in identifying and correlating images [5]. The aim of this paper is to present a novel and holistic multimedia forensic analysis system that can aid the investigation process in analysing, interpreting and correlating

<sup>1</sup> Computer Science Department, Collage of Education for Pure Science, Ibn Al Haytham, Baghdad, Iraq.

<sup>2</sup> Computer Science Department, College of Science for Women, Baghdad University, Baghdad, Iraq.

<sup>3</sup> Security Research Institute, Edith Cowan University, Perth, Western Australia, Australia.

<sup>4</sup> Centre for Research in Information and Cyber Security, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa.

\* In this research, the term image refers to a picture, photograph or other file that is typically associated with file extensions such as BMP, JPEG, and TIFF.

multimedia-based context. The proposed automated framework will be able to analyse a large volume of image sources in an efficient and accurate manner to create the necessary annotation and features (AF) that can be utilised to inspect, correlate and analyse the evidence. This will reduce the cognitive burden placed on the investigator when handling large volumes of data and thus provide more timely analysis of the data.

The remainder of the paper is organized as follows. Section II provides an overview of the current state of art within object and face recognition. Building upon this, Section III is devoted to the proposed M-FAT system architecture and processes that underpin the approach. Section IV provides an illustrative case study that highlights the advantages of the proposed approach. Section V presents a discussion. The conclusion is listed in Section VI.

## II. BACKGROUND

Image recognition can best be analysed under two methods (i.e., object and the more specialised face recognition) in order to comprehend the current capabilities and limitations. Efforts have been made to narrow down the search environment, so that the investigation is focused to the most current states of the art in object and facial recognition. The research methodology has utilized a range of keywords (object-based image retrieval (centric object retrieval, non-centric object retrieval), and multiple object-based image retrieval, followed by automatic image annotation studies, facial recognition, partial or disguise faces, facial aging, illumination, face pose and expression) to research related studies from various academic databases IEEE, Google Scholar, and Science Direct. The keyword “forensic” is used to find which studies are more related with this field.

Five criteria are applied to select the papers, these are: all publications less than two pages long (including posters, presentations, abstracts, or short theoretical papers) are excluded; non-peer-reviewed publications are eliminated; the language of this literature review is English; site number, impact factor, and publication year.

### 1) Object Recognition

Few studies have focused upon image analysis for the purpose of digital forensics and identifying and extracting evidence from images [6]. An analysis of these studies is summarized in Table I.

Some of these studies have offered good procedures for FIA and achieved high retrieval accuracy. However, they suffer from the fact that it deals with a specific criminal case. In addition, they have suffered from limitations in their work, such as undetermined number of images that used for experiments or analysis, or they only use a small volume of pictures. In addition, no criteria was applied to evaluate the performance, or no comparison with other studies was performed [7–10]. Moreover, the special characteristics of forensic images are different from characteristics of standard images; therefore, the image features that are suitable to describe standard image databases are inefficient for forensics. For example, the background of forensic photographs is typically far more complicated than those used within the experimental studies,

because the target object could be damaged, deficient, or the object may appear very small in the picture [3]. In addition, the clarity and accuracy of forensic image retrieval are essential requirements for any investigation; however, some real-life images suffer from noise or losing blocks such as losing a number of bits, when sending the image through a wireless channel, and thus require enhancement before analysis [11]. Manual image annotation is yet another challenge, because annotating image manually needs big cost, time consuming, etc. [12].

The findings highlight that there has been little work performed on the subject of extracting evidence from images or solving criminal cases through FIA. Moreover, very few studies are able to overcome the challenges of finding and discovering forensically interesting and suspicious or beneficial patterns within huge datasets while taking into account the requirements of accuracy and speed.

In order to overcome the above problems, research from existing fields such as Object-Based Image Retrieval (OBIR) and Automatic Image Annotation (AIA) could be employed on forensic images to retrieve specific evidence and thus to solve many of the current challenges of image analysis within the forensic domain. However, the forensic examiner needs an automatic system that can recognise multiple objects in the same image, although these objects may differ in size, colour, shape, texture, and orientation. Despite a considerable amount of literature having been published in OBIR, the main limitation is focusing upon having a single main object only. They tend not to focus upon real-life complex imagery. The experiments for these studies were also conducted on only a small and very specific number of images [13–15]. In addition, image retrieval accuracy decreases dramatically with an increasing number of images [16, 17]. Furthermore, there is a substantial gap between low-level content features (color, shape, etc.) that are used for OBIR and semantic concepts (e.g., keyword, text, descriptor) used by humans to interpret images. Moreover, in this approach, users must have an example or a query image at hand, because the query must be an image [18].

As it already mentioned, previous OBIR methods suffer from several kinds of issues. Consequently, AIA systems could be used instead of an OBIR to describe images with words in place of using image features. AIA is a process of automatically assigning words to a given image and it suggests a promising way of achieving more efficient image retrieval and analysis, by bridging the semantic gap between low-level features and high-level semantic contents in image access [19]. This will enable the ability to search based upon keywords and solve problems presented by OBIR systems. Therefore, AIA is considered a highly valuable tool for image search, retrieval, and archival systems [20]. However, AIA studies suffer from multiple problems such as there is no standard annotation database that has been utilised to evaluate system performance, and most studies conduct experiments using unrealistic image databases [21, 22]. In addition, there is a great disparity in system performance, because of the divergence in segmentation, features, and classifier approaches, as well as the number of images that used in the systems assessment [23, 24].

TABLE I. SUMMARY OF FIA STUDIES

Ref.	Object Extraction Method	Features Extraction	Performance (%) *		Database Name	#Images
			Precision	Recall		
[3]	---	Colour and texture		62 70	forensic and Corel databases	400 forensic images and 800 images
[6]	Background subtraction algorithm	Scale-Invariant Feature Transform (ASIFT) and min-hash technique	85	---	Three videos	203 vehicle object images
[7]	---	Colour, texture, and shape	---	---	---	---
[8]	---	Grey Level Co-Occurrence Matrix (GLCM), texture	---	---	fired bullets, firing pins, extractor marks, ejector marks, and cartridges	50 images
[9]	Region Of Interest (ROI)	Histogram, texture, entropy and Speeded-Up Robust Features (SURF)	98	---	---	250 images
[10]	---	Colour ,texture and shape	---	---	---	---
[11]	---	Filtering algorithm and Reconstructing algorithm	median filter	---	---	---
[12]	---	Scale-Invariant Feature Transform (SIFT)	90	--	tattoo images from Michigan State Police	64,000 tattoo images

\* Some results are approximated from studies.

Furthermore, it should be noted that studies have proposed solutions to the problems of multiple objects retrieval and AIA associated with image retrieval systems, and have achieved high retrieval accuracy. Even then, there is still a problem that none of these studies tested images related to forensic cases and real-life complex and diverse imagery. This makes it impossible to determine whether these studies would achieve high performance in FIA. Moreover, the forensic case images are changeable that makes it difficult to build for each case own AIA system.

## 2) Face Recognition

Face recognition has become more popular in forensics; however, a number of issues within a forensic context still need to be addressed. The efficiency of face recognition is affected by internal and external factors. Internal factors include uncooperative people in front of the camera such as pose variation, facial expression, faces occluded, accessories and aging [25]. External factors are unrelated to the user, such as light factors, camera quality, and more than one person in the same location, which could obscure the subject's face [26]. In order to improve the efficiency of the forensic facial recognition, these barriers should be thoroughly investigated. Some of the prior facial recognition studies are summarized in Table II and then discussed in the paragraphs that follow.

Focusing upon facial aging, several studies adopted generative face images according to age progression to minimize the age gap in face matching technique. [27] introduce one example of these studies; they used the craniofacial growth during formative years up to age 18 to improve the recognition accuracy. However, the drawback was ignoring the face texture

growth such as fat tissue (that could be an essential feature in the analysis process). While [28] generated series of age-progressed face photos between 1 and 80 years. They dealt with face shape and texture changes, which made the results close to the reality, but they depended on human decisions instead of an automatic identification system.

In comparison, other studies preferred to use the discriminative approach to solving the facial aging issue in face recognition system by using the local features of the face, which they consider is more robust to age variation [29, 30]. Moreover, they combined multi-feature descriptors to obtain more face discriminative information that could support the recognition system. However, the results are vulnerable to other issues such as pose change that could produce the low accuracy [29].

Human interaction in front of surveillance cameras has added new challenge in the forensic system. For example, head pose (e.g., frontal face or not), and partial face or occluded (e.g., face hidden by glasses, hat, and scarf). In spite of numbers of researchers have sought to overcome the facial pose issue in face recognition the limitations are low pose degree (e.g., 20° degree), one direction pose (e.g., horizontal face), and determine individual images required to process the system [31]. In addition, recognition accuracy decrease with an increasing of face pose degree [32]. Other researchers preferred to correct the face pose by creating a 3D face viewing from a 2D image [33, 34]. In some cases, the 3D model makes the system more robust due to the high discriminative information. Nevertheless, the 3D model needs additional time for processing images database.

TABLE II. SUMMARY OF FACIAL RECOGNITION STUDIES

Ref.	Approach	Recognition Accuracy (%) <sup>++</sup>	Database (Subjects, Images)
Facial Aging			
[27]	Shape growth modelling	15.0	Private database (109,233)
[28]	Automatically age progression	---	The Google Images (-, 40000)
[29]	Discriminative model	83.9	MORPH album 2 (20569,78207)
		47.5	FG-NET (82,1002)
[30]	Multiview discriminative model	65.2	MORPH album 2 (20569,78,207)
		91.8	FG-NET (82,1002)
Facial pose			
[31]	Mosaicing scheme	96.76	CMU PIE (68, 494)
		97.06	WVU Multispectral (40, -)
[32]	Gabor-based method	86.8	FERET (200, 1196)
		67.6	CMU PIE (68, 494)
[33]	3D transformation model	99	CMU-PIE (68, 494)
		95.6	FERET (200, 1400)
[34]	3D features model	95.31	FERET (200, 1400)
Illumination Factor			
[35]	The maximum filter	98.9	Yale B (10, 5760)
		94.44	extended Yale B (38, -)
[36]	the shadow compensated technique	99	CMU-PIE (68, 494)
		92.3	Yale B (10, 5760)

<sup>++</sup> Some results are approximated from studies.

Illumination factors also play a key role in the matching processing, holding a significant impact on the overall system performance. A number of studies have attempted to minimize illumination effects on images to increase the recognition accuracy. One of these studies proposed a method of filtering images with illumination variation to obtain smooth images for face recognition [35]. Moreover, [36] proposed a shadow compensated technique that adding weighted average intensity to light angles instead of shadow variations on the facial image. The problems of illumination in previous face recognition studies were due to limitations such as face pose, light angle, the capture environment (indoors, outdoors, night, etc.), and the face image noise and its effect on recognition accuracy.

Regarding multiple image issues, a small number of studies have tried to cope with multiple challenges within the face recognition system [37, 38]. They investigated a face recognition system based on facial expression, face pose, and illumination issues. Their systems only applied and evaluated

three issues on non-real life images individually. However, in their system the database does not include real-life images. On the other hand, [39] proposed a study to identify any suspect person in a large crowd of people with uncontrolled captured images. However, their system focused on partial face images rather than other image issues.

In order to improve the speed of the facial matching system, Park and Jain filtered database by using gender, and ethnicity as demographic information that does not change over time [40].

As demonstrated above, existing studies have attempted to deal with the different effects application of facial recognition. To the best of our knowledge, there have not been attempts to solve all issues together in one system. Additionally, [41] conducted a study of the Boston Marathon bombings of 2013 and analysed the reasons why the automated face recognition system failed to identify the suspected persons at the time. Their study concluded that forensic facial recognition system operates under unconstrained conditions of people in the presence of digital surveillance cameras. Therefore, the current forensic systems require further investigation to overcome the drawbacks of them.

### III. M-FAT ARCHITECTURE

The objective of the proposed system is about incorporating image analysis within a single case management-based system that goes beyond the current state of the art both within forensics and within their specific specialist domains. The key requirements are:

- Acquire and process a wide variety of base forensic images and live sources (e.g., computer, mobile, cloud, CCTV).
- To analyse and create the necessary AF (object-based or facial) to describe the nature of the image.
- To provide a range of forensic analyses and correlation capability to aid an investigator in querying the image source.

In analysing and creating the necessary AF, the proposed M-FAT will seek to overcome the aforementioned weaknesses of existing annotation and facial recognition systems to provide an effective and robust multimedia forensic analysis tool. The proposed framework is illustrated in Figure 1. This architecture consists of the following processes:

- Evidence selection: The first stage of the proposed system involves the forensic investigator collecting all videos and images from different sources such as CCTV camera, mobile, digital camera, computer images, hard drive and manual data. The system will refine the collection data through the exclusion of irrelevant images based on image metadata as identified by the investigator, in order to facilitate the process of selecting the target image. The output from evidence selection goes to image annotation and face feature extraction in order to recognize evidences and save the results in process evidence database, or to M-FAT manages to select the requested evidence.

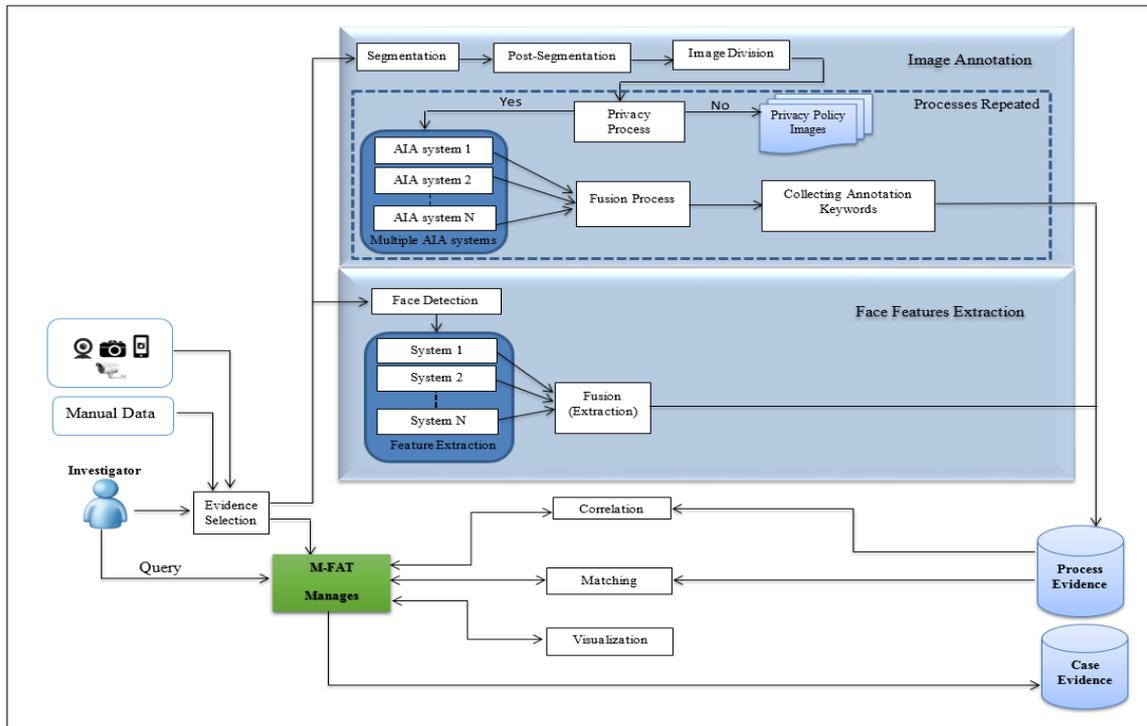


FIGURE 1. M-FAT ARCHITECTURE.

- **Image annotation:** This stage explains the process for extracting and annotating objects (evidences) for each image. It consists of the following processes:

- **Segmentation:** An image segmentation process will be carried out in order to divide the image into regions that permit a more meaningful and easier analysis. Forensic images include multiple objects with complex backgrounds; therefore, more than one algorithm will be used in order to extract multiple objects efficiently and effectively.

- **Post-segmentation:** Various problems that affect the accuracy of object extraction may appear after the segmentation process. These problems include noise, bumpy boundaries, unlinked boundary segments, objects with common boundaries, weak segments, over-segmentation, and edge segments that do not belong to realistic boundaries. It is thus very important for the post-segmentation process to resolve these problems, because the accuracy of extracting objects (evidence) will be inefficient without this step.

- **Image division:** In this process, the image will be divided into small pieces; and each piece contains one main object, making it ready for the annotation process. Each piece will be called an Object Image (OI). The main goals of this stage are to increase the annotation accuracy and to maintain image privacy before sending it to multiple external AIA systems.

- **Privacy process:** Sometimes an OI includes a label or text in its content, such as a name, a car registration number, or a personal address, which may be considered as important information. Thus, this process will reveal whether the OI includes any private information and, if so, the image will be saved in a separate list in order to address it individually. The images that are saved in the separate list will be tackled

separately by hiding important information through the use of a mask and then send it to external AIA systems, or by sending it to one secure external AIA system. If there is no private information, the OI will be sent to the next stage directly.

- **Multiple AIA systems:** In this process, the OI will be sent to multiple external existing AIA systems such as CLOUD VISION API and Microsoft Cognitive Services (Computer Vision API) for annotation. Then, the results from these systems will be collected and sent to the next stage. Multiple AIA systems will be employed in order to annotate these objects, and then their outputs are fused in order to improve the accuracy of annotation results over the results that can be achieved through employment of a single automatic annotation system. The AIA systems will describe the OI with words that name the object; however, these systems are unable to describe features of the object such as color or shape. Therefore, the proposed system will improve upon the annotation process by adding an annotation of these features to the external AIA system results.

- **Fusion process:** This process will be utilised to fuse the results from multiple AIA systems to provide more efficient results than the individual systems involved in the process. Combining annotation results from different systems will result in performance improvement.

- **Collecting annotation keywords:** The result from the previous process will be saved along with the image in order to build the final image annotation. The privacy process, multiple AIA systems, and fusion process stages are repeated depending on the number of objects extracted from the forensic image. Then, the final result (image with their annotated keywords) will be saved in a process evidence database.

- **Face Feature Extraction:** This stage determines faces from image and extracts their features. A detailed of each process will explain below:

- **Face Detection:** In the first process, the facial area will be determined and extracted from the image. In order to get the best performance outcome, a number of techniques will be used to seek best accuracy. After cropping the face from the entire image, the face will be normalized; and various face components such as eyes, mouth, and nose can be located.

- **Feature Extraction:** In this process, a multi feature extraction systems will be utilised to extract effective face features that can be used to improve face recognition quality (each one could to focus on different face features). The results of each system will save as vectors then sent them to the next process.

- **Fusion:** The fusion engine will build to fuse face features from multiple feature vectors. This process will tried to improve the overall accuracy by increasing the dimension of feature space. The final features will consider as evidences and save in the process evidence database.

- **M-FAT Manages:** The M-FAT manages is an interface between the investigator and the underlying system to provide the ability to search, correlate and visualize the data. The results from this interface will be saved in the case evidence database. Based on the requirements, the investigator can recall the aforementioned results that could be used as a potential evidence.

- **Matching:** This stage includes establishing a search engine connected to the process evidence database, which has the capability of accommodating single, or multiple keywords or query image. After that, the system will retrieve all images that satisfy the search conditions. For example, if the investigator inserts the text 'red car' with requested face image, then the system will retrieve all images that contain all conditions.

- **Correlation:** A Decision Support System (DSS) will be used at this stage to facilitate the role of the investigator by finding correlations between retrieval images based on metadata and AF, in order to construct the crime scene. This process will assist the investigator to find relevant pieces of evidence from among others.

- **Visualization:** Data visualization enables the investigator to see analytics presented visually, and assisting them to better understand complex concepts. For instance, google map, graph and report will be used to present the results.

#### IV. CASE STUDY

To help illustrate how the proposed framework would operate, a child abduction example is presented. In this example, it is assumed that a child has been kidnaped. Intelligence provides a rough last location for the child and information that they were seen getting into a red car. In order to solve a child abduction case, an investigator starts to collect all preliminary evidence that may help to find the child as fast as possible. For example, narrowing the timeframe of

abduction, examining properties of the car that a witness believes was involved in the abduction, determining the location of the abduction, and any information about suspect (e.g., face description, age, and gender). The next step of the investigative process would involve collecting all available imagery (e.g., videos from surveillance cameras at the crime scene and from nearby surveillance systems). Manual analysis of the sources in and around the timeframe would provide an investigator with an image of the child's face and of the car she was forced into. Timely analysis and evidence and the reporting of the investigators findings are critical to the safe recovery of the child.

The current solution would involve teams of investigators manually trawling through the footage from possibly dozens of evidence sources. The use of a manual human matching process is a laborious and time-consuming resulting in examining large volumes of image data and given the pressurised nature of the task likely to result in a high proportion of human error.

The proposed system will permit an investigator to select the necessary evidence sources, automatically process all of the footage. The investigator will be able to select the objects of interest – in this case the face of the child and the car that she got into. The system will then perform facial and object recognition across the evidence sources, providing an investigator with a prioritised set of results with which to interact with. The system will refine the retrieval results based on metadata (time, location, and date of the abduction) in order to reduce the number of retrieval results. The investigator will be able to target image (the suspect's car) from the retrieval results, and the DSS will provide further correlation and analysis functions that would enable the target car or face to be tracked across the different evidence sources. The resulting visualisation would provide the graphical map of the resulting journey alongside the image sources utilised to identify the path of the car. Where multiple paths are possible, the system will provide a probabilistic measure indicating which to investigate first as shown in Figure 2.

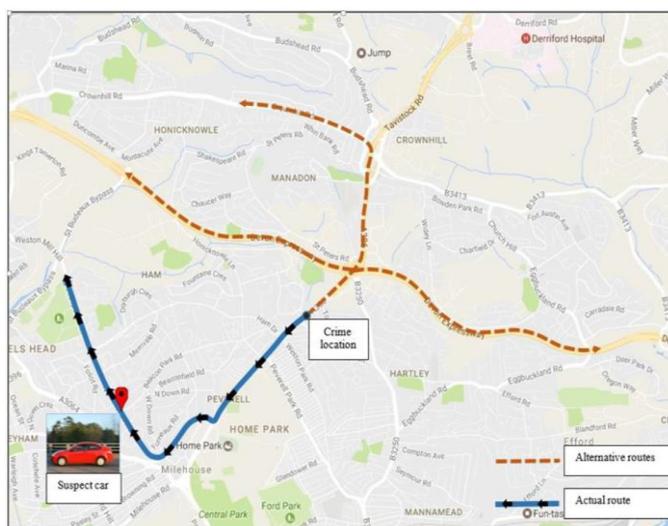


Figure 2. Example of car tracking result.

## V. DISCUSSION

In reviewing the literature, no study was combined between object and face in order to describe image contents. In addition, most studies have some drawbacks in their works especially that they have not been tested on real-life complex imagery and whenever the number of images is increased, their retrieval accuracy will decrease dramatically. Furthermore, there is not study that deals with all facial image issues together in one system. As a result, the premise of the proposed M-FAT is to combine between object and face in one system that meet investigator requirements, and capitalise upon existing research and systems in a multi-algorithmic manner to benefit from the different feature extraction and classification approaches. In order to achieve the proposed system requirements: Firstly, various techniques such FTK and Mobile Phone Examiner Plus (MPE+) will be used to acquire images from different sources (e.g. CCTV, computer and mobile) base on the source type. In addition, the acquired images that need to be investigated, suggesting that these images are usually large in number, vary in quality, unconstrained illumination, various orientation, object size, irregular background, and contain multiple objects. As a result, these images are voluminous and need processing often in near real-time and in doing so maintain the level of accuracy. The quality of images is another an important issue in image analysis, because the reliability of any inspection task is based on that quality. Therefore, the image under consideration should be checked first to determine whether the image quality is sufficient to allow for a meaningful and reliable analysis

Secondly, different multiple object and facial recognition systems that have the ability to recognise different objects and faces with different characteristics from the image will be examined and their results will be fused, in order to improve the evidence extraction process. The objective of using multiple systems is to overcome limitations of each system individually and looking for different reliable information. Previous researches within biometrics in particular have shown this to be beneficial [42, 43]. However, further work needs to explore the extent to which a multi-systems approach would work in this specific application.

Finally, accuracy and speed of retrieving images are the biggest challenges facing the use of image analysis in digital forensics. However, once annotated or query image, merely looking at all the results of a single or set of keywords or image will not necessarily diminish the investigative task. Therefore, the proposed system will tackle this challenge by applying additional knowledge to the retrieved images with the aim of enabling the investigator to ask and filter evidence using a wider range of information. Therefore, it is important to develop the DSS that can link the annotation and image feature alongside relevant metadata to enable investigators to ask higher-level more abstract questions of the data. Consequently, more investigation will be required to find the best correlation methods to provide the necessary functionality that required building a relation between various types. Moreover, additional work required to find the best way to present the results in

different models based on investigator requirements such as symbol map, node-link diagram, and chart.

Regarding to operational considerations that should be taken for the proposed system: big storage is essential in order to save images and their features (AF and metadata) in a database, addition to the retrieval results that also should be saved to review it again in any time. In addition, the use of publically available AIA or facial recognition systems results in benefitting from the latest developments of image analysis, without having to develop and manage the system, it does introduce the problem of submitting evidence to an external untrusted source for analysis. Therefore, pre-processing procedures should be introduce the necessary privacy are required.

## VI. CONCLUSIONS

Current forensic tools provide a basic level of analysis for multimedia-based content. With increased volumes of data to process and the timeliness of analysis often being key issues, specific tools need to be developed. The paper proposed a holistic multimedia architecture using a multi-algorithmic approach to enhance the power of final recognition result. Current stage of the research has managed to utilise multiple systems and fuse their results to recognise image contents and has identified that compared to the individual systems the fusion results are more promising. This has shown positive signs in terms of the feasibility of the proposed-system-achievement making the research goal attainable. Future paper(s) will discuss the completed results.

Future work however needs to investigate how the proposed system will be able to support the investigator in providing complex and high-level questioning of the resulting forensic data.

## ACKNOWLEDGMENT

This project is part of a PhD research currently being carried out at Centre for Security, Communications and Network Research (CSCAN), Plymouth University, U.K. The deepest gratitude and thanks to Baghdad University, Ministry Of Higher Education & Scientific Research and Higher Committee for Education Development in Iraq (HCED) for funding this PhD research.

## REFERENCES

- [1] A. K. Jain, B. Klare, and U. Park, "Face Matching and Retrieval in Forensics Applications," pp. 20–28, 2012.
- [2] A. Singh, "Exploring Forensic Video And Image Analysis." 2015.
- [3] H. Yuan and L. Ying, "Study on forensic image retrieval," in 2014 9th IEEE Conference on Industrial Electronics and Applications, 2014, no. 112, pp. 89–94.
- [4] R. B. Hanji and V. Rajpurohit, "Forensic Image Analysis - A Frame work," *Int. J. Forensic Comput. Sci.*, vol. 8, no. 1, pp. 13–19, 2013.
- [5] M. Al Fahdi, N. L. Clarke, F. Li, and S. M. Furnell, "A suspect-oriented intelligent and automated computer forensic analysis," *Digit. Investig.*, vol. 18, pp. 65–76, 2016.
- [6] Chao-Yung Hsu, Li-Wei Kang, and H.-Y. Mark Liao, "Cross-camera vehicle tracking via affine invariant object matching for video forensics applications," in 2013 IEEE International Conference on Multimedia and Expo (ICME), 2013, pp. 1–6.

- [7] C. Wen, D. Ph, and C. Yu, "Image Retrieval of Digital Crime Scene Images," pp. 37–45, 2005.
- [8] R. S. Choraś, "Texture Based Firearm Striations Analysis for Forensics Image Retrieval," in *Advances in Intelligent Systems and Computing*, vol. 184 AISC, 2013, pp. 25–31.
- [9] K. V Shriram, P. L. K. Priyadarsini, and A. Baskar, "An intelligent system of content-based image retrieval for crime investigation," vol. 7, pp. 264–279, 2015.
- [10] S. A. Gulhane and A. A. Gurjar, "Content based Image Retrieval from Forensic Image Databases," vol. 5, no. 3, pp. 66–70, 2015.
- [11] A. Aljarf and S. Amin, "Filtering and Reconstruction System for Gray Forensic Images," vol. 9, no. 1, pp. 20–25, 2015.
- [12] J. Lee, W. Tong, R. Jin, and A. K. Jain, "Image Retrieval in Forensics: Application to Tattoo Image Database," *IEEE Multimed.*, 2011.
- [13] J. Wu, X. Wang, and H. Xing, "Regional objects based image retrieval," in *2011 Chinese Control and Decision Conference (CCDC)*, 2011, pp. 1273–1277.
- [14] M. Mohammadpour and S. Mozaffari, "A method for Content-Based Image Retrieval using visual attention model," in *2015 7th Conference on Information and Knowledge Technology (IKT)*, 2015, pp. 1–5.
- [15] M. Shamsujjoha, M. S. Ahmed, F. Hossain, and T. Jabid, "Semantic modelling of unshaped object: An efficient approach in content based image retrieval," in *2014 17th International Conference on Computer and Information Technology (ICCIIT)*, 2014, pp. 30–34.
- [16] N. Gupta, S. Das, and S. Chakraborti, "Revealing What to Extract from Where, for Object-Centric Content Based Image Retrieval (CBIR)," in *Proceedings of the 2014 Indian Conference on Computer Vision Graphics and Image Processing - ICVGIP '14*, 2014, pp. 1–8.
- [17] N. W. U. D. Chathurani, S. Geva, V. Chandran, and V. Cynthujah, "Content-Based Image (object) Retrieval with Rotational Invariant Bag-of-Visual Words representation," in *2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS)*, 2015, pp. 152–157.
- [18] D. Zhang, M. Monirul Islam, and G. Lu, "Structural image retrieval using automatic image annotation and region based inverted file," *J. Vis. Commun. Image Represent.*, vol. 24, no. 7, pp. 1087–1098, Oct. 2013.
- [19] C. Jin and S.-W. Jin, "Automatic image annotation using feature selection based on improving quantum particle swarm optimization," *Signal Processing*, vol. 109, pp. 172–181, Apr. 2015.
- [20] S. H. Amiri and M. Jamzad, "Automatic image annotation using semi-supervised generative modeling," *Pattern Recognit.*, vol. 48, no. 1, pp. 174–188, 2015.
- [21] Y.-F. Huang and H.-Y. Lu, "Automatic Image Annotation Using Multi-object Identification," in *2010 Fourth Pacific-Rim Symposium on Image and Video Technology*, 2010, pp. 386–392.
- [22] M. Hidajat, "Annotation Based Image Retrieval using GMM and Spatial Related Object Approaches," vol. 8, no. 8, pp. 399–408, 2015.
- [23] Y. Xia, Y. Wu, and J. Feng, "Cross-Media Retrieval using Probabilistic Model of Automatic Image Annotation," vol. 8, no. 4, pp. 145–154, 2015.
- [24] J. Majidpour, E. Khezri, H. Hassanzade, and K. S. Mohammed, "Interactive tool to improve the automatic image annotation using MPEG-7 and multi-class SVM," in *2015 7th Conference on Information and Knowledge Technology (IKT)*, 2015, pp. 1–7.
- [25] S. Z. Li and A. K. Jain, *Handbook of Face Recognition*. 2010.
- [26] X. Xu, W. Liu, and L. Li, "Low Resolution Face Recognition in Surveillance Systems," vol. 2014, no. January, pp. 70–77, 2014.
- [27] N. Ramanathan and R. Chellappa, "Modeling Age Progression in Young Faces," 2006.
- [28] I. Kemelmacher-shlizerman, S. Suwajanakorn, and S. M. Seitz, "Illumination-Aware Age Progression," 2014.
- [29] Z. Li, U. Park, and A. K. Jain, "A Discriminative Model for Age Invariant Face Recognition," vol. 6, no. 3, pp. 1028–1037, 2011.
- [30] D. Sungatullina, J. Lu, G. Wang, and P. Moulin, "Multiview Discriminative Learning for Age-Invariant Face Recognition," 2013.
- [31] R. Singh, S. Member, M. Vatsa, and S. Member, "A Mosaicing Scheme for Pose-Invariant Face Recognition," vol. 37, no. 5, pp. 1212–1225, 2007.
- [32] L. A. Cament, F. J. Galdames, K. W. Bowyer, and C. A. Perez, "Face recognition under pose variation with local Gabor features enhanced by Active Shape and Statistical Models," vol. 48, pp. 3371–3384, 2015.
- [33] A. Asthana, T. K. Marks, M. J. Jones, K. H. Tieu, and R. Mv, "Fully Automatic Pose-Invariant Face Recognition via 3D Pose Normalization," 2011.
- [34] D. Yi, Z. Lei, and S. Z. Li, "Towards Pose Robust Face Recognition," pp. 3539–3545, 2013.
- [35] A. Nabatchian and M. Ahmadi, "An Efficient Method for Face Recognition under Illumination Variations," pp. 432–435, 2010.
- [36] S. Choi, C. Choi, and N. Kwak, "Face recognition based on 2D images under illumination and pose variations," vol. 32, pp. 561–571, 2011.
- [37] F. Bhat and M. A. Wani, "Elastic Bunch Graph Matching Based Face Recognition Under Varying Lighting, Pose, and Expression Conditions," vol. 1, no. 8, pp. 51–59, 2015.
- [38] M. Sultana, M. Gavrilova, and S. Yanushkevich, "Expression, Pose, and Illumination Invariant Face Recognition using Lower Order Pseudo Zernike Moments," 2014.
- [39] S. Liao, A. K. Jain, and S. Z. Li, "Partial Face Recognition: Alignment-Free Approach," vol. 35, no. 5, pp. 1193–1205, 2013.
- [40] U. Park and A. K. Jain, "Face Matching and Retrieval Using Soft Biometrics," vol. 5, no. 3, pp. 406–415, 2010.
- [41] J. C. Klontz, E. Lansing, and E. Lansing, "A Case Study on Unconstrained Facial Recognition Using the Boston Marathon Bombings Suspects The Boston Marathon Bombings - Investigation Timeline," pp. 1–8, 2013.
- [42] K. V Awalkar, S. G. Kanade, and D. V Jadhav, "A Multi-modal and Multi-algorithmic Biometric System Combining Iris and Face," 2015.
- [43] G. Sathish, "Multi-algorithmic IRIS Recognition," vol. 38, no. 11, pp. 13–21, 2012.

# Best of Two Worlds: Secure Cloud Federations meet eIDAS

Thomas Zefferer  
A-SIT Plus GmbH  
Seidlgasse 22  
1030 Vienna, Austria  
thomas.zefferer@a-sit.at

Dominik Ziegler  
Know-Center GmbH  
Inffeldgasse 13/6  
8010 Graz, Austria  
dominik.ziegler@tugraz.at

Andreas Reiter  
IAIK  
Graz University of Technology  
Inffeldgasse 16a, 8010 Graz, Austria  
andreas.reiter@iaik.tugraz.at

**Abstract**—The federation of information technology (IT) systems is a common approach to bundle capabilities and get the best results for all participants. Cloud computing and electronic identity (eID) are only two out of many domains, for which federated solutions have been a topic of scientific and corporate interest during the past years. Recently, the H2020 project SUNFISH has introduced a new cloud-federation approach called ‘Federation as a Service’ (FaaS). FaaS enables secure cloud federations, where data owners remain in full control of their data and workflows. In this paper, we identify shortcomings of the FaaS approach in terms of secure and reliable user authentication. In this sense, data security and protection mechanisms are only as good as the applied authentication measures. We solve this issue by proposing the integration of an existing pan-European federation of national eID systems into FaaS. This increases security guarantees of FaaS by using a trustworthy and liable eID solution that has a strong legal basis in the form of the EU eIDAS Regulation. A first successful implementation and deployment of the proposed solution demonstrates its feasibility and shows the great potential of combining federation solutions from the cloud domain and the eID domain.

**Keywords**—cloud computing, electronic identity, federation, security

## I. INTRODUCTION

The federation of IT systems has become an important topic in both research and industry. In general, a federation consists of different entities agreeing on a certain standard of operation and thus achieving interoperability. The advantages are apparent: by joining a federation, each federation member can benefit from functionality of all other members of the federation.

One area with a high potential to benefit from federation concepts is cloud computing. If multiple clouds are interconnected, i.e. federated, users from one cloud can benefit from the combined functionality of the entire federation. For instance, two cooperating companies could agree to federate their private clouds in order to benefit from each other’s infrastructure and data-processing capabilities. Despite its high potentials, the federation of clouds is a complex proposition, whose implementation raises several challenges, many of them related to data security. In the example given above, the two companies might certainly agree to exchange certain data with the other party, as part of using its infrastructure, but will

naturally refrain from granting universal access to all data processed in their own private clouds. Maintaining control of own data is hence the key challenge in federated cloud environments.

Overcoming this challenge has been the goal of the European Union (EU) funded research project SUNFISH<sup>1</sup>. SUNFISH has introduced the concept of Federation as a Service (FaaS). Its goal is to enable secure cloud federations that provide a higher degree of functionality by allowing the cross-cloud exchange of data, but are still able to meet security requirements of these data. The proposed FaaS concept and its implementation developed by SUNFISH achieve this.

The FaaS solution developed by SUNFISH bases on reliable policy-definition and policy-enforcement methods. These methods restrict data access to authorized users with the necessary assigned privileges. This approach requires users to be reliably identified and authenticated beforehand. Only if the identity of users is reliably verified by means of secure authentication schemes, policy-enforcement methods can work effectively. The FaaS solution developed by SUNFISH relies on the implicit assumption that end users are reliably authenticated within their home cloud-environment. Accordingly, all federation members build an implicit circle of trust and assume that all members act responsibly when authenticating its users. However, in certain scenarios, this assumption might be unrealistic. If a member cloud of the federation fails to authenticate reliably its users, restricting data access to users with certain roles or privileges is actually pointless.

In this paper, we propose a solution to this problem. Concretely, we propose to combine SUNFISH’s FaaS solution with the federation of European national eID solutions defined by the European Union’s eIDAS Regulation [11]. The eIDAS Regulation defines an interoperability solution to federate existing national eID solutions such as the Austrian Citizen Card [5], the Belgian eID card [2], or the Swedish eID [12]. This solution can be used to reliably identify and authenticate users across Europe. Backed by the EU eIDAS Regulation, this eID solution has a strong legal foundation. In addition, well-defined requirements and established governance processes guarantee an adequate level of security. The eIDAS-based federated eID solution can hence be regarded as ready-to-use

<sup>1</sup> <http://www.sunfishproject.eu>

building block that provides secure and reliable user identification and authentication. We propose to integrate this building block into SUNFISH's FaaS concept and implementation. We show how to combine the federation concepts behind FaaS and the eIDAS-based eID solution on conceptual and architectural level and demonstrate a working proof of concept.

Accordingly, the remainder of this paper is structured as follows. Section II provides relevant background information on SUNFISH's FaaS concept and the eIDAS-based eID federation. Section III then introduces in detail our proposal to combine eIDAS-based user authentication with SUNFISH's FaaS concept. Findings obtained during evaluation of the proposed solution are discussed in Section IV. Finally, conclusions are drawn in Section V.

## II. BACKGROUND AND RELATED WORK

This section elaborates on the two baseline technologies combined in this work. While Section II.A focuses on the research project SUNFISH and its FaaS concept, Section II.B introduces the EU eIDAS Regulation and its definition of a secure eID federation across Europe.

### A. SUNFISH: Secure Cloud Federations

SUNFISH is a H2020 project funded by the European Union with the goal to enable secure data sharing and federation of clouds [10]. SUNFISH ensures that data owners maintain control of their data and can define flexibly who is allowed to access data for which purposes and to which extent. For this purpose, SUNFISH has developed FaaS, an extended and flexible data-security governance model, which supports a variety of scenarios. FaaS relies on classical eXtensible Access Control Markup Language (XACML) [7] based approaches, which are already widely used today to enforce data-access control [1] [4].

The core of the XACML enforcement model comprises several well-defined entities including the Policy Enforcement Point (PEP), the Policy Decision Point (PDP), the Policy Administration Point (PAP) and the Policy Information Point (PIP). This separation of components yields a clear assignment of responsibilities with regard to policy enforcement. The PEP is the contact point for applications and issues access-decision requests to the PDP. The responses respectively the decisions are then enforced by the PEP. The policy store itself is logically separated into the Policy Retrieval Point (PRP) and the Policy Administration Point (PAP). The flow for a usual policy-decision request starts at the PEP, where the request is generated, and is passed on to the PDP. The PDP gathers potentially missing attributes from connected Policy Information Points (PIPs) and retrieves a list of matching policies from the PRP. The evaluation result is returned to the PEP where the decision is finally enforced.

FaaS, as introduced by SUNFISH and described in detail by Suzic and Reiter [9], extends this generic XACML approach with additional components. Furthermore, it provides a concrete implementation of the XACML approach for federated cloud environments and closes identified gaps of the XACML specification. Figure 1 illustrates the FaaS concept

developed by SUNFISH. Different cloud environments (each a member of the cloud federation) are modeled as so-called tenants. A special infrastructure tenant operates common services like the policy-decision service and the policy store. The Blockchain technology [13] [14] is used by the federation to guarantee integrity of stored policies.

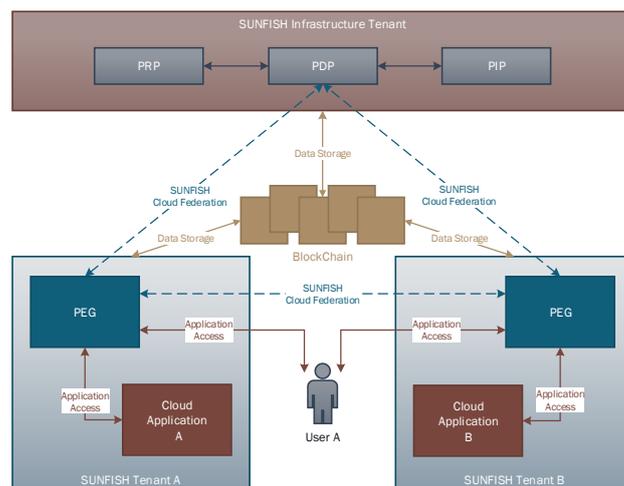


Figure 1. Basic building blocks of the FaaS concept.

All tenants dedicate computational resources to the federation, which are used to deploy and operate services and applications. The classical XACML model is extended by introducing the Policy Enforcement Gateways (PEG) entity. The PEG extends the responsibilities of the PEP by means of gateway functionality. PEGs are deployed at the edge of each tenant. Acting as a gateway, they protect the entire communication entering or leaving this tenant. The PEG analyzes the traffic and issues policy-decision requests to the PDP. Based on the policies defined by the data owner, a decision is derived, e.g., if an application from a certain tenant is allowed to access a service deployed in another tenant or if a user may access a particular service.

Summarizing, the FaaS concept developed by SUNFISH enables the federation of clouds while still leaving control of the data in the hands of the respective data owners. Apparently, the FaaS concept does not focus explicitly on the authentication of users. Instead, FaaS implicitly assumes that users are authenticated reliably in their respective home tenants and that user-specific policies can therefore be enforced reliably.

### B. eIDAS: Secure Identity Federations

The Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market (eIDAS Regulation) [11] is the legal basis for the provision of cross-border eID and trust services in the European Union. Although covering different kinds of trust services, the regulation has a strong focus on eID and on achieving interoperability between existing national eID solutions of EU Member States (MS). Essentially, the eIDAS Regulation defines an interoperability framework that enables European citizens to use an eID issued by MS X for

identification and authentication at a service provided by MS Y. Technical foundations of the eIDAS Regulation have been developed in several European large-scale pilots (LSPs) such as STORK<sup>2</sup> and STORK 2.0<sup>3</sup> [6].

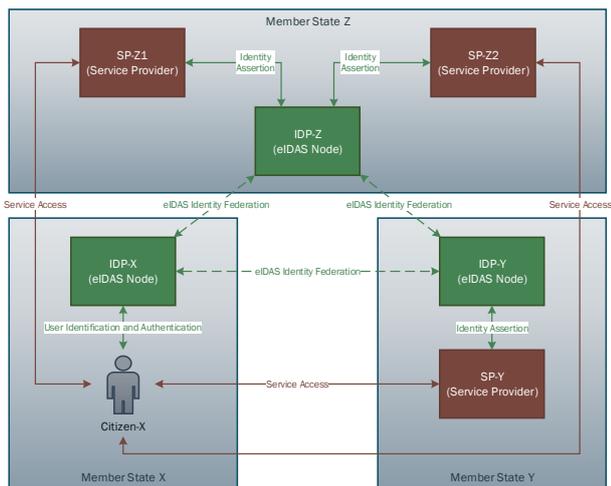


Figure 2. eIDAS-based eID federation.

The general architecture of the eIDAS interoperability framework is shown in Figure 2 by means of three exemplary EU MS. Each MS operates an own identity provider (IDP) denoted as eIDAS Node. Each national eIDAS Node is able to identify and authenticate citizens from the Node's own MS and accepts authentication requests from Service Providers located in the same MS. In order to allow citizens access to services in other MS as well, all eIDAS Nodes are federated, i.e. build a cross-border circle of trust.

The functionality provided by this eID federation is explained best by means of a concrete example. Assume that (according to Figure 2) a citizen from MS X accesses a service provided by MS Y, i.e. SP-Y. To identify and authenticate the citizen, SP-Y sends an authentication request to the eIDAS Node of MS Y, i.e. IDP-Y. As the citizen is from MS X, IDP-Y is unable to authenticate the citizen itself. Therefore, it sends an authentication request to the eIDAS Node of MS X, i.e. IDP-X. IDP-X identifies and authenticates the citizen, issues an identity assertion, and returns this assertion to IDP-Y. Due to the established eID federation, IDP-X and IDP-Y mutually accept issued identity assertions. This way, IDP-Y supplies SP-Y with a valid identity assertion, which is finally used by SP-Y to identify and authenticate the citizen from MS X.

In summary, the eIDAS interoperability framework assures that all European citizens can continue to use their existing national eIDs. At the same time, national IDPs only need to support their own national eID solution. Support for foreign eIDs is achieved by federating IDPs (eIDAS Nodes) of other MS. The technical interoperability framework defined by the eIDAS Regulation is currently being set up in EU MS.

### III. PROPOSED SOLUTION

The main problem tackled in this paper is the potentially weak realization of user authentication in FaaS-based cloud federations. This section introduces a solution to this problem by extending the concept of FaaS with secure user authentication provided by the existing eIDAS-based eID federation.

#### A. Relevant Use Cases and Requirements

Requirements to be met by the proposed solution have been derived beforehand from relevant use cases. Use cases and derived requirements are detailed in this section.

Identification of relevant use cases has been based on three general assumptions. First, we have assumed that each FaaS tenant features an IDP. Second, we have relied on the assumption that the IDP of a FaaS tenant is able to identify and authenticate users originating from this tenant. Third, we have assumed that applications deployed in a FaaS tenant can request the IDP of the same tenant to identify and authenticate users. All three assumptions can be regarded as realistic and comply with usual cloud deployments.

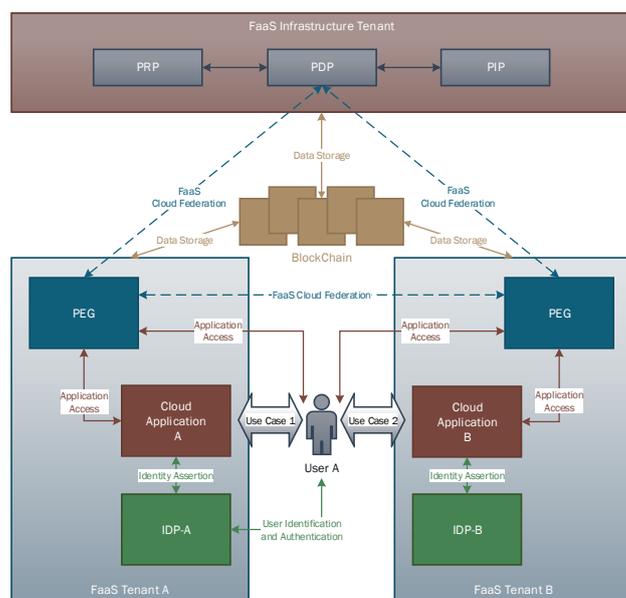


Figure 3. Relevant use cases

Considering the three assumptions made, yields the more detailed model of SUNFISH's FaaS concept shown in Figure 3. Note that Figure 3 still resembles Figure 1, i.e. shows the original FaaS concept. In addition, Figure 3 puts an additional focus on user authentication via IDPs. Figure 3 shows that two basic authentication scenarios can be distinguished, yielding the following two general use cases:

- **Use Case 1:** A user originating from Tenant A authenticates at a cloud application deployed in Tenant A. This means that user and application originate from the same FaaS tenant.
- **Use Case 2:** A user originating from Tenant A authenticates at a cloud application deployed in

<sup>2</sup> <https://www.eid-stork.eu>

<sup>3</sup> <https://www.eid-stork2.eu/>

Tenant B. In contrast to Use Case 1, user and application originate from different tenants in this use case.

The first use case resembles a classical setup of an identity management system consisting of a user, a service provider (application), and an IDP. The IDP authenticates the user on behalf of the service provider. In the context of SUNFISH’s FaaS-specific PEG component acting as gateway between the user and the application to be accessed. This component must be appropriately integrated into the authentication process.

The second use case is even more complex. In this use case, user and application originate from different tenants. Hence, there does not exist any IDP that is able to do both, receive authentication requests from the application and authenticate the user. Instead, two different IDPs are involved and need to interact in order to complete a successful user-authentication process. In addition, the same challenge as in Use Case 1 applies, i.e. the FaaS-specific PEG component needs to be integrated appropriately into the authentication process.

From the two sketched use cases, two general requirements can be derived for the proposed solution:

- **Requirement 1:** User identification and authentication functionality must be integrated into FaaS-based cloud federations such that FaaS-specific architecture components like the PEG are adequately addressed.
- **Requirement 2:** IDPs deployed in different tenants must be federated, in order to assure that users can identify and authenticate at applications deployed in other tenants.

Obviously, Requirement 2 suggests reliance on an established eIDAS federation solution. In Europe, such a solution is defined by the eIDAS Regulation and currently set up in EU MS. It is hence reasonable to employ the full potential of this approved solution and integrate it into the FaaS cloud federation model. The architecture that results from combining SUNFISH’s FaaS concept with the eIDAS-based eID federation is introduced in the following section.

*B. Architecture*

We have combined SUNFISH’s FaaS concept with the eIDAS-based eID federation according to the architecture shown in Figure 4. Note that in addition to architectural components, Figure 4 also shows implementation-related entities to support a better understanding. These entities are elaborated in more detail in Section C. For the sake of simplicity, the architecture shown in Figure 4 sketches only two MS participating in the eID federation. In reality, the federation spans all EU MS that have successfully set up the eIDAS-based interoperability framework.

Figure 4 illustrates how the proposed solution meets Requirement 1 defined in Section A. In order to combine the two federation approaches, the solution transfers the conceptual role of the service provider from the cloud application to the PEG. Accordingly, the PEG interacts with the corresponding

IDP and receives identity assertions issued by this IDP. This implies that the user authenticates at the PEG instead of the cloud application itself. Accordingly, the PEG can implement access-control mechanisms based on the user’s confirmed identity. As a downside of this approach, the cloud application no longer receives identity assertions from the IDP. Consequently, the cloud application cannot use identity information contained in the assertion to provide identity-based features or to apply identity-based access-control schemes itself. To remove this drawback, the proposed solution foresees that the PEG supplies the cloud application with required identity and role attributes. This way, the cloud application benefits from the user’s identity information without assuming the role of a service provider.

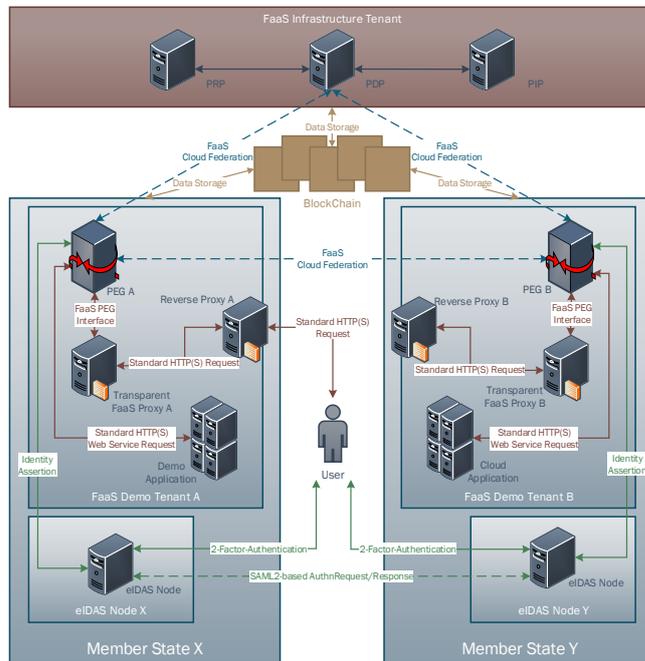


Figure 4. Architecture of the proposed solution

Requirement 2 defined in Section A is implicitly met by integrating the eIDAS-based eID-federation framework. Reliance on this framework and on eIDAS-compliant IDPs guarantees that users can authenticate at arbitrary cloud applications, independent from the tenant, in which the respective application is deployed. In theory, each PEG can access and use each IDP. However, in practice each PEG will most likely be able to communicate with one IDP only, i.e. with the eIDAS Node of the MS, in which the PEG itself is deployed. Note that this does not limit functionality, as all eIDAS-compliant IDPs are federated. Benefits of this federation also become apparent from the process-flow description provided in Section D.

*C. Implementation*

Before elaborating on the proposed solution’s basic process flow, we introduce a few implementation details to support a deeper understanding of the proposed solution. To evaluate its feasibility, we have implemented all relevant building blocks

of the proposed solution and deployed them in an evaluation environment. Details on evaluation results are provided in Section IV. Our implementation has been based on components developed by SUNFISH. The deployment spans across Europe by integrating the Austrian and Swedish eIDAS infrastructures with their already available eIDAS Nodes into the SUNFISH infrastructure. These countries have been chosen as their national eIDAS infrastructures are sufficiently set up. In addition to the eIDAS Nodes, each MS is assumed to host a computational FaaS tenant in this evaluation setup. An integral part of the FaaS model is the Infrastructure Tenant, thus also part of the evaluation setup. The Infrastructure Tenant hosts components related to policy evaluation and supports the PEGs of other tenants in enforcing defined policies.

In addition to the architecture, Figure 4 also shows some implementation-related components. When digging more into technical details, it becomes apparent that the PEG actually comprises the gateway components itself as well as two proxy components. First, the Reverse Proxy serves as a basic router to handle incoming HTTP(S) requests. This approach allows mapping multiple services to the same instance of the Transparent FaaS Proxy. Consequently, the Reverse Proxy can provide transport-layer security on a domain basis or load balancing for increased scalability and flexibility. Additionally, this entails that services can be migrated transparently to the FaaS model, without the need to change existing application configurations. Second, the Transparent FaaS Proxy serves as intermediate layer between legacy applications and the FaaS cloud-federation infrastructure. This enables an efficient integration of legacy applications into the FaaS infrastructure, without interfering with existing workflows. By aiming at a transparent deployment of services and by bridging the gap between legacy applications and managed cloud federations, a better acceptance of the overall system can be achieved.

D. Process Flow

To bring the introduction of the proposed solution down to a round figure, a typical user-authentication process is sketched in this section. Focus is put on Use Case 2 as defined in Section A, since this use case is regarded more complex and thus more challenging.

Use Case 2 assumes that a user (User Y) from MS Y wants to access a cloud application located in FaaS Demo Tenant A. It is further assumed that the user can only be authenticated by eIDAS Node Y, while the accessed cloud application can communicate with eIDAS Node X only. This setup requires IDP X (eIDAS Node X) and IDP Y (eIDAS Node Y) to be federated, in order to complete a successful user-authentication process and to enable identity-based access-control mechanisms.

The overall process flow for this scenario is illustrated in Figure 5. In the beginning, the user triggers the process by requesting access to a Demo Application. User Y uses a client that is unaware of the underlying FaaS infrastructure. Hence, the user enters the public URL of the Demo Application. The FaaS infrastructure intercepts the user's request using the Reverse Proxy in FaaS Demo Tenant A (1).

The Reverse Proxy forwards the intercepted request to the Transparent FaaS Proxy located in the same tenant. The Transparent FaaS Proxy acts as adapter for legacy (FaaS-unaware) applications and integrates them with the FaaS policy-enforcement infrastructure (2). The Transparent FaaS Proxy transforms the original request into a format suitable for the FaaS infrastructure (3). This request is then forwarded to the PEG deployed in FaaS Demo Tenant A (4). The PEG issues an authorization request to the PDP (5). As User Y is not authenticated at that time, the PDP denies access to the requested Demo Application (6). Consequently, the PEG requests an identity assertion from eIDAS Node X (7). The eIDAS Node X initiates the user-authentication process by displaying a selection screen of all available and supported eIDAS-compliant IDPs (8). User Y selects the preferred eIDAS node (i.e. eIDAS Node Y) (9). Based on this selection, eIDAS Node X requests an identity assertion from eIDAS Node Y using the eIDAS-based interoperability protocol (10).

The eIDAS Node Y authenticates User Y by means of the user's national eID (11) (12). After successful verification of User Y's identity, eIDAS Node Y issues an eIDAS-compliant identity assertion (13) and forwards it to eIDAS Node X (14). The eIDAS Node X transforms the received assertion into its national format (15) and returns the transformed identity assertion to the PEG (16). The PEG extracts required identity attributes from the assertion (17), but still needs to verify whether the provided attributes are sufficient to grant access to the Demo Application. Thus, the PEG again issues an authorization request to the PDP, this time including available user information (18). If the data provided meets all defined access policies, the PDP grants access to the requested Demo Application (19). In the last step, the PEG forwards the initial user request to the Demo Application located in FaaS Demo Tenant A and includes available user information (20). This completes the user-authentication process.

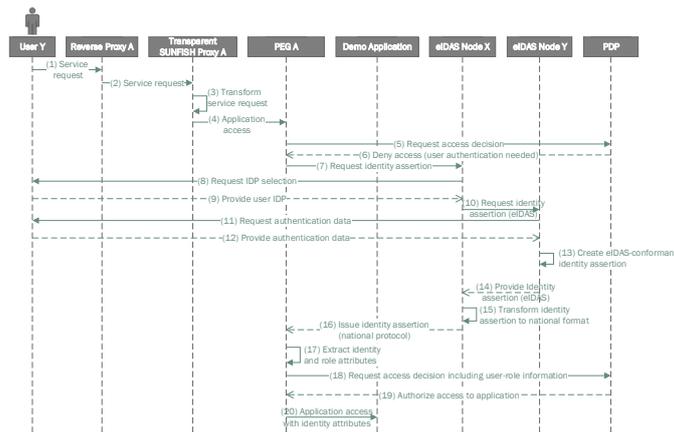


Figure 5. Process flow of Use Case 2

Note that the process flow for Use Case 1 is merely a subset of Use Case 2's process flow sketched above. The main difference is the lack of communication between IDPs, as in Use Case 1 the initially requested IDP is already able to authenticate the user. Thus, we do not elaborate on Use Case 1 in more detail at this point.

## IV. EVALUATION

We have conducted a first evaluation of the proposed solution to verify its feasibility and to learn its benefits and drawbacks. Therefore, we have created a working prototype implementation. The prototype has been deployed and operated in an evaluation environment during an evaluation period. As expected, several lessons have been learned during this evaluation period. From these lessons learned, we were able to derive benefits and open issues of the proposed solution.

The conducted evaluation has shown that the proposed solution has indeed several advantages over SUNFISH's initial FaaS concept. Overall, the proposed solution provides a *higher level of security*, as it takes away the responsibility of user authentication from members of cloud federations and assigns this task to a highly specialized and approved solution, i.e. the eIDAS-based eID federation. Since the eIDAS Regulation defines strict security requirements for eID solutions being part of this federation, the overall security of FaaS solutions is raised. Another advantage of the proposed solution is its *strong legal foundation* in the form of the eIDAS Regulation. This implies that responsibilities and liabilities are clearly defined on legal level, again leveraging the quality and reliability of the used eID federation. Its implementation and operation have shown that the proposed solution also provides the advantage of *easy integration*. Due to its generic architecture, the eIDAS-based eID solution integrates smoothly into the FaaS concept. Information exchange between these two worlds, i.e. cloud federation and eID federation, take place via well-defined interfaces only, yielding a neat integrated solution. Finally, evaluation has shown that the proposed solution is in *full conformance with SUNFISH's FaaS approach*. All concepts of FaaS and its underlying processing model can be left unmodified when integrating eIDAS-based user authentication. The eIDAS-based eID federation does not affect negatively functionality of FaaS-based cloud federations by any means.

In addition to these benefits, lessons learned during evaluation have also yielded some open issues. One challenge the proposed solution needs to face is the rather *slow take-up of the eIDAS Regulation*. Although the regulation is already in effect, its implementation in the EU MS is still an ongoing process. Accordingly, it cannot be assumed that all EU citizens already have assigned a compliant eID issued by their MS. However, it can be expected that the situation will improve in future, as MS continue to implement the eIDAS Regulation. Another issue to be considered is the existence of *decoupled eID systems*. In most corporate enterprises, proprietary eID systems are established and used. Users are not identified by means of their eIDAS-based eIDs (even if they have one), but by means of internal eIDs issued by the respective enterprise or institution. Integrating our proposed solution to such a scenario leads to two decoupled eID systems in place. In this case, means must be applied to map eIDAS-based eIDs to legacy eIDs used internally. Although such a mapping is expected to be technically feasible in most cases, it introduces additional effort.

Overall, the conducted evaluation has shown that the proposed solution is feasible and significantly improves the SUNFISH's original FaaS concept in several aspects.

## V. CONCLUSIONS

In this paper, we have proposed a solution that enhances SUNFISH's FaaS concept by integrating a highly secure and reliable pan-European federated eID solution, i.e. the eIDAS-based eID federation. We achieved an integration of eIDAS-based user authentication in full conformance with the original FaaS processing model. Our solution outsources the security-critical task of user identification and authentication to an approved solution that is backed by European law and applicable throughout Europe. This way, our solution eliminates a potential weakness of FaaS solutions, i.e. the weak implementation of user-authentication schemes, and enhances the security of SUNFISH's FaaS concept.

For future work, we plan to tackle remaining issues that we have derived from lessons learned during evaluation. In particular, we will work on solutions to problems that arise from the fact that the eIDAS-based eID solution might be decoupled from proprietary eID systems used in the respective members of the cloud federation. Despite these open issues to be addressed, we believe the proposed solution is mature enough to be integrated into FaaS-based cloud federations, in order to raise their level of security.

## REFERENCES

- [1] A. Bertolino, T. Y. Le, F. Lonetti, E. Marchetti, and T. Mouelhi. Validation of Access Control Systems. pages 210{233, 2014.
- [2] A. Fairchild and B. de Vuyst. The Evolution of the e-ID card in Belgium: Data Privacy and Multi-Application Usage. In The Sixth International Conference on Digital Society, pages 13{16, Valencia, 2012.
- [3] D. Hardt. The oauth 2.0 authorization framework. 2012.
- [4] S. Kasem-Madani and M. Meier. Security and Privacy Policy Languages: A Survey, Categorization and Gap Identification. CoRR, page 13, 2015.
- [5] H. Leitold, A. Hollosi, and R. Posch. Security Architecture of the Austrian Citizen Card Concept. In 18th Annual Computer Security Applications Conference, 2002. Proceedings, pages 391{400, 2002.
- [6] H. Leitold, A. Lioy, and C. Ribeiro. STORK 2.0 : Breaking New Grounds on eID and Mandates. Proceedings of ID World International Congress, (Idm):1{8, 2014.
- [7] B. Parducci and H. Lockhart. eXtensible Access Control Markup Language (XACML) Version 3.0. OASIS Standard, (January):1{154, 2013.
- [8] N. Sakimura, J. Bradley, M. Jones, B. de Medeiros, and C. Mortimore. Openid connect core 1.0. The OpenID Foundation, 2014.
- [9] B. Suzic and A. Reiter. Towards Secure Collaboration in Federated Cloud Environments. Workshop on Security, Privacy, and Identity Management in the Cloud, In Press, 2016.
- [10] B. Suzic, A. Reiter, F. Reimair, D. Venturi, and B. Kubo. Secure Data Sharing and Processing in Heterogeneous Clouds. Procedia Computer Science, 68(316):116{126, 2015.
- [11] The European Parliament and the Council of the European Union. REGULATION (EU) No 910/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC, 2014.
- [12] E. Wihlborg. Secure electronic identification (eID) in the intersection of politics and technology. International Journal of Electronic Governance, 6(2):143{151, 2013.
- [13] D. G. Wood. Ethereum: A Secure Decentralised Generalised Transaction Ledger. Sante Publique, 28(3):391{397, 2016.
- [14] Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. *Www.Bitcoin.Org*, 9. <https://doi.org/10.1007/s10838-008-9062-0>

## **Session 17: Infonomics and e-Technology**

Title: Venmo: Exposing a User's Lifestyle  
(Authors: Husna Siddiqui, Aspen Olmsted, Brendan Keane)

Title: Challenge Based Visual Speech Recognition Using Deep Learning  
(Authors: Philip McShane, Darryl Stewart)

Title: The Study of the Local Community Products (OTOP) Website Characteristics toward Buyer Decision using Eye Tracking  
(Authors: Nagul Cooharajanone, Krittika Akasarakul, Thipsuda Wongkhamdi, Phonkornkrit Pruetthiwongwanich, Kanokwan Atcharyachanvanich)

Title: Data visibility and trust enhancement of enterprise customers in cloud computing services  
(Authors: W.P.Yuen, K.B.Chuah)

# Venmo: Exposing a User's Lifestyle

Husna Siddiqui, Aspen Olmsted, Brendan Keane

Department of Computer Science

College of Charleston

Charleston, SC 29401

siddiquih@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—The focus of this work is to investigate the implications of the intersection of payment processing and social networking. Venmo is a peer-to-peer mobile payment service with a social twist. We analyze the information that is publicly available on Venmo. This allows us to see patterns of a user's day to day life, who they interact with, and what they spend money on. In this paper, we modify an existing Chrome extension, Money Trail, by adding a map reduce method to track common messages that are used in a user's transaction. Analyzing the occurrence of common words exposes a user's personal activities and budget habits.

**Keywords**- API; Venmo; security; payment; social engineering

## I. INTRODUCTION

In today's world, social networking is the heart of all communication. Media outlets are used to share personal information, professional development, photos, as well as to keep friends and family updated about one's travels and interests. [1] Now social media is also being used to share information about payment transactions. Venmo is a payment application that allows users to transfer money between one another. It provides users with a UI, user interface, to easily pay someone for things like rent, food, coffee and more. Venmo allows users to connect with their friends on Venmo, send payments, and charge other users with a Venmo account. [4] The defining quality of Venmo is the social media feature that encourages users to share their transactions with others. This aspect of Venmo is what makes it vulnerable to many security risks and social engineering attacks.

In this paper, we investigate the implications of openly sharing your payment transaction details. To do this, we use empirical analysis and common knowledge about the structure of the Venmo platform. Additionally, we use Money Trail, a chrome extension that creates visualizations of public data that can be accessed through Venmo's API, application programming interface. The organization of this paper is as follows. Section II will describe the works that are related to our research area. Section III explains the motivation for our research. Section IV details the underlying framework that we use as our foundation and describes our contribution. We conclude our work in Section V.

## II. LITERATURE REVIEW

Social engineering techniques can be utilized to manipulate people into giving up confidential information. Schmookler explains that in the cyber world, the weakest link in the security chain is the human who accepts a person or scenario at face

value. Security comes from the knowledge of knowing who and what to trust. The popularity and mass use of social networking sites, as well as its security concerns, have made these sites an avid candidate for the research community. [6] A common problem that occurs on Venmo is that users accidentally pay the wrong person. There is an obvious UI issue here. Users are unable to distinguish friends from non-friends. [2] Moreover, Venmo is not prompting users to notify them that they are sending money to a user who is not a friend which could help prevent an erroneous scenario.

In the paper, Making it Rain with Cloud Payment Processing, Piazza brings to light an issue with the lack of proper documentation for CardConnect, a payment processing platform. Following CardConnect's incomplete documentation leads developers to introduce security vulnerabilities in their organization's system. Piazza used Siteground to host a standard Magento install with CardConnect's Magento plugin. During the checkout process, a user can view the form items that are sent to CardConnect and create a request that mimics a successful transaction. This is because a client-side dependent model allows for trivial manipulation of the form data. [3] CardConnect failed to properly document that the request must be sent back to a web service for authentication before a transaction can be successfully processed.

Kraft, Mannes, and Moldow performed a security audit on Venmo to detect technical and social vulnerabilities. They exploit the Venmo API to find bugs that allow them to steal money from others, and trick people into sending money to them. The authors found that the Venmo API contains a call that publicly leaks information about transactions that are set to be shared only with friends. [2] This violates a user's assumptions that their transaction information is only being shared with a known audience. The authors highlight that Venmo users can update their username and other profile information at any time. This means that a user with malicious intent can update their username to persuade someone that they are someone else. [2] They also emphasize that the lack of visual distinction between charges from friends versus non-friends makes it easy for a malicious user to attempt social engineering attacks on the platform.

Khanna created a Chrome extension, Money Trail, which can identify relationships between users on Venmo and how much time they spend together. It can also identify members of private social organizations, attendees of private events, and a

users' food purchases. [3] "After a user logs into Venmo.com and navigates to any Venmo user's newsfeed page, the script runs in the background to detect the source of the historical transaction data populating the newsfeed. It then requests a copy of that historical transaction data and loads it into the visualizations." [3] These visualizations allow users to easily discover how much data they are publicly exposing on Venmo.

### III. MOTIVATION

Venmo's popularity has skyrocketed in the past year. The company processed an astounding \$17.6 billion in payments last year. This is more than double the volume from 2015. [4] With this high increase in growth, Venmo becomes an avid candidate for the research community. Being that Venmo is a payment application, the most crucial aspect of the application is security. By default, Venmo sets all transactions as public. This information includes the names of the payer and the recipient, date, and a transaction message. Analyzing this data can give insights into a user's financial and social life.

```
// Gets and parses json data on user
// transaction from hitting venmo feed endpoint
function getRestMessages(endpoint) {
    next = endpoint;
    background_reqs++;
    console.log("requesting "+background_reqs);
    $.ajax({
        type: "GET",
        url: endpoint+"?limit=50",
        processData: false,
        success: function(json) {
            var stringArray = new String("");

```

Figure 1. getRestMessages() function

### IV. OUR CONTRIBUTION

Our implementation adds value to the existing Money Trail script, by exposing significant details that are publicly available to anyone that has a Venmo account. Fig. 1 shows a snippet of our getRestMessages() function. We augment this function to collect information on recurring payment words using map-reduce. The getRestMessages() function uses an asynchronous AJAX call to hit the following public endpoint: <https://venmo.com/api/v5/users/feed>. The only two parameters required are a Venmo user id and the limit for how many transactions we want to return.

```
// map reduce the stringArray into counts of each word
var counts =
    stringArray.replace(/[\s]/g, "").split(/\s+/).reduce(function(map, word){
        map[word] = (map[word]||0)+1; return map;
    }, Object.create(null));
```

Figure 2. Map reduce the string array created from the response object

If the asynchronous call returns successfully, then a Javascript object is returned with data about the user's transactions. We modify the function first by building an HTML table with column headers for future use. Then, we build one string of all the returned messages in the JSON objects by

concatenating each message into a variable called stringArray. Common irrelevant words are scrubbed out, and a word must occur more than once to be recorded. As Fig. 2 shows, after the full single string of messages is built, a hashmap is created by performing a map reduce on the string while keeping track of how many times each word occurs in all the messages returned. Now that the hash map of each word's occurrence count is built, the program iterates through the hash map and outputs each word and count to the HTML table. This table is prepended to the Venmo webpage.

Table I. Sample Keyword Mapping

Word	Occurrences
Rent	4
Power	8
Hotel	2
Uber	2
Haircut	3
Tip	2
Drinks	4
Wifi	4

### V. CONCLUSIONS

In this paper, we contribute to the Money Trail extension, by adding a map reduce function to track the use of keywords used in a Venmo user's transaction history. Table I shows a sample dataset of a user's keywords and the number of occurrences in their Venmo history. This simple information can give us insights into a user's personal activities and their spending habits. Recurring utility payments including water, the internet, and electricity bills to another Venmo user can indicate and expose a user's fellow roommate or close friend. This information can potentially be used for social engineering attacks.

### REFERENCES

- [1] E. M. J. M. Ben Kraft, "Security Research of a Social Payment App," 2014.
- [2] A. Khanna, "Venmo'ed: Sharing Your Payment Data With the World," 29 10 2015. [Online]. Available: [techscience.org](http://techscience.org). [Accessed 10 06 2017].
- [3] M. P. a. A. Olmsted, "Making it Rain with Cloud Payment Processing Vulnerabilities," in *IEEE Xplore Digital Library*, 2017.
- [4] L. Gensler, "Forbes," 14 02 2017. [Online]. Available: <https://www.forbes.com/sites/laurengensler/2017/02/14/venmo-customer-service/#1d8f0d131cfd>.
- [5] A. O. Z. D. Husna Siddiqui, "Engineering your social network to detect fraudulent profiles," in *Information Society (i-Society)*.
- [6] Schmoekler, S.L. (2015, 12). Lexology. Retrieved 3 2016, from <http://www.lexology.com/library/detail.aspx?g=e4646c20-f27a-495d-a4a0-df4d4741610>

# Challenge Based Visual Speech Recognition Using Deep Learning

Philip McShane  
EEECs  
Queens University Belfast  
Belfast UK

Darryl Stewart  
EEECs  
Queens University Belfast  
Belfast UK

**Abstract-** We present a novel approach to liveness verification based on visual speech recognition within a challenge-based framework which has the potential to be used on mobile devices to prevent replay or spoof attacks during Face-based liveness verification. The system uses model visual speech recognition and determines liveness based on the Levenshtein Distance between a randomly generated challenge phrase and the hypothesis utterances from the visual speech recognizer. A Deep learning-based approach to visual speech recognition is used to improve upon the state of the art for the use of visual speech recognition for liveness verification.

## I Introduction

Alternatives to the use of passwords are increasingly being considered as means of securing access to electronic devices such as laptops and phones. The most common approaches towards user authentication for gaining access to these devices make use of passwords, user IDs, identification cards and PINS. These techniques have a number of limitations: Passwords and PINs can be guessed, stolen or illicitly acquired by surveillance or brute force attack. There have been many high-profile hacks emanating from password breaches in recent times. These hacks allow malicious individuals to gain access to a system using the credentials of a valid user without the user being present.

In order to enhance security, alternatives to the password-based approaches have been considered and these have primarily been focused on forms of Biometric authentication. A number of different biometrics have been proposed, with

the most popular involving recognition of the Face [1], Voice [1] or Fingerprint [1][2]. These systems, while more secure than passwords, also have some limitations. Fingerprint scanning systems are accurate, fast and robust, however, they can be susceptible to forms of ‘spoofing’ whereby false fingerprints, can be used to fool the sensor [2]. A further limitation is the additional cost of having a dedicated fingerprint sensor within the device means that few devices have offered fingerprint scanning as an authentication process.

Speech recognition systems can be deployed inexpensively and universally to all mobile device types as they use only the standard microphone in the device. Voice has been shown to be highly accurate and reasonably robust in quiet environments. The performance can be affected by the presence of loud and/or time-varying background noises. Furthermore, in some environments, it may be considered inappropriate or indiscrete to speak clearly into a microphone.

Face recognition has been shown to be highly accurate and can be robust to changes in the user's environment, appearance, variations in pose and illumination conditions. A key concern with face recognition systems is that they may be susceptible to spoofing attacks where an unauthorized user holds a photograph in front of the camera and gains access as the person in the photo [3]. These forms of attack are more likely to be successful in the unsupervised, remote access use cases involving mobile devices. The security of remote unsupervised face recognition systems would be significantly improved by ensuring that “liveness” detection is included in the authentication process, thereby ensuring that the authorized user is present and responds when prompted for input by the system.

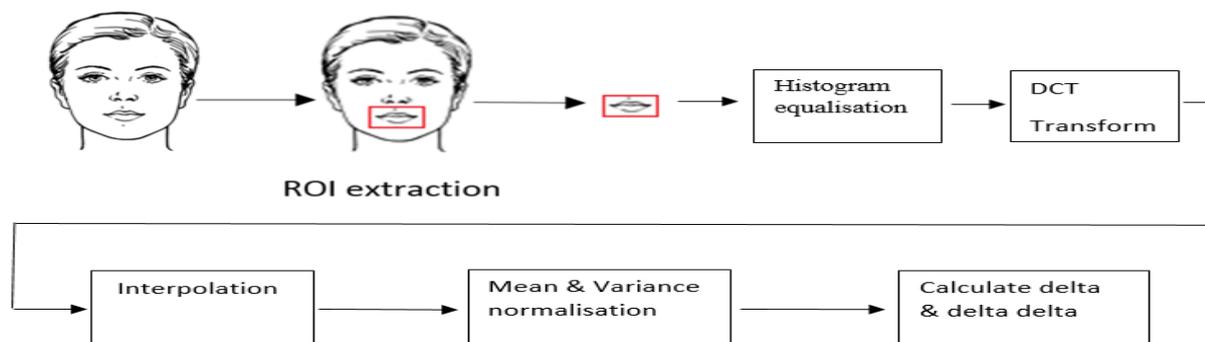


Figure 1. The feature extraction process

In this paper, a means of liveness verification based on visual phrase verification algorithm which uses a visual speech recognition system within a challenge-based verification framework. Specifically, the process of verification involves the user being challenged to say a randomly generated string of digits which they will then speak into the phone's camera. Visual speech recognition will be performed on the video and if the visual recognition system is confident that the video contains lip movements which match the challenge phrase, then the 'liveness' of the user will be verified. The challenge phrases are randomly generated at each verification attempt to limit the possibility of replay attacks using previously recorded videos.

For practical use, this approach to liveness verification would be combined with other biometric authentication processes such as face recognition in order to improve the overall security and robustness of the biometric system and would not inconvenience the user significantly beyond the standard face capture process.

Visual speech recognition has been the focus of extensive research in recent years and has matured to the point that it can be used robustly for limited vocabulary tasks [4][5]. Prior research on the use of visual speech recognition for biometric applications have focused on the use of the visual information combined with audio [6][7] and most of the research has focused on using visual speech as an alternative means of verifying the user's identity not for verifying liveness. Evano and Besacier [8] investigated liveness verification based upon an analysis of the synchronicity of visual and audio features and reported an Equal Error Rate of 14.5% using the XM2VTS dataset. In [10] a liveness verification system based on only using visual information was proposed that is based on speech recognition with an SVM (support vector machine) to recognize digits that had been individually segmented. A speech recognition rate of 68% was reported on the XM2VTS dataset, using the approach in [10] with only the visual modality. In this paper, the aim is to show an improvement over previous works through the use of deep learning.

## II. Visual speech recognition

Visual speech recognition aims to determine the text spoken by an individual based on the movement of their lips. When a visual speech recognition system receives a video the first step to performing recognition is first to determine where in the images the lips are located and to extract the lip region to be used as the region of interest (ROI). For the system that is used in this paper, the Dlib image processing library was used [9]. Dlib provides a facial landmark detector that has been used to locate and extract the ROI from each video frame, this process is described in [9].

Once the frames have been extracted from the video histogram equalization is applied to it in order to make the system more robust to changes in illumination. The frame is then converted to grayscale and then down sampled to a 16x16 image. After this, a DCT (discrete cosine transform) transform is applied to the frame.

The DCT transform was chosen as it was shown to give good performance in [5]. A triangular mask is then applied to the result of the DCT transform and from this the 15 lowest frequency coefficients are selected with the DC component being removed, leaving 14 DCTs. The DC component is removed as initial experiments showed that the system performed better when the DC component was not present. Mean and variance normalization is then applied to the feature vectors. The number of features is increased through cubic spline interpolation to 100 fps, as this was found to increase the performance of the visual speech recognizer. From the 14 DCTs, differential and acceleration coefficients are calculated. These are then concatenated with the 14 DCTs to give a feature vector of 42 coefficients.

Deep learning approaches have shown promise in solving problems in areas such as computer vision [11] [12], audio speech recognition [13] and natural language processing [14]. In order to create a visual speech recognition system that is capable of performing to a level comparable with audio based speech recognition software a deep learning based approach was chosen. By incorporating such an approach, the aim is to produce a system that would be suitable for real-world applications.

For this work, we have employed a hybrid system for performing visual speech recognition. The term hybrid refers to a speech recognition system in which a DNN (deep neural network) and HMM (hidden Markov model) are combined [15]. The DNN is used to provide the posterior probability estimates for the HMM states. The HMM models the long-term dependencies needed to take account of the temporal dimension of speech. For this work, we employed a DNN-HMM trained on DCT features. The use of DNN-HMM recognizers has shown significant improvement in the performance of speech recognition systems over prior approaches [13][16]. The architecture of the DNN can be seen in Fig. 2.

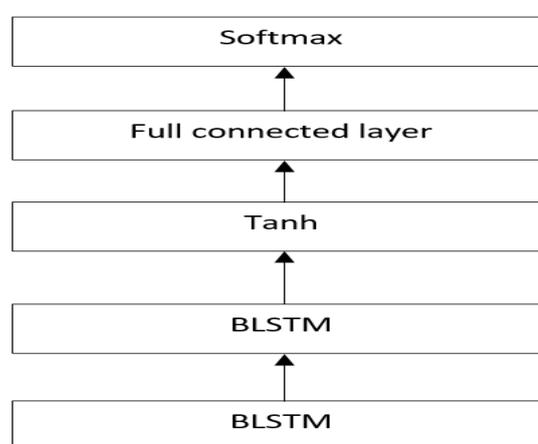


Figure 2. DNN architecture

Prior to training the DNN a DBN (deep belief network) of stack RBMs (restricted Boltzmann machines) was trained. This process is used to initialize the parameters of the hidden layers in the DNN. This is done via a greedy layer-wise procedure with each RBM trained and then stacked to produce the DBN. The RBM's are trained via approximate stochastic gradient descent[[kaldi]. After this pre-training step, the DNN is trained using sMBR (state level minimum Bayes risk) sequence discriminative training as this is suggested as the best criteria for sequence discriminative training in [17] [18].

### III. The Liveness verification algorithm

The output of a speech recognition system is a single highest-likelihood hypothesis phrase and the performance of a recognition system is commonly measured by performing recognition on a set of test utterances and calculating its average Word Error Rate (WER) [19]. WER for a single test utterance is calculated as:

$$WER = \frac{S + D + I}{N} \tag{1}$$

Where  $S$  is the number of substitution errors found in the hypothesis phrase,  $I$  is the insertion errors,  $D$  is deletions errors and  $N$  is the total number of words in the correct transcription.  $S$ ,  $D$  and  $I$  are determined through the use of dynamic programming during the calculation of the Levenshtein distance between the correct transcription of the spoken utterance and the hypothesis phrase.

Ideally, when a speaker says the challenge phrase the result of visual speech recognition would be a perfect match but visual speech recognition is not yet perfect and typically may operate at WERs of between 10% - 40% depending on the user and the quality of the video provided. Therefore, given that a recognition system will operate at a certain average WER, it seems plausible that if a challenge phrase of sufficient length is compared to the output of the recognizer and the Levenshtein distance is within an Acceptable Levenshtein Distance (ALD) threshold then it could be postulated that the challenge phrase was probably spoken as opposed to a random phrase and therefore liveness could be verified. Given this setup, the probability of a successful spoofing attack using a video containing the correct number of random digits can be expressed as in Equation 2.

$$P = \frac{1}{w} \cdot \frac{v^\epsilon}{w} \cdot \binom{w}{\epsilon} \tag{2}$$

Where  $P$  is the probability of a match being found with a challenge phrase containing  $w$  words chosen from a vocabulary of  $v+1$  word types and where the system allows  $\epsilon$  errors. Taking a specific example, the probability of a random digit string video being used successfully for a spoof attack where  $w=20$  and  $\epsilon=12$  is  $3.4 \times 10^{-10}$ . Therefore, while ideally the  $\epsilon$  would be kept as low as possible, even where the recognizer is not completely accurate.

### IV. Lattice-based Phrase Verification.

Aside from the single highest-likelihood hypothesis phrase, it is also possible to generate an N-best list of phrase hypotheses ranked according to their likelihoods from what is known as the recognizer's *search lattice*. The N-best list typically includes phrases which are plausible slight variations of the highest ranked hypothesis. For example, if a user was challenged to say the following phrase:

“one two seven three nine zero eight six four”

TABLE 1: 3-best list

Rank	Hypothesis
1	one two six three nine six eight six six
2	one two seven three nine zero eight six six
3	one two three seven zero eight six four five

then the resulting 3-best list might be as in **Error! Reference source not found.**

As can be seen in this example, the second-ranked hypothesis contains fewer errors than the best hypothesis and it is not unusual for the correct transcript or a close match to it to be found elsewhere within the N-best list rather than at the very top. The maximum length of an N-best list is primarily determined by the beam width and other pruning parameters during recognition but in practice, the correct phrase is generally found close to the top and in our experiments always within the top 50 phrases. Therefore, we allow the system to perform phrase verification with each of the hypothesis phrases in the top 100 phrases and if any of the phrases are verified based on the search of the N-best list, then the liveness is determined to be positive. This potentially allows the ALD threshold to be reduced slightly leading to a reduction in False Rejection Errors (FRR).

### V. XM2VTS dataset

For the experiments, the XM2VTS dataset [20] was chosen. The XM2VTS dataset was created for research into the multimodal identification of human faces. It has also been used for biometrics research into the use of voice for the purposes of liveness verifications. [8] [10] It has been chosen in preference to other possible datasets, as a liveness verification system is used as an addition to other biometric systems and by using a dataset used for that task a better understanding of how the proposed system would enhance the security of existing biometric systems is shown.

The XM2VTS dataset is a multi-model dataset comprised of 295 speakers saying the phrases “zero one two three four five six seven eight nine”, “five zero six nine two eight one three seven four” and “Joe took father’s green shoe bench out”. The focus of our experiments is on digit recognition by visual speech recognition so only the digit string phrases have been used. The data is split between training and testing data based on the Lausanne protocol [21]. This protocol divides the dataset into training and test for the training and testing of biometric systems. The protocol specifies two distinct configurations for the dataset. We use Configuration II of the protocol as the starting point for selecting our training and test data. Specifically, we selected the videos from the 70 speakers in the test partition as our test data for the speaker independent liveness verification experiments. The videos from the speakers that are not in the test set are used when training the recognizer. As none of the videos from the speakers present in the training set were used for our experiments the results reported indicate how the system would perform under speaker independent conditions and are therefore a good indication of how the system would perform when presented with data from new speakers, as would occur when such a system would be deployed for practical use.

The two 10-digit sequences were combined within one video to give the 20-digit phrase “zero one two three four five six seven eight nine five zero six nine two eight one three seven four”. Only the 20-digit videos were used during training of the recognizer. Using this model, a word accuracy of 86.3% was obtained using the 20-words videos. To allow for investigation into the effect of varying the length of challenge phrases, we segmented the videos in the test set to generate new videos from the test data which contained digits strings of 6, 10 and 15 digits.

This was achieved by segmenting the 20-digit videos into videos containing shorter phrases based upon word boundaries for each digit in the video. These were obtained by performing forced alignment of the audio from the videos using a highly accurate (99% word accuracy) audio-based speech recognizer. A variety of phrases were generated using these boundaries by moving a window of size  $w$  one word at a time over the 20-digit phrase. As a result of generating the videos based on this approach, it was possible to expand the number of phrases that were used in our experiments. The variety of phrases can be seen by looking at the first few 10-digit videos generated from the original 20-digit videos were “zero one two three four five six seven eight nine”, “one two three four five six seven eight nine five”, “two three four five six seven eight nine five zero” etc. While running the experiments, each video was tested as a possible spoof attack case and as a valid user test. The spoof attacks were set up as 1000 random challenge phrases of the correct length containing digit strings that did not match the actual content of the video were created. This simulates the possibility of an attack where the attacker poses a video of the correct user saying a phrase different to the one the system prompts the user to say.

VI. Experiments

Experiments were conducted using the visual speech recognizer on videos containing different length phrases. For practical use, a shorter phrase is preferable as it would take less time for a user to say, however, a longer phrase might be desirable where a stronger level of security is required.

Chart 1. FAR/FRR for 6-word phrases

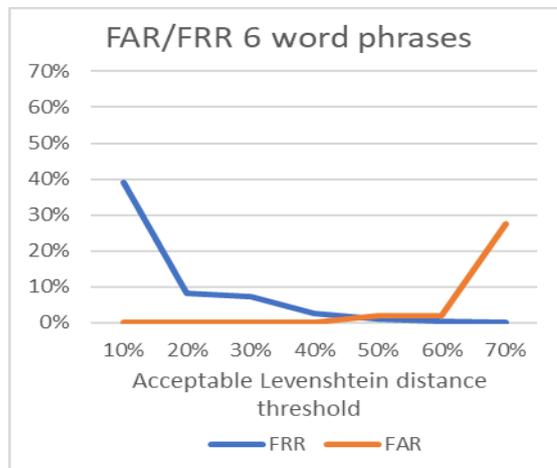


Chart 2. FAR/FRR for 10-word phrases

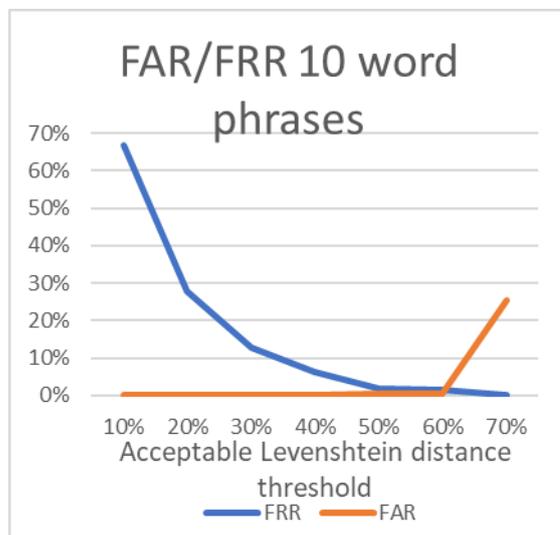


Chart 3. FAR/FRR for 15-word phrases

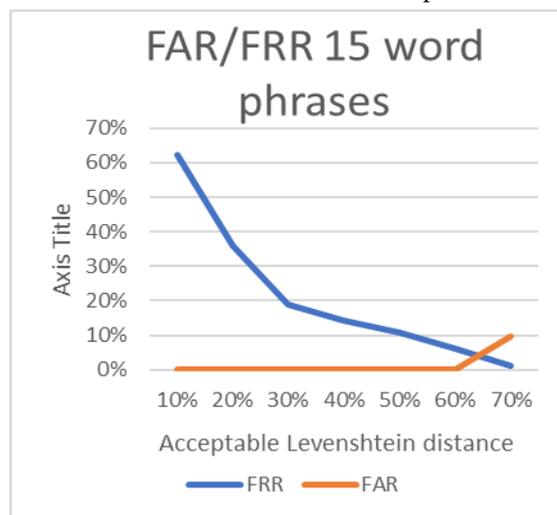
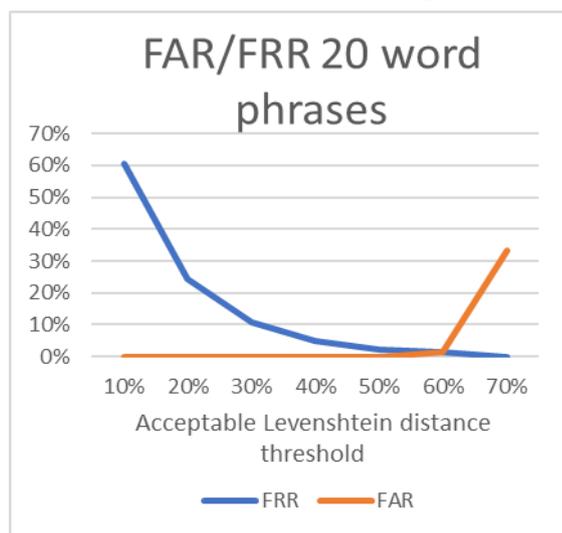


Chart 4. FAR/FRR for 20-word phrases



The average durations for the videos of 6, 10, 15 and 20 words in length were 2, 4, 6 and 8 seconds respectively. The results of the experiments can be seen in charts 1-4. These charts show the FRR and FAR (false acceptance rate) when the ALD threshold is set to different values. It is shown that the FRR stays below one percent even with ALD thresholds as high as 40%.

## VII. Future work

In this work, the use of deep learning based visual speech recognition as the basis for challenge-based liveness verification has been investigated. The performance of the system on a variety of phrase lengths has been shown and the appropriate ALD thresholds for the different phrase lengths are indicated. Future work will look at improving the performance of the visual speech recognition system and how to make it more robust to noise that such system would encounter when used in real-world conditions.

## REFERENCES

- [1] A.K. Jain, A. Ross, S. Prabhakar "An introduction to biometric recognition" IEEE Trans. Circuits Systems Video Technol., 14 (1) (2004), pp. 4–20.
- [2] C. Roberts, "Biometric attack vectors and defenses," Computers and Security, vol. 26, no. 1, pp. 14–25, 2007.
- [3] B. Biggio, Z. Akhtar, G. Fumera, G. L. Marcialis, F. Roli, "Security evaluation of biometric authentication systems under real spoofing attacks", IET Biometrics, vol. 1, no. 1, pp. 11-24, Mar. 2012.
- [4] S. Dupont and J. Luetttin, "Audio-visual speech modeling for continuous speech recognition", IEEE Transactions on Multimedia, vol. 2, no. 3, pp. 141-151, 2000.
- [5] A. Pass, J. Zhang, and D. Stewart, "An investigation into features for multi-view lipreading," in Proc. 17th IEEE Int. Conf. Image Process., Sep. 2010, pp. 2417–2420.
- [6] G. Chetty and M. Wagner, "Biometric person authentication with liveness detection based on audio-visual fusion", IJBM, vol. 1, no. 4, p. 463, 2009.
- [7] M. Alam, M. Bennamoun, R. Togneri, F. Sohel, "A Joint Deep Boltzmann Machine (jDBM) Model for Person Identification Using Mobile Phone Data", Multimedia IEEE Transactions on, vol. 19, pp. 317-326, 2017, ISSN 1520-9210.
- [8] N. Eveno, L. Besacier, A speaker independent "liveness" test for audio-visual biometrics, in 9th European Conference on Speech Communication and Technology (Lisbon, 4–8 September 2005).
- [9] Kazemi, Vahid, and Josephine Sullivan. "One-millisecond face alignment with an ensemble of regression trees." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1867-1874. 2014.
- [10] E Benhaim, H Sahbi, and G Vitte, "Designing relevant features for visual speech recognition, " in Acoustics, Speech, and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, 2013, pp. 2420-242.
- [11] Y. LeCun, K. Kavukcuoglu, C. Farabet et al., "Convolutional networks and applications in vision." in ISCAS, 2010, pp. 253–256.
- [12] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, ser. CVPR '14. Washington, DC, USA: IEEE Computer Society, 2014, pp. 1701–1708.
- [13] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in 2013 IEEE international conference on acoustics, speech and signal processing. IEEE, 2013, pp. 6645–6649.
- [14] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuska, "Natural language processing (almost) from scratch," Journal of Machine Learning Research, vol. 12, no. Aug, pp. 2493–2537, 2011.
- [15] N. Jaitly, P. Nguyen, A. Senior, and V. Vanhoucke, "Application of pre-trained deep neural networks to large vocabulary speech recognition," in Proc. INTERSPEECH, September 2012.
- [16] R. Seymour, D. Stewart, and J. Ming, "Comparison of Image Transform-Based Features for Visual Speech Recognition in Clean and Corrupted Videos", EURASIP Journal on Image and Video Processing, vol. 2008, pp. 1-9, 2008.
- [17] B. Kingsbury, "Lattice-based optimization of sequence classification criteria for neural-network acoustic modeling," in Proc. IEEE, ICASSP, April 2009, pp.3761–3764.
- [18] B. Kingsbury, T. N. Sainath, and H. Soltau, "Scalable minimum Bayes risk training of deep neural network acoustic models using distributed Hessian-free optimization," in Proc. INTERSPEECH, September 2012.
- [19] K. Zechner and A. Waibel, "Minimizing word error rate in textual summaries of spoken language", Proceedings of the 1st North American chapter of the Association for Computational Linguistics Conference, p.186-193, April 29-May 04, 2000.

[20] K. Messer, J. Matas, J. Kittler, and J. Luetin, "Xm2vtsdb: The Extended M2VTS Database," Proc. Second Int'l Conf. Audio- and Video-Based Biometric Person Authentication (AVBPA '99), pp.72-77, 1999.

[21] J. Luetin and G. Maître, "*Evaluation Protocol for the XM2FDB Database (Lausanne Protocol)*", Communication 98-05, IDIAP, Martigny, Switzerland, 1998.

# The Study of the Local Community Products (OTOP) Website Characteristics toward Buyer Decision using Eye Tracking

Nagul Cooharajanane

Department of Mathematics and Computer Science  
Faculty of Science, Chulalongkorn University  
Bangkok, Thailand  
e-mail: nagul.c@chula.ac.th

Thipsuda Wongkhamdi

Technoprenurship and Innovation Management Program  
Graduate School, Chulalongkorn University  
Bangkok, Thailand  
e-mail: Thipsuda.W@student.chula.ac.th

Kanokwan Atcharyachanvanich

Faculty of Information Technology,  
King Mongkut's Institute of Technology Ladkrabang  
Bangkok, Thailand  
e-mail: kanokwan@it.kmitl.ac.th

Krittika Akasarakul

Department of Mathematics and Computer Science  
Faculty of Science, Chulalongkorn University  
Bangkok, Thailand  
e-mail: Krittika.ak@student.chula.ac.th

Phonkornkrit Puetthiwongwanich

Department of Mathematics and Computer Science  
Faculty of Science, Chulalongkorn University  
Bangkok, Thailand  
e-mail: Phonkornkrit.pr@student.chula.ac.th,

**Abstract**— Thai government has launched a policy to encourage local community people to sell their products (OTOP) through e-commerce. One Tambon One Product or OTOP was derived from the concept of One Village One Product (OVOP) in Japan. Currently, people sell their products through the e-commerce web portal such as Facebook rather than having their own OTOP official websites. Therefore, this study aims to study and develop the prototype of the official website. The results show that social influence, perceived ease of use, reliability of website and reliability of product have a positive relationship to the intention to buy a product. Researchers further investigated using eye tracking device to analyze the participants' purchasing behavior. The results show that participants spent time looking at images on the center of the homepage, the product menu bar at the bottom of the center of the homepage, and the components from the product page, respectively. Moreover, participants did not spend time looking or even clicking the social network icon such as Facebook. The results of this study and implications are further discussed.

**Keywords**—Rural sub-district; Local Community Enterprise; OTOP; SMEs; Eye tracking; Web page.

## I. INTRODUCTION

Since the Ministry of Digital Economy and Society has a policy to develop the country to be the learning society, the ministry has established the ICT learning center in various communities across the country. In each sub-district, there is at least one learning center providing knowledge for the public and private sectors in order to use ICT effectively, using the knowledge to generate income for people in the community [1].

One Tambon One Product or OTOP Project is a project that stimulates local businesses based on One Village One

Product (OVOP), the concept of Oita City, Japan. OTOP is a project that helps solve economic problems of people in various communities by using local wisdom and different resources in each community to develop products that can be sold in both domestic and international markets [2].

Each ICT Center in the community is assigned to be center for selling OTOP products online. Currently, there is a website 'www.thaitambon.com' [3] which is an e-commerce web portal that contains many OTOP products. The website provides information of the products that the manufacturers can consign their products for free. According to the preliminary survey, it is found that there were only a few OTOP manufacturers consigning products in the website. Most OTOP manufacturers chose to consign their products through other channels such as the department stores, or other e-commerce websites.

This study aims to use the factors affecting the user's buying decision by developing a prototype website for the community, and to make a comparison between the simulated community website and Thai Tambon website. In the part of the simulated website, it is designed to resemble www.furanotourism.com [4], which is a Japanese Tourist Promotion website. It will be further developed in the field of e-commerce and then considered what users will buy from the website. Moreover, the reasons why the customers buy the products from the website and factors affecting the buying decision will be investigated in order to use these factors to develop the actual community website.

The objectives of this study are to study the nature of website that affects the purchase of community products, and to develop a simulated website for buying and selling

community products. The results of this research are expected to be used in the design and development of e-commerce website for OTOP community enterprise so that the OTOP product community can apply the results to its website to increase revenue. Moreover, the website developers can apply the factors affecting the purchase of community products in the development of the community website.

## II. BACKGROUND & RELATED WORK

### A. Web design for optimizing user experience

Usability is an important feature for software indicating that user interface is easy or difficult to use. It consists of 5 parts: Learnability, Efficiency, Memorability, Errors and Satisfaction [5].

For the website, the ability to use is very important. If the website is difficult to use, such as content on the website is difficult to read, the first page of a website will not attract or confuse the users. As a result, it makes users to have a negative attitude toward the website immediately. The most important thing about e-commerce is that if the users cannot find the product that they want, users will not be able to make purchases.

In the website design, the designer should focus on making everything easy to see. Especially on the large website, the website should not have small icons, and the colors of the menu and links should be different from the website's background color [6]. In addition, if the website has a long article, it should be segmented for easy reading, or highlighted on particular sections to help keep users interested [7].

Moreover, the beauty issue is related to Nelson's problem solving which is 'recognition rather than recall'. This makes the website easier to use and more user-friendly [8]. Specifically, the section of selling a product should have a large product image in order for the user to see all the details without having to rely on reading the details [9].

### B. OTOP Project Concept (One Tambon One Product)

OTOP project was established for each community to use local wisdom of product development. The government helps and supports the community by providing modern knowledge and commercializing the products from the community to domestic and international markets with trading system networks and internet. Moreover, in order to promote and support the local development process, the government helps build a strong and self-reliant community to increase local people's income by using local wisdom resources. Importantly, value-added products and services should be focused in the development process because they are seen as 'the selling points' for both domestic and international markets. This study aims to achieve several objectives including (1) to create jobs in the community, (2) to strengthen the community so that people can develop their community on their own, (3) to promote local wisdom, (4) to promote human resource development, and (5) to promote community initiative in product development in accordance with the way of life and local culture [10].

## III. METHODOLOGY

We firstly conducted a survey to collect user data in Nakhon Ratchasima province by asking for the cooperation from the ICT Learning Center. Participants were asked to complete the survey and the result shows that four factors: (1) Reliability of product, (2) Reliability of website, (3) Perceived ease of use, and (4) Social Influence affected the users' buying intention [11].

We used the result from [11] to develop our prototype website. The website consists of three main sections: the header, body and footer. The header consists of the logo of the website and menu bar. The body is the section which is different in each page, and consists of content such as pictures and articles. The last part is the footer appearing at the end of the page, and consisting of social media bar, contact information, sub menu, etc.

The prototype website was created through Wix which is web design and development tool. Also, Wix is the very famous online website design with the full functionality and ease to organize. To design the prototype web, we searched for OTOP products and then analyzed the existing OTOP websites. After that, we tested with the participants iteratively. The prototype website is shown in Fig. 1.



Figure 1. The homepage of the prototype website

In this study, the prototype website is based on Dantum district community in Nakhon Pathom province which dried tomato is the OTOP product under the brand "Maechuy". The header of the website is the logo of Don Tum district community enterprise and the menu includes shopping cart icon. The body part contains rotating images showing all the products and other information of "Maechuy" products. At the end of the page, it allows users to send the message, and follow in Instagram by providing the social media icon. The Instagram icon is connected to Instagram's account of Maechuy, called maechuytomato. Lastly, the menu is written in Thai.

In addition to the section of the homepage, there is also the section of the product page. This section has the header and footer similar to the homepage and body section.

From the consideration of Maechuy website, we intended to add more attractive images, community articles and social media icons for eye tracking analysis.

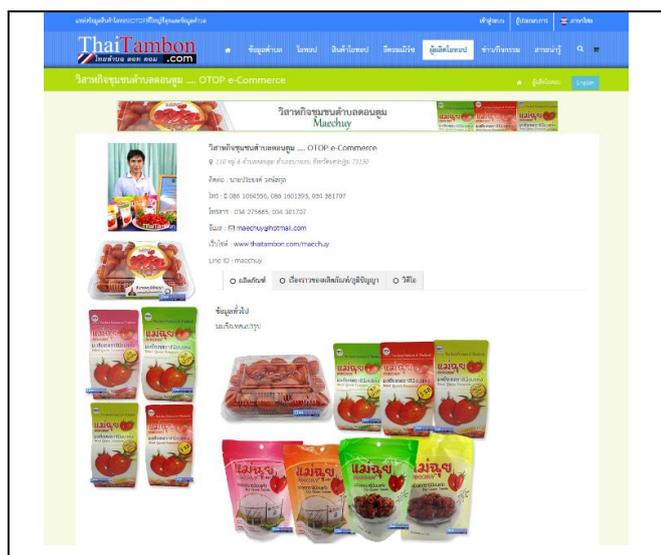


Figure 2. The homepage of Maechuy at ThaiTambon.com

#### IV. EXPERIMENT

##### A. Eye tracking

The eye tracking device, using Tobii Studio program, was used to conduct the experiment in this study. We used the Area of Interest function in Tobii eye tracking to frame the point of interest, and extract the statistical data for comparing the viewing pattern from a video that matches the statistical data. The participants were asked to visit our prototype website which consists of five pages as follows:

- 1) Home: Rotating set of images in the middle of the page
- 2) Product: Product images related to scenario
- 3) About Us: The history of Maechuy tomato
- 4) Manufacturing Process
- 5) Contact Us: Contact information including maps and phone number

##### B. Test procedures

1) Calibration: Calibration is a process of indicating the precision of the eye movement detector testing with the user's eyes. The calibration must be done before every eye-tracking test in order to make sure that the data collected from eye-tracking is more accurate.

2) Experimental preparation: The tool used in this test is called Tobii studio, which is the program using in conjunction with the Tobii eye movement detector.

3) Test method: To study user's behavior, Tobii studio and Tobii eye movement detector were used to store video data.

##### C. Data collection

Data were collected from the participant's gaze plot video according to the following criteria:

- 1) Viewing sequence in the video which is more than 80% of the total looking duration.
- 2) Viewing sequence in the video which must be in the vicinity of the area of interest.

##### D. Data Analysis

The data were analyzed from 3 sources including video sequencing, questionnaire and statistical data.

1) Analysis of video: The data were observed through the viewing sequence of participants in the test. The similarities of the gestures and expressions of each participant were observed and recorded by the researchers. As a result, this helped the researchers analyze the user's behavior and summarize the viewing patterns of the participants easily. Lastly, the results were compared with the statistical data obtained from the Area of Interest in Tobii studio program.

2) Analysis of questionnaire: There are 9 questions in this research. The participants were asked to complete the questionnaire online. The questionnaire is divided into 2 parts as follows:

Part 1: Personal information.

Part 2: Attitude toward the nature of the website, and the reasons for choosing the website service based on the interview.

3) Analysis of eye movement data: The movement of the eyes can be divided into two broad categories which are the heat map and the gaze plot. Both of them are diagrams that help study the movement of the eyes. The heat map determines the density of the viewing at each point.

In this test, the situation was created for the users in order to make every user to have the same purpose of using the website. The situation is as follows:

"In the coming week, you are going back to the old school party and then buy some souvenirs to the teacher. You and your friends decide to buy souvenirs from the OTOP products of Maechuy. Maechuy's products are food including fresh and dried tomatoes with a variety of flavors. It is found that there are two websites selling these products: Thai Tambon website which is OTOP consignment web, and the official website of Maechuy. Therefore, you will need to look at any website in order to buy the products.

There are the products to be ordered as follows:

- 1) 8 cartons of fresh tomatoes
- 2) 4 cartons of dried tomatoes, plum size 160 g.
- 3) 2 bags of original dried tomatoes, size 40 g.

#### V. FINDINGS

##### A. Demographic Data

We conducted the experiment with a questionnaire from 28 participants. Personal information of the participants is shown in Table 1.

TABLE I. THE DEMOGRAPHIC DATA

Characteristic	Frequency	Percent
<b>Gender</b>		
Male	20	71.4
Female	8	28.6
<b>Age</b>		
Under 20	0	0

Characteristic	Frequency	Percent
20 and less than 30	28	100
more than 30	0	0
Income per month		
Under 6,000 BTH	10	35.7
6,000-12,000 BTH	10	35.7
12,000-20,000 BTH	6	21.4
20,000-30,000 BTH	2	7.1
Over 30,000 BTH	0	0
<b>Total</b>	<b>28</b>	<b>100.0</b>

From Table 1, there are 20 male, accounted for 71.4%, and 8 female, accounted for 28.6%, participants. All of participants are between 20-30 years old. Most of the participants have income lower than 6,000 Baht (35.7%) and between 6,000-12,000 Baht (35.7%), followed between 12,000-20,000 Baht (21.4%), and between 20,000-30,000 Baht (7.1%), respectively.

*B. Behavior analysis of heat map*

1) *Homepage*

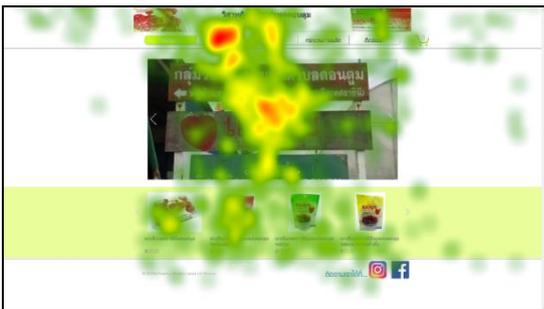


Figure 3. Heat map on the homepage of simulated website

Fig. 3 shows that the highest viewing density is at the product menu next to the homepage menu followed by About Us, and the photo slide in the center of the page. The least viewing density is where the user looked at the slides of the first and second products.

2) *Product page*

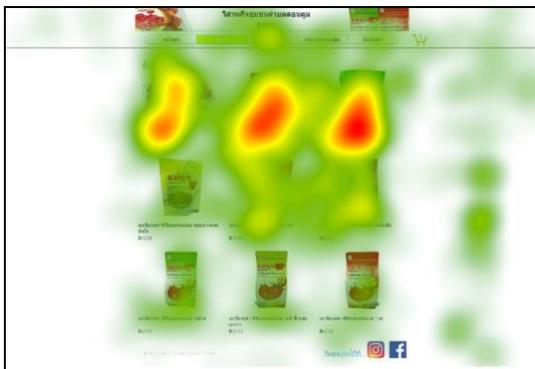


Figure 4. Heat map on the product page of simulated website

Fig. 4 shows the highest viewing density of the page including the top three items. Most users looked at the product on the right side, followed by the product in the middle and the left side, respectively. Moreover, it is noticed that users looked at About Us as well.

3) *About Us page*

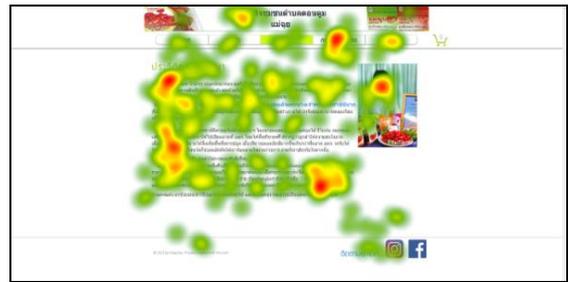


Figure 5. Heat map on About Us page of simulated website

Fig. 5 shows that in About Us page, the viewing patterns are quite fragmented. The highest viewing density on this page is at the beginning of the paragraph of each article, the last paragraph of the article, manufacturer page and manufacturing process menu. The second highest density is Contact Us menu and the badge, Furthermore, it is observed that the users looked at the social media section in this page more than the other pages.

4) *Manufacturing process page*

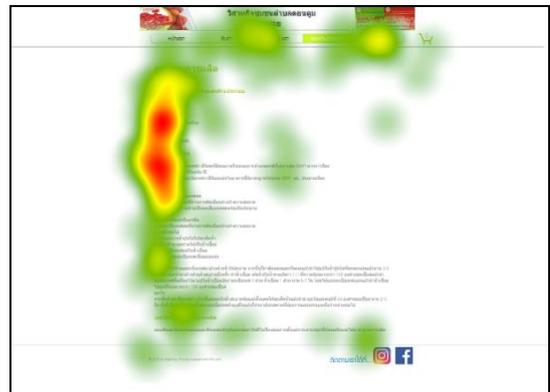


Figure 6. Heat map on manufacturing process page of simulated website

Fig. 6 shows that in the manufacturing process page, there is an article with sub-topics about the ingredients of the tomato. It is found that the users were more interested in the subheadings than the details. However, the usage time on this page is shortest comparing with other pages. Moreover, the users looked at the side menu of the manufacturing process which they looked at Contact Us more than About Us in this page.

5) *Contact Us page*



Figure 7. Heat map on Contact Us page of simulated website

Fig. 7 shows the contact information page. Users looked at the location on the map most followed by the contact details. Moreover, the density of looking at About Us is similar to the density in the manufacturing process page.

From the heat map of the simulated website, it can be concluded that the users looked at the menu bar in sequence, For example, if the user is in the first page, they will look at the menu that leads to the next page which is product page. Moreover, the users were more interested in the images than the articles on the website.

C Additional Statistical Data

a) Time Spending

TABLE II. STATISTICS TABLE OF TIME SPENDING

Location	Visit Duration (sec)	Page	Rank
Center-pic	11.18	Homepage	1
Product-list	5.40	Homepage	2
Fresh-tomato-pic	5.14	Product page	3
Original-pic	4.70	Product page	4
History-content	3.01	About Us page	5
Map-pic	1.68	Contact Us page	6
Buay-pic	1.32	Product page	7
3Header-pic	1.06	Homepage	8
Original-name	0.93	Product page	9
Product-button	0.71	Homepage	10
Owner-pic	0.70	About Us page	11
2.1-Content	0.70	Manufacturing process page	12
Header-pic	0.63	Product page	13
Fresh tomato-name	0.56	Product page	14
Buay-name	0.44	Product page	15
...	...	...	...
Instagram-icon	0.02	About Us page	70
Mnufac-button	0.02	Manufacturing process page	71
Cart-icon	0.00	Homepage	72
Facebook-icon	0.00	Homepage	73
Facebook-icon	0.00	Product page	74
Cart-icon 2	0.00	About Us page	75
Facebook-icon	0.00	About Us page	76
Aboutus-button	0.00	Manufacturing process page	77
Cart-icon	0.00	Manufacturing process page	78
Facebook-icon	0.00	Manufacturing process page	79
Followus-pic	0.00	Manufacturing process page	80
Home-button	0.00	Manufacturing process page	81
Instagram-icon	0.00	Manufacturing process page	82
Cart-icon	0.00	Contact Us page	83
Contact-name	0.00	Contact Us page	84
Contactus-button	0.00	Contact Us page	85
Facebook-icon	0.00	Contact Us page	86

Table II shows what the users spent time looking at. The longest viewing is the images on the center of the homepage with 11.18 seconds per use on average, followed by a product menu bar at the bottom of the center of the homepage with 5.4 seconds per use on average.

The third and the fourth longest duration are the components in the product page. Moreover, it can be seen that top 15 longest viewing are mainly in the homepage and product page besides one component from Contact Us page and another from manufacturing process page.

It is noticeable that the users were interested in the image centered on the page and the slide area of the actual product. The users often clicked the product menu more than other menus.

Users were interested in product images most. Although the topic of the product page was at the third longest viewing duration, the users spent only 0.2 second. Based on statistical data, the users spent time looking at pictures before the product's names and prices. Also, the users spent time looking at the product image at the top longer than the product image below.

The main element of About Us page is an article and a picture of the manufacturer. When observing from the time spent in About Us article, the users spent time reading the article only 3.01 seconds.

At about us page button, it makes sense that the users did not want to read this section of the article because it is not related to their buying decision. Hence, it can be seen that the viewing duration is less than 1 second per time.

For Contact Us page, the longest viewing component is a map image which took 1.68 seconds to view. Also, it has the highest viewing and clicking number.

For the time spending, it can be seen that the participants spent time mostly for viewing the homepage as same as product page (4 of the first 10 most time spending). However, if we consider top 15 of most time spending, the participants spent time mostly on product page (7 of the first 15 most time spending). Therefore, when SMEs need to develop their websites, they should focus more on product page and homepage because their customers spend most of time looking in these sections.

D. Attitude toward Website

TABLE III. RESPONSES TO THE QUESTIONNAIRE IN THE FIELD OF ATTITUDE INFORMATION.

Characteristic	Frequency	Percent
1. Which website between the consignment website and official website is more reliable?		
• Consignment website	4	14.3
• <b>Official website</b>	<b>24</b>	<b>85.7</b>
2. Do you think having social media tab helps you make a purchase?		
• <b>Yes</b>	<b>16</b>	<b>57.1</b>
• No	12	42.9
3. Do you think having the multiple views of product images helps you to make a purchase?		
• <b>Yes</b>	<b>26</b>	<b>92.9</b>
• No	2	7.1
4. Do you think a page with articles such as About Us affects your purchase?		
• <b>Yes</b>	<b>18</b>	<b>64.3</b>
• No	10	35.7
5. Which shopping basket function do you like?		
• Thai Tambon website	4	14.3
• <b>Our prototype website</b>	<b>24</b>	<b>85.7</b>
6. In this test, which web services do you use?		
• Thai Tambon website	10	35.7
• <b>Our prototype website</b>	<b>18</b>	<b>64.3</b>
Total	28	100.0

Table III shows that 24 participants considered that the official website was more reliable (85.7%), while only 4 participants considered that the consignment website was more reliable (14.3%).

16 participants agreed that the social media tab could help make a decision to buy the products (57.1%), while 12 participants disagreed (42.9%).

26 participants agreed that the pictures of multiple products help make a buying decision (92.9%), while 2 participants disagreed (7.1%).

18 participants agreed that the webpage containing the article such as About Us had an effect on the buying decision (64.3%), while 10 participants disagreed (35.7%).

24 participants were satisfied with the features of the shopping basket function in simulated website (85.7%) while 4 participants preferred the features Thai Tambon website (14.3%).

18 participants decided to use the services of simulated website (64.3%), while 10 participants chose Thai Tambon websites (35.7%).

From question 2 in Table II, the participants agreed that having social media tab could help them make a purchasing decision (51.1%). However, from the eye tracking data in Table III, they did not spend much time with social media tab such as Facebook and Instagram.

On the other hand, from question 3 they, the participants agreed that having the multiple views of product images could help them make a buying decision (92.9%) which is inconsistent with the eye tracking data that they spent a lot of time on viewing products. From question 4, they agreed that a page with articles such as About Us page affected their purchase (64.3%) which is in consistent with the data from eye tracking that the participant spent time on history content and the owner's picture in About Us page.

## VI. CONCLUSION

We developed the prototype website from previous research [11] based on the argument that social influence, perceived ease of use, reliability of website, and reliability of product have significantly positive effects on the intention to buy a product on the community website.

We used eye tracking to analyze the participants' purchasing behavior. According to the result from the questionnaire, the participants agreed that having the multiple views of product images could help them make a purchase, accounted for 92.9%. This finding is also inconsistent with the eye tracking data which they spent a lot of time viewing product. Also, they agreed that a page with articles such as About Us affected their purchasing decision, accounted for (64.3%) which is in consistent with the data from eye tracking that the participant spent time on history content and the owner's picture in About Us page.

However, when using questionnaires, participants agreed that having social media tab could help them make a purchase (57.1%), but from eye tracking data, they did not spend time looking on social media tab such as Facebook and Instagram.

In summary, it is found that some features of a website affect the users' decision to buy community food and beverages such as attractive images, community articles, articles that are related to security policy, payment method and ordering methods.

Moreover, if the website section is reliable and updated regularly or contains the instruction for use, the social media bar and the multiple images of the products will play less important role in influencing the user's buying decision for this type of product. Furthermore, this study demonstrates that website layout and ease of use are the most important features for e-commerce website.

## ACKNOWLEDGMENT

This research was supported by the "Chulalongkorn Academic Advancement into Its 2nd Century Project (CUAASC)."

## REFERENCES

- [1] Ministry of Digital Economy and Society, September 2016. "Thai Community ICT Learning Center Project", <http://www.thaitelecentre.org/> [Accessed: 1-September-2017].
- [2] Community Development Department of Thailand, June 2016. "OTOP", <http://www.cdd.go.th/content/download/documents> [Accessed: 1-September-2017].
- [3] Thailand One Tambon One Product (OTOP) Website, August 2016. "Thaitambon Website", <http://www.thaitambon.com/otop> [Accessed: 4-September-2017].
- [4] Japanese Tourist Promotion website, August 2016. "Japanese Tourist Promotion website", <http://www.furanotourism.com/en/> [Accessed: 4-September-2017].
- [5] Nielsen, J., June 2012. "Usability 101: Introduction to Usability", <https://www.nngroup.com/articles/usability-101-introduction-to-usability/> [Accessed: 6-September-2017].
- [6] Whitenon, K., June 2015. "Menu Design: Checklist of 15 UX Guidelines to Help Users", <https://www.nngroup.com/articles/menu-design/> [Accessed: 14-September-2017].
- [7] Loranger, H, August 2017. "7 Tips for Presenting Bulleted Lists in Digital Content", <https://www.nngroup.com/articles/presenting-bulleted-lists/> [Accessed: 4-May-2017].
- [8] Gumussoy, C.A., "Usability guideline for banking software design". *Computers in Human Behavior*, Vol 62, p: 277-285.
- [9] Schade, A., 2014. "Ecommerce UX: 3 Design Trends to Follow and 3 to Avoid", <https://www.nngroup.com/articles/e-commerce-usability/> [Accessed: 8-September-2017].
- [10] Community Development Department of Thailand, June 2016. "OTOP", <http://www.cdd.go.th/content/download/documents> [Accessed: 4-September-2017].
- [11] Akasarakul K., Cooharajanane N., Lipikorn R., "A Study of Factors Influencing Intention to Purchase". 2017 18th IEEE/ACIS International Conference on June 26-28, 2017. *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, Kanazawa, Japan; pp.213-218.

# Data Visibility and Trust Enhancement of Enterprise Customers in Cloud Computing Services

W.P. Yuen, K.B. Chuah

Department of Systems Engineering and Engineering Management  
City University of Hong Kong  
Hong Kong  
wpuyen2-c@my.cityu.edu.hk / mebchuah@city.edu.hk

*Abstract*—This Cloud computing is a pervasive technology and has been a platform in IT for several years. Cloud service providers (CSP) have developed and offered different service platforms to accommodate different needs of enterprise subscribers. However, there still exists the situation of enterprise customers' hesitation and reluctance to deploy their core applications using cloud service platforms. Our survey results show that security is the perceived major concern of existing and prospective enterprise customers of cloud services. The research investigates the trust between enterprise customers and cloud service provider with regards to the perceived security of cloud services. Enterprise customers expected to be reassured of cloud service security in a more visible way from cloud service provider. The perception of data security in cloud computing platform can be enhanced by data visibility. The term data visibility has been widely used in the IT industry especially from ICT product and solution vendors. However, there is no common practice guideline or standard in industry. This paper defines the characteristic and elements of data visibility and proposes a model and architecture as best practice reference for the industry. We expect that the enhancement of data visibility can earn the trust of enterprise customers in adopting public cloud services.

*Keywords*- data visibility; perception of cloud security; cloud trust; data visibility elements

## I. INTRODUCTION

The pay-per-use business model of cloud computing allows their tenants to adjust their investment in IT resources dynamically, enterprise can reduce the CAPEX (capital expenditure) and OPEX (operating expenditure) on both infrastructure such as servers, network, security, storage, application, and IT operation, administration/management.

It is often attractive to enterprises especially small and medium enterprise (SME). When enterprise adopted the public cloud business service, they might lose the control over their owned data. Different concerns from customer's perspective have appeared such as security, data integrity, performance, compliance etc. Among all these concerns, security is perceived as the major concern by both existing and prospective enterprise customers of cloud services.

Data security and data privacy issues in different service models are closely related to the visibility of cloud. They made

enterprises not comfortable with cloud platform (especially SaaS model) due to lack of traffic visibility in their data being stored and secured.

Today, subscribers can only trust the audit report provided by CSP with standard code of practice and certification such as ISO27001/27002/27017/27018. However, those standards only require documenting all the steps for manual procedures that touch production environment. SLA (service level agreement) is supposed to be another source for customer to trust on the service from cloud provider, but most of the SLA covers service availability (e.g. 99.999% per year) more than security, confidential and integrity aspects [1].

According to passed literature studies, perception of data security in cloud computing platform can be enhanced by data visibility.

The term data visibility has been widely used in the ICT industry especially data security related product and solution vendors. There is no standard definition or guideline to explain this term in ICT industry. This paper defines the characteristics and the dimensions of data visibility, proposes the conceptual model, framework and architecture on cloud platform for enterprise customers as best practice reference in adopting public cloud services. We aim at investigating the data visibility for the consumer to enhance trustworthy cloud computing. The related work on this paper is to propose a solution to address the above problem with the following contribution:

- Define the core elements and characteristics of data visibility in cloud computing service.
- Propose the conceptual model, framework and reference architecture of data visibility model as best practice framework.

## II. LITERATURE REVIEW

### A. Security Issues

There are different research and studies related to security, threat and vulnerabilities of cloud computing services. NIST (National Institute of Standards and Technology) proposed 9 key security issues in 2011 [2], CSA (Cloud Security Alliances) proposed 12 top threats in cloud [3]. European

Union Agency for Network and Information Security (ENISA) in 2015 raises 9 risks and 12 security questions [4].

TABLE I. TABLE OF MAJOR SECURITY ISSUES

Cloud Concerned item	CSA	NIST	ENISA
Data breaches	CSA 1	NIST 4.9	
Data loss	CSA 8	NIST 4.8	
Data protection	CSA 2	NIST 4.7	ENISA Q8
Availability		NIST 4.8	ENISA Q3
Security - APT/DDoS attack	CSA 7, 10, 11		
Software isolation	CSA 12	NIST 4.6	ENISA Q7
Shared technologies vulnerabilities	CSA 12	NIST 4.4	
Account hijack	CSA 5		
Data location	CSA 9	NIST 4.2	
Unauthorized access	CSA 2	NIST 4.5	ENISA Q6
Control		NIST 4.1, 4.2, 4.5	
Malicious insider	CSA 6	NIST 4.3	
Insufficient due diligence	CSA 9		ENISA Q5
SLA		NIST 4.8, 3.3	ENISA Q4
Privacy		NIST 4.2, 4.7, 4.9	
Compliance	CSA 8, 9	NIST 4.2	ENISA Q12
Monitoring	CSA 3, 6, 12	NIST 4.1, 4.4-4.6, 4.9	ENISA Q9
Forensic		NIST 4.9	
Instance response	CSA 5	NIST 4.9	

### B. Security and Visibility

Security and data privacy issues in different service models are closely related to the visibility of cloud platform from subscriber’s point of view. It makes most enterprise are still not comfortable cloud platform (especially SaaS model) due to lack of traffic visibility about the way their data is stored and secured [5].

The Guideline on Security and Privacy in Public Cloud Computing from NIST [2] indicate that visibility is a continuous monitoring of information security requires maintaining ongoing awareness of security controls, vulnerabilities, and threats to support risk management decisions. Service arrangements should include some means for the organization to gain visibility into the security controls and processes employed by the cloud provider and their performance over time. Ideally, the consumer would have control over aspects of the means of visibility to accommodate its needs, such as the threshold for alerts and notifications, and the level of detail and schedule of reports.

There are different research papers and articles about the relationship between data traffic visibility and security in cloud computing environment in recent years.

- Reference [6] highlight the requirement of visibility that it is necessary to review and access log and audit trail than only authorized user are allowed to access the data. Malicious insider is one of the threats in cloud, and customers require enough visibility for monitoring.
- Reference [7] mentioned that acquisition of evidence should be the role of provider. However, they propose a data owner (user) approach to collect evidence for forensic by applying agency-based software to collect all activities in cloud. CSP should enable the visibility to cloud customer by provide an audit report with the

data forensic acquisition and analysis to the data owner.

- Reference [8] raise the question of whenever SaaS environment is used, the service provide will be responsible person for storage of data, in a way in which visibility and control is limited. So how can a customer retain control over their data when it is stored and processed in the cloud?

### C. Visibility and Trust

The major security challenge of customers to adopt public cloud computing service is the third-party control issue. The data owners do not have their control in data processing. CSP should make the management and operation more transparent and auditable by customer. Transparency of what security is enforced, existed risk, and possible breaches on cloud platform. This is “trust-but-verify” where cloud customer should trust in their CSP and CSP should deliver tools to help customer to verify and monitor security enforcements [9].

Lack of Transparency from is one of the challenges of trust to cloud computing from customer’s perspective. There are two issues involved in transparency, one is the physical location of the storage and processing site, the other is the security profiles of those sites [10].

There are many research papers and articles about the relationship between data traffic visibility, security and trust in cloud computing.

- Reference [11] believed that the issue of low trust on cloud computing is an obstacle. They presented a model for trust in cloud computing, accounting for important elements which shape the users trust and a way of evaluating each element’s importance. In order to achieve them, it is definitely need to a mechanism for user to visualize those elements.
- Reference [12] suggested a model that users form perceptions of security control that strongly determine how much trust they put in online services. This model can also be applied on enterprise users have the perception of security by data visibility in trust to use cloud service.
- Reference [13] mentioned that most enterprise still not comfortable with SaaS due to lack of control, lack of trust, and lack of visibility with might cause different security threats
- 50% of cloud users believe identity and access monitoring/management is the cloud service provider’s responsibility. If CSP does not adhere to these security audits, then it leads to an obvious decrease in customer trust [14].

### III. REQUIREMENT OF DATA VISIBILITY IN CLOUD

Customers are ultimately responsible for the security and integrity of their own data in the cloud. This shared-responsibilities idea was proposed by major incumbent CSP, and is mentioned in different standard cloud institutes.

Customers may not feel comfortable to adopt cloud computing service for their key applications and data in lacking transparency and visibility. On the other hand, they need to part of the responsibilities in data security. Some CSP provide optional service by offering activities log or an online interface for customer to access their activities. In fact, the trusted data visibility features should comprehensive enough to cover all the activities on user's data in the entire cloud platform including the administration or operation staffs in physical and virtual environment.

Besides, the data visibility information should be support either in pushing mode to notify data owner for any events, or in pulling mode by user requested, or both [15].

NIST [1] also indicate that the organization (cloud enterprise customers) need to manage security and privacy risks, as appropriate for each level of the organization involved in decision making. Tenants or data owners should have control over aspects of visibility to accommodate their needs, such as the threshold for alerts, and user-defined monitoring policies with different level of reports.

However, the term data visibility is widely used in ICT nowadays. Products like NPB (network packet broker), application performance monitoring products, network traffic analytic products, data traffic analytic products, and system performance monitoring products, use this term as their product features in marketing. So, there is no official or standard definition or practice guideline on "data visibility" in ICT industry.

According to the result of an industry survey in Melbourne (see Fig. 1), 77% of surveyees agreed that monitoring/visibility as part of cloud service is important to enterprise users in APAC.

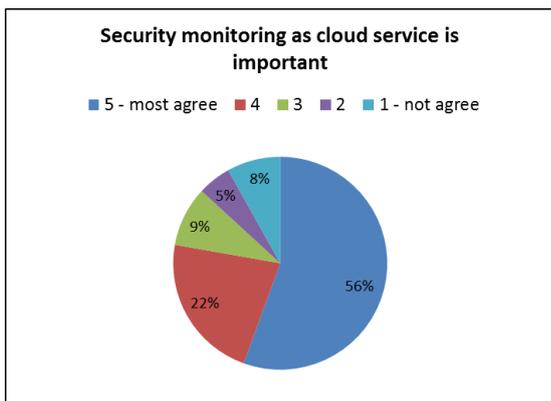


Figure 1: Industry survey result in Melbourne Jul.2013 (sample size 74)

The result of another industry survey in Singapore (see Fig. 2) shown that 75% of surveyees from APAC agreed that more visibility in network can enhance security.

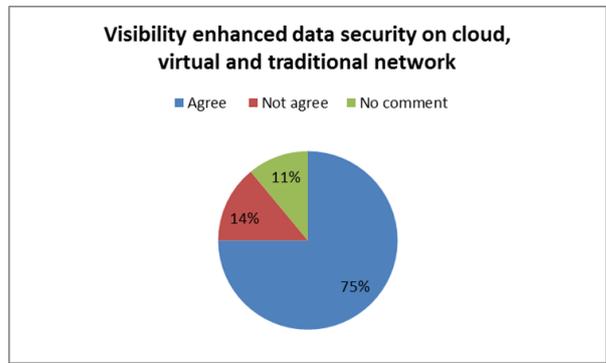


Figure 2: Industry survey result in Singapore 2013 (sample size 264)

#### IV. PROPOSED DATA VISIBILITY MODEL

##### A. Properties of data visibility model

The proposed data visibility model on cloud computing service will have the following properties:

- 1) *Customer centric* – Data should belong to data owners, its policies, magnitude and dimension (both depth and width) should be defined and driven by data owners or customers. Data owners need to establish a mechanism to mandate the enforcement of their own policies to ensure data confidentiality and integrity.
- 2) *End to end* – Required data status should be reflected across different domains from network, server, virtual/container, application to storage [14].
- 3) *Bi-directional/Interactive* – Visibility alert in the form of pop up log message is not enough, it is desirable if it can be driven by user to acquire enough data status at any time. Besides, any data status triggered within the condition defined by user should be reflected.
- 4) *Real-time/Low latency* – The industry does not require monitoring equipment to report any events within a specified period of time. CSCC (Cloud Standard Customer Council) [16] suggests that the elapsed time from when a service is invoked to when it is completed (typically measured in milliseconds). The frequency of collection should be by every minute.

##### B. Core elements of data visibility model

The 5 core elements of the proposed data visibility model are adopted from the well-known "5W1H" model widely used in problem management analysis. They are as follows:

- 1) *Timestamp* – When the activities happened? An accurate time should be provided on the data in different stages. The timestamp in log of each event should be accurate enough for business application use. Eg. PTP/NTP timestamp
- 2) *Data assets* – What kind of data has been involved? Should be detailed enough to identify what data type and data source involved in event or state changes. Eg. Database, table, object, filename.

3) *Data location* – The actual location of physical storage or any virtual machine when data is at rest. Eg, IP address, domain name, mac address and geolocation.

4) *Actor* – Who is the person access that specific data? Data can be access or manipulated in different domains by different people in application or system level. Those information can be acquired by aggregating from different IAM systems in cloud.

5) *Action* – How do the system or application be processed/accessed? Is the data being inquired by normal user? Or modified by unauthorized people? Or dropped during transmission? Those information could be acquired from log/SIEM of different domains and application to identify the data status is changed in each event.

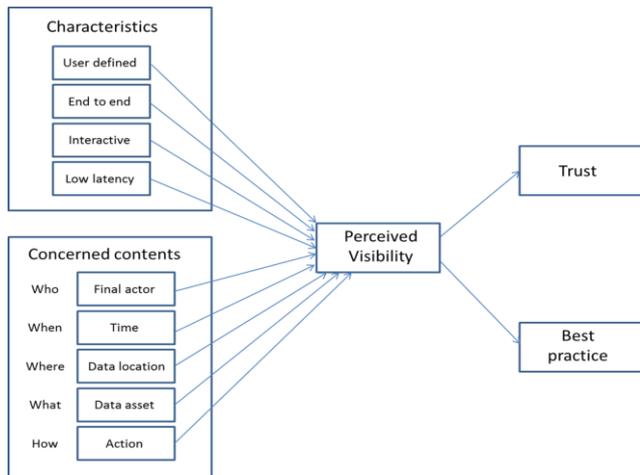


Figure 3. Research framework of data visibility in cloud

C. Conceptual model

The proposed data visibility mechanism is to enable cloud customers to construct their required visibility condition in cloud. User visibility policies are defined according to their own requirements, and those policies will forward to SaaS as selection conditions on relevant activities information (see Fig. 4).

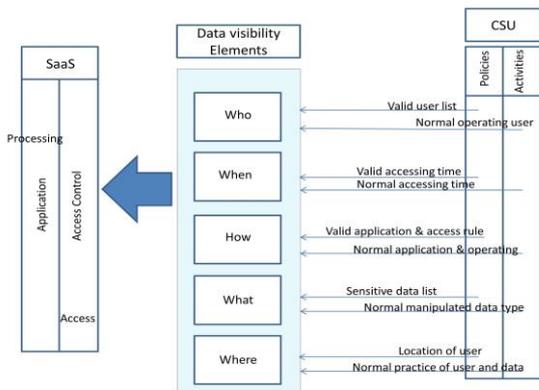


Figure 4. Data visibility requested from tenant to CSP

SaaS as application service provider, in most cases, are the single point of contact to customer on behalf of other cloud service provider (PaaS/IaaS). So, besides the normal application data communication, SaaS can also take the role of visibility policies interface to customers. Visibility features can be interactive between customer and provider, or it can also be alert based unicast to specific customer.

For some uncertain activities which happened in data center, SaaS need to get those information from PaaS/IaaS. So, there is an information flow between them (see Fig. 5). The major characteristics of this prototype are customer centric and end-to-end visibility.

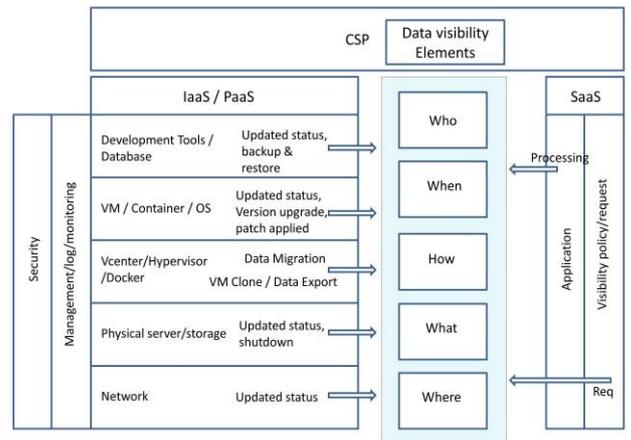


Figure 5. Data visibility conceptual model

V. DESIGNED ARCHITECTURE

A. Components

There are three major components in the data visibility architecture. User requested Interface, Policy mapping module and Collecting & filtering module.

1) *User Requested interface (URI)* – Users input their required visibility on specific data assets, such as user, timing, access method and location. Users can make ad hoc inquiry, and define data visibility policies.

2) *Policy mapping module (PMM)* – The inquiry received from URI will be converted into instruction rule resident into memory buffer (or TCAM). These instructions will be forwarded to Collecting & filtering module (CFM) for visibility data collection. Then PMM will map the specific data with source of inquiry and egress to users

3) *Collecting & filtering module (CFM)* – It is a collector of all different logs and events from IaaS and SaaS. Based on the instruction from PMM, the collected data will be filtered before sending back to PMM. This can make sure the visibility data meet the requirement and also discard any unrelated traffic so as to reduce the network loading

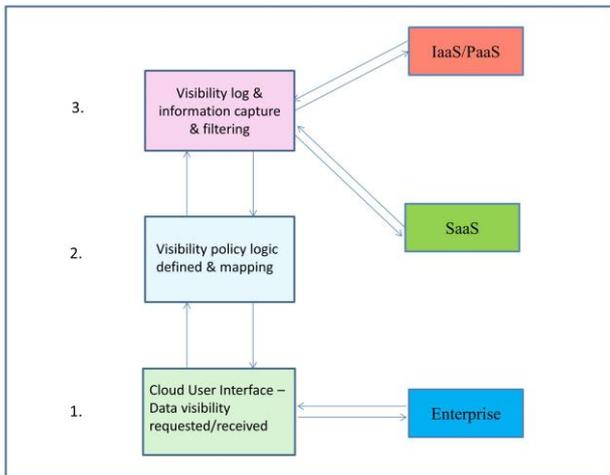


Figure 6. Architecture component

VI. COMPONENTS OF DATA VISIBILITY ARCHITECTURE

A. Architecture

Visibility requested by user through URI will be converted into policies from PMM, the instruction will then be sent to CFM to collect suitable visibility data from different CSP. There, all collected visibility data will be filtered and sent back to PMM for mapping into specific policies which belong to data owners. Finally, the visibility information will be forwarded to customer ends (see Fig. 7).

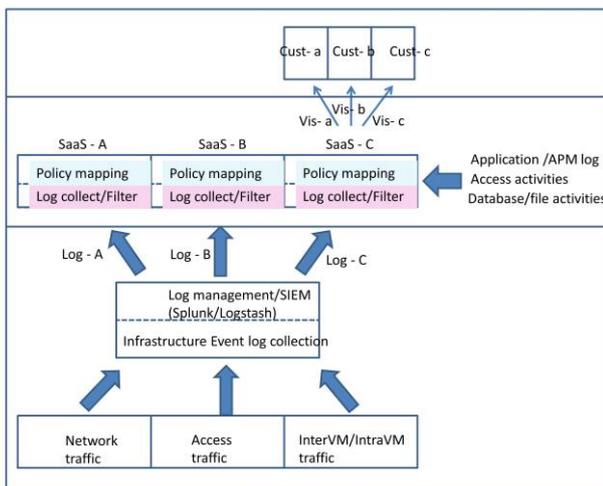


Figure 7. Architecture of data visibility in cloud

Log and event traffic from infrastructure activities will be collected to log management or SIEM such as Logstash or Splunk. Related traffic will be distributed into different SaaS. In SaaS, application activities events, access information, database management events will be collected, and aggregated with the infrastructure traffic into Collecting and Filtering

Module (CFM) for processing. Relevant data visibility information will be filtered, mapped and egress to relevant data owner/customer. The entire visibility traffic across the cloud is centralized and managed for data owner (see Fig. 8).

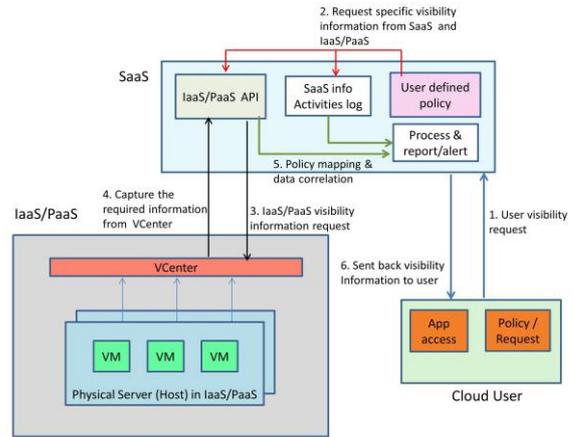


Figure 8. Visibility information walkthrough

VII. DATA VISIBILITY INFORMATION WALKTHROUGH

A. Visibility data taxonomy

1) Infrastructure traffic

- Event from networking and security device including NMS, IDS (host/hypervisor/network based), IPS, Firewall, DPI, load balancer, NAS and NPB
- Event and log from access device - Identity and Access Management (IAM)
- Event and log from VM/Hypervisor and host - VMI, /var/log/auth.log, /var/logs/syslog, Hyper-V log, VMM, external monitoring tools such as Solarwinds

2) Application traffic, log and event

- Event from APM, Lifecycle Management, Configuration Management, Security Management -Security Services, DB Policy Management, HTTP Server, Database Identity Management, Database Access Management, and Directory Services Management, and application log from any external devices.

VIII. CONCLUSION

This model is expected to enable the visibility information flow through different domains across cloud to end users. When designing the conceptual mode, operation and practice of cloud providers also need to be considered, so as to ensure that the model is practical.

This conceptual model has been introduced to 20 different enterprises, system integrators and CSPs in Asia for their

comment. The feedbacks are positive, and worthy for exploration and development.

The next stage of this research is to develop the conceptual model to a prototype. Then, put to industry stakeholders including different cloud providers, cloudbrokers (solution integrator in APAC market), and also enterprise customers, especially in SME for valuation.

It is expected to be the practice reference model for cloud provider of SI when they deliver their cloud-based solutions to enterprises.

#### REFERENCES

- [1] C. Rong, S. T. Nguyen, and M. G. Jaatun, 2013 Computers and Electrical Engineering 39 (2013), *SciVerse ScienceDirect*, pp. 47-54
- [2] NIST Cloud Computing Security Reference Architecture, May 2013 NIST Special Publication 500-299, *NIST Cloud Computing Security Working Group, NIST Cloud Computing Program*.
- [3] The Treacherous 12 - Cloud Computing Top Threats in 2016, Feb 2016 *CLOUD SECURITY ALLIANCE*.
- [4] M.A.C. Dekker, D. Liveri, Cloud Security Guide for SMEs *Cloud computing security risks and opportunities for SMEs* April 2015, European Union Agency for Network and Information Security.
- [5] Subashini & Kavitha, Journal of Network and Computer Applications – A survey on security issues in service delivery models of cloud computing, July 2010.
- [6] P. Kuppuswamy, Saeed QY Al-Khalidi, Analysis of security threats and prevention in cloud storage: Review Report, International Journal of Advanced Research in Engineering and Applied Sciences, Vol 3, No.1 Jan 2014 ISSN:2278-6252.
- [7] Alqahtany, Clarke, Furnell & Reich, A forensically-enabled IAAS cloud computing architecture, Proceedings of the 12th Australian Digital Forensics Conference. Held on the 1-3 December, 2014.
- [8] M. Mohaned Sirejudeen, K. Subramanian, Security issues on Data Transfer Under Clouds- An Overview, International Journal of Information Technology Infrastructure, Vol 3, No.5 Sep-Oct 2014.
- [9] H. Tianfield, Security Issues In Cloud Computing, 2012 IEEE International Conference on Systems, Man, and Cybernetics October 14-17, 2012, COEX, Seoul, Korea.
- [10] K. Khan and Q. Malluhi, Establishing Trust in Cloud Computing, Sep/Oct 2010 IT Pro, IEEE Computer Society, 1520-9202/10.
- [11] A. Rashidi and N. Movahhedinia, A Model for User Trust in Cloud Computing, International Journal on Cloud Computing: Services and Architecture(IJCCSA),Vol.2, No.2, April 2012, DOI: 10.5121/ijccsa.2012.
- [12] Roy, Ow & Kim, Security Assurance: How Online Service Providers Can Influence Security Control Perceptions and Gain Trust, Decision Sciences, Volume 42 Number 2, May 2011, Decision Sciences Journal 2011, Decision Sciences Institute.
- [13] A. Amin Soofi, M. Irfan Khan, Ramzan Talib, Umer Sarwar, March 2014 Security issues in SaaS Delivery model of Cloud Computing, *International Journal of Computer Science and Mobile Computing*, Vol 3 Issue 3, page 15-21.
- [14] V. Om. Gupta, Y. Rai, March 2015 Threats and Vulnerability in Cloud Computing, *International Journal of Advent Research in Computer and Electronic (E-ISSN:2348-5523) special issue*. National Conference "Convergence 2015".  
G. Katsaros, G. Gallizo, R. Kubert, T. Wang, A Multi-level Architecture for Collecting and Managing Monitoring Information in Cloud Environments <https://pdfs.semanticscholar.org/a92b/ca07f0b7ec4ab-22f1cc2ef8adba2f76979a4.pdf>
- [15] Cloud Standards Customer Council, April 2015 Practical Guide to Cloud Service Agreements version 2.

## **Session 18: Infonomics and e-Technology**

Title: Automation of Cyber-Reconnaissance: A Java-based Open Source Tool for Information Gathering

(Authors: Ahana Roy, Louis Mejia, Paul Helling, Aspen Olmsted)

Title: Towards a Security Baseline for IaaS-Cloud Back-Ends in Industry 4.0

(Authors: Elisabeth Bauer, Oliver Schluga, Silia Maksuti, Ani Bicaku, David Hofbauer, Igor Ivkic, Markus G. Tauber, Alexander Wöhrer)

Title: Secure E-Mail Communication – Comparison and Selection of Encryption Solutions Using an Utility Value Analysis Approach

(Authors: D. Fischer, B. Markscheffel, K. Scherr)

Title: Towards Comparing Programming Paradigms

(Authors: Igor Ivkic, Markus G. Tauber, Alexander Wöhrer)

# Automation of Cyber-Reconnaissance

## A Java-based Open Source Tool for Information Gathering

Ahana Roy, Louis Mejia, Paul Helling, Aspen Olmsted

Department of Computer Science  
College of Charleston, Charleston, SC 29401

**Abstract**— In this paper, we propose a tool which captures a footprint of an organization, useful for information gathering phase during penetration testing. It has been found that there is a dearth of an easy-to-access tool which would help in the first phase of such tests; Reconnaissance. This Java-based tool helps in locating and saving organization specific data. Such data repositories will help in vulnerability assessment and profiling of an organization.

**Keywords:** *Cyber Attack, Information Assurance, Information Security, Cyber Reconnaissance*

### I. INTRODUCTION

Reconnaissance refers to the preparatory phase where a penetration tester seeks to gather as much information as possible about a target of evaluation prior to launching a penetration test. It involves three phases: foot printing, scanning and enumeration of the network. In this research, we will be dealing with automating foot printing of an organization. Foot printing is the blueprint of the security profile of an organization, undertaken in a methodological manner. It discovers all information available about the target that is available through public domain sources. It is a time-consuming process to browse through web pages and collect information; hence in this paper, we investigate the problem of tedious web search and propose an efficient way to extract, organize and store data from search engines using a new command line tool, SearchSimplified. Information gathering techniques can be roughly classified into the following:

- **Active:** This includes intrusive reconnaissance that sends (specially crafted) packets to the targeted system, for example, port-scanning. Advanced network enumeration techniques avoid direct communication with the targeted host [1].
- **Passive:** This includes reconnaissance that either does not communicate directly to the targeted system or that uses commonly available public information, not normally identifiable from standard log analysis [2]. This paper focusses on this category.

The organization of the paper is as follows. Section II describes some open source tools and related work. Section III provides information on how a search engine like Google could be exploited for organization profiling. Section IV we provide the structure of the code and libraries in use. In Section V, we show the program out and disseminate the results. We conclude and discuss future work in Section VI.

### II. RELATED WORK

Open-source intelligence (OSINT) is intelligence collected from publicly available sources. There are lots of tools available over the internet [3] which includes all publicly accessible sources of information such as social networking sites, video sharing sites for example.

Michael Henriksen created a command line tool Gitrob [4], which looks for sensitive information in GitHub repositories. It is written in Ruby. Firstly, it collects all the public repositories of a specific organization and then makes a compiled list out of it. Then, it gathers all filenames listed in each repository and compares it with sensitive patterns of data or signatures. All the members, repositories, and files will be saved to a PostgreSQL database. Then, it will start a Sinatra web server locally on the machine, which will serve a simple web application to present the collected data for analysis. A new version [5] was released recently which serves as an improvement to the former version. Susanne Young discusses in [6] how third-party data breaches cost significantly more per lost record than local data breaches. She emphasizes how open source tools such as Recon-ng, search engines, Shodan, Search Diggity help provide businesses a valuable profile of a company's strengths and weaknesses. The paper focuses on how the ubiquity of internet-connected systems can be exploited by criminals and vandals to find (and compromise) victims. Also, the potential dangers of not keeping external systems of an enterprise (email server, DNS server) up to date are cited here. The paper gives some examples of damages incurred to companies due to password leakage through banners.

Recon-ng is a full-featured web reconnaissance framework written in Python. It provides a powerful environment in which open source web-based reconnaissance can be conducted quickly and thoroughly [7].

### III. USING GOOGLE FOR PROFILING

Google's cache system, advanced query operators, such as site: filetype: intitle: or even translation services, makes it a major tool in the passive information gathering arsenal. A Google dork query, sometimes just referred to as a dork, is a search string that uses advanced search operators to find information that is not readily available on a website. Google dorking, also known as Google hacking, can return information that is difficult to locate through simple search queries.

- **Using the cache system:** Google keeps snapshots of crawled pages in its own repository. You may have

experienced it using the “cached” link appearing on search results pages. The advanced operator cache: is used to jump directly to the cached snapshot of a Web site without performing a query. This is a quite simple and effective way to browse Web pages without any direct connection to the target.

- Discovering network resources: Google can help with the network discovery phase. Google search is an alternative to DNS queries, by combining the site: operator and logical NOT, a hacker can obtain a list of public servers. For example, “site:microsoft.com -www.microsoft.com” will reveal msdn.microsoft.com, directory.microsoft.com, partnerconnect.microsoft.com, officelive.microsoft.com, and so forth. Moreover, the link: operator finds pages that link to the queried URL; it can be used to provide important clues about relationships between domains and organizations. The allintitle: intitle: and inurl: operators can be used to detect the presence of Web-enabled network devices, such as routers. For example, inurl:tech-support, inurl:show Cisco OR intitle:“switch home page” site:example.com searches Cisco’s Web-enabled devices on the domain example.com.
- Harvesting system files, configuration files, and interesting data using Google directives: Hundreds of Google searches can be found in [8]. This book describes in depth advanced operators and how to use them to find passwords (clear or hashed), user names, Web-enabled devices, and so on. The main idea is to combine operators, such as intitle: inurl: and site: with specific sentences.
- The Google Hacking Database (GHDB) is a database of queries that identify sensitive data. It is an authoritative source for querying the ever-widening reach of the Google search engine. In the GHDB, you will find search terms for files containing usernames, vulnerable servers, and even files containing passwords [9].

#### IV. SOFTWARE IMPLEMENTATION

Jsoup [10] is a Java library that provides convenient API to extract and manipulate data using the best of DOM (Document Object Model), CSS (Cascade Style Sheets) and jquery-like methods. It scrapes and parses HTML from the company (as provided by the user) URL. The method *CheckIfCompanyExists* uses regex and jsoup to scrape the search results and extract the domain name. The methods *GetHostNames* and *GetIPAddress* use *java.net.InetAddress* class to provide host names and IP address of the company in search. We are also using dnsjava [11] for DNS lookups. Mail Servers with preferences can be extracted along with DNS servers of the company in search. The method *SearchForFiles* searches for company files according to the format chosen by the user from command line (pptx, xls x, docx, txt and pdf). Like filetype directive, many other directives could be also used (for example, inurl, intitle, intext, allinurl,

allintitle, related) to dig specific information related to an organization. Fig. 1 shows the class diagram.

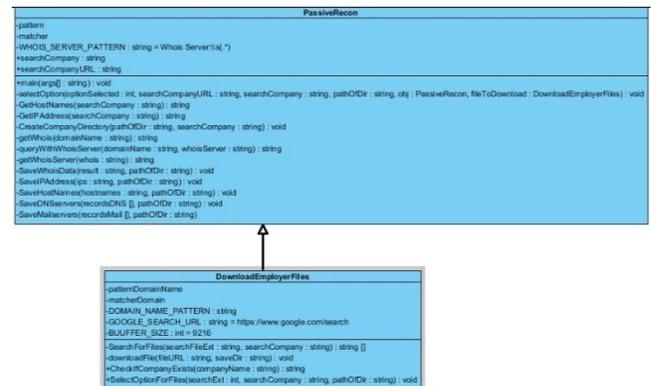


Figure 1. Class diagram of SearchSimplified tool

We have also implemented the functionality of extracting results from WHOIS search page. Using regular expression matching, the tool parses the server name, once a match occurs, the target in search is queried with the server and results are saved in the local machine. Whois is effective in providing useful information of the organization in search in a consolidated manner.

#### V. RESULTS

Fig. 2 displays a sample program output after it’s executed. The user enters the name of the organization and proceeds to query mail servers of the organization.

When the user selects “Company files”, he is asked to choose from a list of file types related to the organization in search. The tool searches for available files and if a match is found, these are downloaded and saved with timestamps in the users’ local machine. This is demonstrated in Fig. 3. Currently, the tool is iterating over google search results of the company name as entered by the user and providing network information as well as certain organization specific files. The tool also records and saves whois lookup data and all other relevant information in the user’s local machine. The types of information gathered by the tool is depicted in Table 1.

```

C:\test>java -jar PassiveRecon.jar
Enter the company name on which you want to gather information.
CofC
What information you want to gather?
1. IPs
2. Hostnames
3. DNS servers
4. Mail servers
5. WHOIS data
6. Company files
7. E-mail addresses
Please enter 1, 2, 3, 4, 5, 6 or 7: 4
Host mxh-00177701.gs.lb.pphosted.com. has preference 20
Host mxa-00177701.gs.lb.pphosted.com. has preference 10
Mail server list saved in C:\PassiveRecon\temp\CofC\Mail_servers_201704251118.txt
Information gathering and saving of data complete
    
```

Figure 1. Mail servers displayed of the company in search

```

Enter the company name on which you want to gather information.
CofC
What information you want to gather?
1. IPs
2. Hostnames
3. DNS servers
4. Mail servers
5. WHOIS data
6. Company files
7. E-mail addresses
Please enter 1, 2, 3, 4, 5, 6 or 7: 6
Enter the file format you want to download
1. PowerPoint or pptx
2. Document or docx
3. Excel Sheet or xlsx
4. Text or txt
5. PDF
Please enter 1, 2, 3, 4 or 5: 5
Content-Type = application/pdf
Content-Disposition = null
Content-Length = 685044
FileName = privacypolicy080409.pdf
File was downloaded in C:\PassiveReconTemp\CofC
Content-Type = application/pdf
Content-Disposition = null
Content-Length = 178860
FileName = CIA95.pdf
File was downloaded in C:\PassiveReconTemp\CofC
    
```

Figure 3. Company file formats

TABLE 1. Types of information gathered by the tool

<b>Network Information</b>	IP addresses
	Host names
	DNS Servers
	E-mail Servers
<b>Organization Profile</b>	Whois data lookup
<b>Organization Files</b>	Powerpoint Presentations (pptx)
	Excel sheets (xlsx)
	Word Document (docx)
	Text (txt)
	Portable Document Format (pdf)

## VI. CONCLUSION AND FUTURE WORK

Most organizations and system administrators are familiar with penetration-testing and intrusion-detection techniques. These techniques are cornerstones in security evaluation and focus mainly on the exploitation of vulnerabilities and suspicious/malicious behavior (e.g., log analysis). However, an organization relying mainly on these techniques may underestimate the huge amount of information that can be anonymously obtained from publicly available content over the Internet. This paper provides the possibility of implementing network-based passive information gathering using Java. Passive techniques are also very useful from an internal perspective; it reduces traffic within the internal network (e.g., passive OS fingerprinting to enumerate OSs in use). The tool we present here is useful for penetration testers to look for organization vulnerabilities. We can then take countermeasures to protect it from potential threats. Code reviews should be conducted, and web pages should be cleaned to avoid comments, prolix banners, version numbers, and so forth. A lot of information can be gathered from error pages, banners, and seemingly innocuous information. Comments can be incredibly information leaking, too.

To sum up, information gathering is the very first step of an attack and probably the most crucial in achieving the attacker’s goal. Information collected in this phase is the raw material that is used to build a firm attack. The attackers can obtain a global view of the target, can focus on the weakest link in security, and can obtain enough information to conduct social engineering. If conducted cleanly via passive techniques using publicly available information, this step is anonymous and practically undetectable. Thus, organizations should be very careful with content anonymously available over the Internet and should take effective, measures.

We present a quick analysis tool for security profiling of organizations (should be only used by penetration testers and ethical hackers). We will be working on creating a master list of data which would contain sensitive patterns or signatures. This master list will be used by the tool to compare with the files in search. Only those files will be downloaded which may contain critical information, which should have been kept private and as a secured entity. In future, we also intend to build a tool in Java that would aid in detection of rising cyber-attacks like Cross-Site Scripting (XSS), Distributed Denial of Service (DDoS) along with the countermeasures that should be taken to mitigate them. Also, we will analyze the possibility of creating an Intrusion Detection System using Java much like *Snort IDS* [12]. We are working to implement a vulnerability assessment tool like *Nessus* as part of the future work.

## REFERENCES

- [1] G. F. Lyon, *Nmap Network Scanning: The Official Nmap Project Guide to Network Discovery and Security Scanning*, January, 2009.
- [2] M. Zalewski, *Silence on the Wire: A Field Guide to Passive Reconnaissance and Indirect Attacks*.
- [3] "OSINT Framework," [Online]. Available: <http://osintframework.com/>.
- [4] M. Henriksen, "Gitrob: Putting the Open Source in OSINT," 12 January 2015. [Online]. Available: <http://michenriksen.com/blog/gitrob-putting-the-open-source-in-osint/>.
- [5] M. Henriksen, "A new version of Gitrob is out," 12 March 2016. [Online]. Available: <http://michenriksen.com/blog/new-version-of-gitrob-is-out/>.
- [6] S. Young, "Using Open Source Reconnaissance Tools for Business Partner Vulnerability Assessment," 2014.
- [7] "LaNMaSteR53 / Recon-ng — Bitbucket," [Online]. Available: <https://bitbucket.org/LaNMaSteR53/recon-ng>.
- [8] J. Long, *Google Hacking for Penetration Testers*.
- [9] "Google Hacking Database (GHDB)," [Online]. Available: <https://www.exploit-db.com/google-hacking-database/>.
- [10] "jsoup: Java HTML Parser," [Online]. Available: <https://jsoup.org/>.
- [11] "dnsjava (2.1.7)," [Online]. Available: <http://www.xbill.org/dnsjava/>.
- [12] M. Roesch, "Snort-Lightweight Intrusion Detection for networks," in *Proceedings of LISA '99: 13th Systems Administration Conference*, Seattle, 1999.
- [13] C. F. M. O. Wade Alcorn, *The Browser Hacker's Handbook*.
- [14] P. Engebretson, *The Basics of Hacking and Penetration Testing: Ethical Hacking and Penetration Testing Made Easy*, Syngress, 2011.

# *Towards a Security Baseline for IaaS-Cloud Back-Ends in Industry 4.0*

Elisabeth Bauer, Oliver Schluga, Silia Maksuti, Ani Bicaku, David Hofbauer, Igor Ivkic, Markus G. Tauber  
University of Applied Sciences Burgenland  
Eisenstadt, Austria

Alexander Wöhrer  
University of Vienna  
Vienna, Austria

**Abstract**—The popularity of cloud based Infrastructure-as-a-Service (IaaS) solutions is becoming increasingly popular. However, since IaaS providers and customers interact in a flexible and scalable environment, security remains a serious concern. To handle such security issues, defining a set of security parameters in the service level agreements (SLA) between both, IaaS provider and customer, is of utmost importance. In this paper, the European Network and Information Security Agency (ENISA) guidelines are evaluated to extract a set of security parameters for IaaS. Furthermore, the level of applicability and implementation of this set is used to assess popular industrial and open-source IaaS cloud platforms, respectively VMware and OpenStack. Both platforms provide private clouds, used as backend infrastructures in Industry 4.0 application scenarios. The results serve as initial work to identify a security baseline and research needs for creating secure cloud environments for Industry 4.0.

**Keywords**—Cloud Computing; IaaS; Service Level Objective; Security; ENISA; VMware; OpenStack

## I. INTRODUCTION

Today's working environment is affected by continuous and rapid change, cost saving aspects and the need for flexibility. This influences social as well as technical environments. Cloud computing offers possibilities to handle the technical challenges while being steadily developed. It is defined by the National Institute of Standards and Technology (NIST) as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. The three cloud service delivery models according to NIST are: Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) [1].

The focus of this paper lies on IaaS, which represents the fundamental building blocks for cloud services. IaaS is composed of highly automated and scalable compute resources, complemented by cloud storage and network capability, which can be self-provisioned, metered, and available on-demand [1]. This is also used to provide private clouds, which can have impact on industry, where cloud computing is a part of the proposed concept for the 4th industrial revolution, referred to

as Industry 4.0. In particular in this domain stakeholder built on dedicated infrastructure rather than on public clouds. Dedicated infrastructure can be based within a trusted providers premises or independent divisions in large enterprises.

The cloud servers and their corresponding resources offered by IaaS providers can be accessed via dashboards or APIs. Furthermore, IaaS customers have direct access to their servers and storage. This implies higher order of scalability and lower investment in capacity planning, physical maintenance and management. Compared to other cloud computing models, IaaS is the most flexible, which allows automated deployment of servers, processing power, storage, and networking. IaaS customers have true control over their infrastructure, compared to customers of PaaS or SaaS services. Thus, IaaS is the most used cloud service model, respectively by cloud service providers and customers [2].

As stated above both, IaaS providers and users, interact in a flexible and scalable environment. To handle organizational and legal aspects, SLA are set up to define clear Service Level Objectives (SLO), to measure the service performance and to deal with security issues, which can be very difficult to define. Therefore, measurable security related SLOs are needed to give a clear definition of the covered security aspects. To address this issue towards a security base line for IaaS platforms in an Industry 4.0 environment, the ENISA guidelines are used to extract measurable security related SLO metrics. The measurable security related SLOs act as an input to assess popular industrial cloud platforms, which provide a lot of features, and to assess popular open source cloud platform, which provide a number of customized services. In this paper, the platforms considered for assessment are the VMware vCloud Suite and OpenStack. Moreover, both are compared to highlight the main differences concerning security related SLOs.

Thus, the contribution of this paper is twofold:

- first, we identify a set of topics that still require research and development and
- secondly, we provide a comparison of popular industrial and open-source platforms focusing on private cloud environments, which are important for Industry 4.0 use cases.

## II. RELATED WORK

Related research work has shown the correlation between security standards and SLAs for cloud infrastructure. For instance, the benefits and costs of introducing SLA matching mechanisms for cloud infrastructures are discussed in [3], but there is no explicit focus on security and IaaS in Industry 4.0 environments. Besides that, in [4] the main security concerns and solutions are identified, while [5] presents security and privacy requirements referring to ISO, NIST and CSA. To show the applicability of these three standards, Hoehl [6] compares their service models, whereas the International Telecommunication Union (ITU) analyzed the current standards and guidelines concerning cloud security in [7].

Within these papers, security related SLAs are defined and discussed generically. The focus lies on the contractual relationship between the cloud service provider and the customer and the specified responsibilities between them. Another outcome of these papers is defining common metrics for measuring security service levels. Rathbun [8] shows the importance of security metrics and answers general questions addressing the reasons for having security metrics. Based on this outcome, it is possible to generate the essential information concerning cloud security assessment for our research work.

The CSA has published a new version of the Cloud Controls Matrix. An overview of all common standards and legal publications concerning cloud computing, including ENISA, ISO, NIST, BSI and more are mentioned in [9]. These metrics give an overview of the current state of the art concerning the offer of security standards and guidelines within cloud computing, but no platform's assessment.

Other related work assesses IaaS solutions such as [2], where security vulnerabilities from inside and outside concerning OpenStack are discussed, but without the consideration of ENISA guidelines.

The related work mentioned above consists of general information about security metrics, guidelines and standards as well as comparative aspects between them. Furthermore, it includes the assessment of different cloud computing offers. This paper relies on ENISA guidelines for monitoring of security in cloud contracts (PRO-SEC) [10]. Based on this guideline the extraction of ENISA specific needs regarding SLOs from the cloud service provider's point of view are discussed. This paper focuses on the analysis of the applicability and implementation of the extracted security SLOs on the industrial software suite VMware vCloud and the open source software services of OpenStack. The comparison of both analyses shows the main differences concerning security of both solutions.

To the best of our knowledge, no publication was identified that evaluates IaaS solutions based on ENISA criteria, comparing VMware to OpenStack cloud platforms.

## III. SECURITY RELATED BASELINES

In the area of security, a variety of standards and guidelines concerning cloud computing are in place. This section, gives a brief overview of these standards and guidelines.

### A. International Organization for Standardization

The ISO standards cover a wide spectrum of topics which are generated out of market needs. Security is covered by the ISO 2700x family, which provides an overview and explains the Information Security Management Systems (ISMS), referring to ISMS family of standards with related terms and definitions. The most relevant standards of ISO 2700x family addressing cloud computing security are listed below:

- ISO/IEC 27001 focuses on the requirements for Information Security Management Systems (ISMS) [11]
- ISO/IEC 27002 provides guidelines and general principles for organizations and their ISMS [12]
- ISO 27017 generally focuses on the protection of the in-formation in the cloud services and is built upon the existing security controls of ISO 27002.

The ISO 2700x family provides a set of security criteria and addresses also cloud technologies such as, ISO 27017. Therefore it provides a comparable basis for the further extraction of security criteria within the ENISA guidelines.

### B. National Institute of Standards and Technology (NIST)

NIST is responsible for developing information security standards and guidelines, including minimum requirements for federal information systems. The most relevant guidelines of NIST addressing cloud computing security are listed above:

- NIST Special Publication 800-53 provides a set of security controls which fulfill the security requirements based on organizations, business processes and information systems, including a risk management framework. Additionally, it provides a set of security controls [13].
- NIST Special Publication 800-144 provides an overview of public cloud computing and guidelines for protecting security and privacy [14].

The NIST publications provide comparable security controls and the setup of the NIST Cloud Computing Program (NCCP) stands for further development in this area. Nevertheless, within this research work the status quo in the year 2016 is taken into account.

### C. Cloud Security Alliance

The CSA security guidance working group provides 14 domains about critical areas of cloud computing and works on continuous improvement of the published best practice methods concerning security. The CSA itself is member driven and provides best practice efforts.

### D. Bundesamt für Sicherheit in der Informationstechnik

The BSI Germany provides catalogues concerning IT Security Management and golden rules for application. These catalogues mainly have an organizational view on security and no direct recommendation for the evaluation of IaaS offers. The BSI provides minimum information security requirements for Cloud Service provider.

E. European Commission Guideline

The European Commission (EU) has published in 2014 the “Cloud Service Level Agreement Standardization Guideline” (C-SIG). This guideline focuses on the international orientated standardization of the contractual relationship between a cloud service provider and a cloud service customer. EU aims to make it easier for cloud services customers to compare cloud services offered by different providers by using standardized SLAs. Using a standard set of metrics in a cloud SLA makes it easier and faster to define the SLA. This also simplifies the task comparing one cloud SLA to another. The SLOs covered in C-SIG are: (i) Performance SLOs, (ii) Security SLOs, (iii) Data Management SLOs, and (iv) Personal Data Protection SLOs. The definition of SLOs in C-SIG relies on standards and guide-lines produced by organizations such as ENISA, NIST or ISO/IEC.

F. European Network and Information Security Agency

ENISA is defined as a center of expertise for cyber security in Europe, build up in 2004. The agency’s role is to guide experts towards security solutions that are adapted to the needs of the internal market. Therefore, the development of security approaches and solutions that are interoperable across the whole EU is one of the agency’s aims. In general, guidelines and recommendations published by ENISA are quite unknown within the ITIL community.

In contrast to ISO, NIST, CSA and BSI, ENISA focuses on the European market and is not a member driven organization such as CSA. In the comparison of the standards, none of them includes unique criteria. Moreover, the standards have many similarities in the provided criteria. Due to that fact that ENISA criteria refers to guidelines and standards mentioned above, the criteria is not very different. Procure secure (PRO-SEC) [10] is a guideline, which concentrates on the reporting and alerting of key measurable parameters, as well as a clear understanding of how to manage the customer’s own responsibilities for security. This customer centric approach provides the basis for understanding security risks, which in combination with measurable parameters shows the special focus of ENISA. Due to these aspects, ENISA is used to extract a set of security related SLOs described in the following section.

IV. ENISA EVALUATION

As mentioned before, ENISA has published a guide for monitoring security including service availability and continuity in SLAs (PRO-SEC) [10]. This security monitoring framework defines the following parameter groups (P.G.):

- Service Availability (SA)
- Incident Response (IR)
- Service Elasticity and Load Tolerance (SELT)
- Data Life Cycle Management (DLCM)
- Technical Compliance and Vulnerability Management (TCVM)
- Change Management (CM)
- Data Isolation (DI)
- Log Management and Forensics (LMF)

Table I lists the ENISA P.G.s, together with a short description, a possible monitoring solution and a set of software-defined-datacenter (SDDC) applicable security related SLOs which have been extracted from C-Sig [15]. A SDDC is a data storage facility where all elements of the infrastructure, networking, storage, CPU and security are virtualized and delivered as a service. The entire process from deployment to monitoring and automation of the infrastructure is abstracted from hardware and implemented in software. Typically, the core architectural components that comprise a SDDC include the following [16]: (i) Computer virtualization (compute), (ii) Software-defined storage (SDS), (iii) Software-defined network (SDN), and (iv) Management and automation software (business logic layer - BLC).

In the next section, the applicability and implementation of this set of security related SLOs within VMware’s SDDC and OpenStack is discussed.

TABLE I. ENISA P.G. AND SECURITY RELATED SLOS

P.G.	Description (D) & Monitoring (M)	Security related SLOS
SA	<b>D:</b> clearly defines when a service is considered available and which functions and services are covered by availability monitoring. <b>M:</b> log examination running service health-checks using monitoring tools	<b>SA01</b> Service uptime <b>SA02</b> Percentage of successful requests <b>SA03</b> Percentage of timely service provisioning requests <b>SA04</b> Average response time <b>SA05</b> Maximum response time <b>SA06</b> Level of redundancy <b>SA07</b> Service reliability
IR	<b>D:</b> relates to how a cloud service provider responds to and recovers from incidents. <b>M:</b> collecting log data about time related events, e.g. incident reporting	<b>IR01</b> Support hours <b>IR02</b> Support responsiveness <b>IR03</b> Average resolution time per incident severity
SELT	<b>D:</b> reflects the automatic scale-up/scale-down of resources (compute, network, storage) over a given period of time and is linked to SA. <b>M:</b> burst tests or real-time monitoring of the affected resources, CPU cores and speed, memory size or network bandwidth	<b>SELT01</b> Number of simultaneous services <b>SELT02</b> Maximum resource capacity <b>SELT03</b> Service throughput <b>SELT04</b> Automatic load balancing <b>SELT05</b> Ratio of failed resource provisioning requests to the total number of resource provisioning requests over a commitment period.
DLCM	<b>D:</b> defines the provider’s data handling practices, for example backup and restore strategy, replication and data export or data loss prevention. The SDDCs DLCM relies on backup or snapshots integrity checks or restoration speed durability. The deletion of data and reporting of data-loss-prevention systems can be included. <b>M:</b> backup time stamps and backup/restore durations	<b>DLCM01</b> Data mirroring <b>DLCM02</b> Data backup/snapshot <b>DLCM03</b> Data backup frequency <b>DLCM04</b> Retention time <b>DLCM05</b> Maximum data restoration time <b>DLCM06</b> Percentage of successful data restorations <b>DLCM07</b> Data deletion type

<b>TCVM</b>	<p><b>D:</b> measures the ability of a service to comply with a technical security policy and the handling of vulnerabilities. Regular scans of all affected systems can be used. Additionally, mailing lists and/or public vulnerability reporting programs should be included.</p> <p><b>M:</b> Monitor the time between vulnerability discovery and patch/fix appliance in relation to the class</p>	<p><b>TCVM01</b> Percentage of timely vulnerability corrections</p> <p><b>TCVM02</b> Percentage of timely vulnerability reports</p> <p><b>TCVM03</b> Reports of vulnerability corrections</p>
<b>CM</b>	<p><b>D:</b> monitors critical changes in security relevant properties of the system, for example the loss of certifications or major system upgrades. It relies on time-related change management processes and triggers or the time to implement security-related change requests.</p> <p><b>M:</b> using the main-stream supports duration, count applicable security-related certifications</p>	<p><b>CM01</b> Critical software change reporting notifications</p> <p><b>CM02</b> Percentage of timely software change notifications</p> <p><b>CM03</b> Software security related certificates</p>
<b>DI</b>	<p><b>D:</b> covers discrete access to a shared pool of resources by legitimate users for legitimate purposes and must be present at all time. Additionally, cryptographic quantitative metrics can be included, for example key lengths or hash algorithms.</p> <p><b>M:</b> penetration tests</p>	<p><b>DI01</b> User authentication level</p> <p><b>DI02</b> Authentication</p> <p><b>DI03</b> User access storage protection</p> <p><b>DI04</b> Cryptographic brute force resistance</p> <p><b>DI05</b> Key access control policy</p>
<b>LMF</b>	<p><b>D:</b> covers the access to information about historical events related to the usage of the service allocated to the customer</p> <p><b>M:</b> regular randomized tests of log availability</p>	<p><b>LMF01</b> Log access availability</p> <p><b>LMF02</b> Log retention period</p>

## V. ASSESSMENT OF IAAS PLATFORMS

This paper focuses on the comparison of two well-known IaaS cloud platforms, industrial and open source, concerning security related SLOs. Therefore, the VMware vCloud Suite will be assessed and compared with the OpenStack software, including its components and services, based on security related SLOs extracted from the ENISA guideline.

### A. VMware

VMware is one of the world’s leading suppliers for x86 server virtualization infrastructure [17]. This paper uses VMware’s cloud management platform vCloud Suite. The VMware vCloud Suite (vCloud) includes VMware’s industry hypervisor vSphere and vRealize, the hybrid cloud management platform for providing a SDDC [18]. VMware’s vCloud Suite comprises virtualization technologies for computing, network, storage and management tools for providing SDDC capabilities to the customer.

Within this section the assessment of the applicability and implementation of extracted security related SLOs for each

P.G., as listed in Table I, is described. The assessment addresses the following three VMware products: (i) the free hypervisor vSphere Hypervisor (ESXi), (ii) the industry hypervisor vSphere Suite (vSphere), and (iii) the whole VMware vCloud Suite (vCloud).

**SA:** ESXi offers an interface for monitoring the Service Availability. The measurement of response times (SA04) or redundancy (SA06) is not possible. vSphere and vCloud provide measurement and reporting tools for these parameters.

**IR:** ESXi is a free hypervisor, but does not offer SLO related support for Incident Response. In short, for ESXi this parameter is not applicable. vSphere and vCloud have different support contracts addressing this parameter.

**SELT:** ESXi allows a manually scale-up and scale-down of resources for virtual machines. Service elasticity or (automatic) load tolerance is not available. vSphere and vCloud support service elasticity and load tolerance for both, the virtual machine and the underlying hardware. High availability is also provided by vSphere and vCloud. Tenant specific resource management, capacity and performance management (SELT04, SELT05) are provided by vCloud only. The measurement of resources is available in all three products (SELT03).

**DLCM:** ESXi does not provide access to the storage API. This API is used by backup applications for snapshot management. Therefore, ESXi does not offer DLCM. Snapshots are available in all three products (DLCM02), both vSphere and vCloud provide comprehensive measurement and reporting tools regarding this parameter.

**TCVM:** In general hotfixes and patches are available for all three products. The Security Response Policy offered by VMware states the correction of security issues in a timely fashion (TCVM03). Integrated patch management is provided by vSphere and vCloud, compliance measurement is provided by vCloud only.

**CM:** ESXi does not offer tools regarding software change management. Major upgrades are installed manually. vSphere offers basic tools regarding change management, like software deployment tools for upgrades. vCloud provide comprehensive measurement and reporting tools supporting software change processes (CM01).

**DI:** ESXi does not provide DI in a multi-tenant environment. vSphere uses pools for data isolation within one security boundary. vCloud offers data isolation in a multi-tenant environment. All three products use secure connections for resource access. Different types of user authentication and user rights management (DI01, DI02) are offered by vSphere and vCloud only.

**LMF:** ESXi offers basic log access (LMF01). vSphere and vCloud provide Log Management, supporting different formats, achievement and access on user basis.

To avoid redundancy, the results of the analysis of the applicability and implementation of ENISA parameters on VMware are only presented in the comparison in Table III.

B. OpenStack

OpenStack is an open-source platform for cloud computing providing IaaS. The platform consists of separate modules, where each offers a specific service, and acts as a software-defined-datacenter. The interaction of all OpenStack services enables a high integration to provide IaaS within SDDC. Within this section the assessment of the applicability and implementation of extracted security related SLOs for each P.G., as listed in Table I, is described.

**SA:** A single-controller high availability mode, which is managed by the services that manage highly available environments, is supported by OpenStack (SA01). Redundant controllers for failovers are not configured (SA06).

**IR:** Currently not applicable on OpenStack.

**SELT:** Load balancing as a service (SELT04) is offered by Neutron. Furthermore, Octavia provides additional capabilities for load balancing (including the usage of compute drivers to build instances which operate as load balancer).

**DLCM:** Cinder services can be used for management of volumes and snapshots (DLCM02) for use with the Block Storage API. The self-service API supports end users to request and consume resources.

**TCVM:** OpenStack offers a vulnerability management process, but the manual installation of hotfixes and patches by root is necessary (TCVM03).

**CM:** OpenStack architecture and processes are documented on the community webpage, but there is no automatism to inform the cloud service customer (CM01). Furthermore, there are no supporting change management services within OpenStack.

**DI:** OpenStack provides a separation of tenants. With the implementation of OpenStack’s Identity API, Keystone provides API client authentication, service discovery and distributed multi-tenant authorization (DI02).

**LMF:** OpenStack provides standard logging levels depending on increasing severity. The possible levels are “debug”, “info”, “audit”, “warning”, “error”, “critical” and “trace”. Horizon offers the dashboard for the end user, with limited access to the content of the log-files (LMF01). A Log Management Data Flow by Monasca is provided.

To avoid redundancy, the results of the analysis of the applicability and implementation of ENISA parameters on OpenStack are only presented in the comparison Table III.

C. Results and Discussion

To compare the applicability of SLOs, as described in the previous sections, the following grades have been defined:

TABLE II. GRADES OF APPLICABILITY

Grade	Applicability of SLO
F	The SLO is fully applicable and natively implemented.
P	The SLO is implemented, but cannot be used without additional administrative tasks, e.g.: a feature is natively implemented, but needs to be installed manually (partially applicable and implemented).
N	The SLO is not applicable or not implemented.

After assessing both VMware and OpenStack, these grades have been used to point out whether the applicability and implementation of ENISA security P.G.s were fully applicable (F), partially applicable (P) or not applicable (N). Table III summarizes the results of the comparison:

TABLE III. COMPARISON OF APPLICABILITY AND IMPLEMENTATION OF ENISA PARAMETER GROUPS BETWEEN OPENSTACK AND VMWARE PRODUCTS

ENISA parameter group	OpenStack	VMware		
		vCloud	vSphere	ESXi
SA	P	F	F	P
IR	N	F	F	N
SELT	F	F	P	N
DLCM	P	P	P	N
TCVM	P	F	P	P
CM	P	F	P	N
DI	F	F	P	N
LMF	P	F	F	P

As presented in Table III, the results of the analysis show a high level of applicability and implementation of ENISA SLOs in VMware’s vCloud Suite. Nearly all ENISA parameters are fully covered by vCloud. The major gap is the lack of Data Loss Prevention. In contrast, OpenStack and its services fulfill most of the ENISA criteria partially.

The results show two main outcomes:

- OpenStack does not address the P.G. of IR, whilst in vCloud and vSphere it is natively implemented and applicable.
- In both, VMware and OpenStack the DLCM is only partially implemented.

Hence the contribution is twofold, first we identify a set of topics that still require research and development (e.g.: IR, DLCM) and secondly, as a practical output, we provide a comparison of popular industrial and open-source platforms focusing on private cloud environments, which are important for Industry 4.0 use cases.

VI. CONCLUSION AND FUTURE WORK

In this work, due to the broad applicability of ENISA, the focus lies on the evaluation of ENISA guidelines concerning security assessment of IaaS solutions. The extracted set of security parameters is used for the assessment of the industrial VMware platform and the open source OpenStack platform.

The results of the assessment have shown a high applicability of ENISA criteria on the VMware vCloud, where only one P.G. is met partially. In contrast to VMware vCloud, OpenStack fulfills most of the parameter groups only partially, but only one is not fulfilled at all. Compared to the existing work, the evaluation shows another way for assessing IaaS platforms. The results show the different levels of applicability

of ENISA criteria on an industrial platform, where we expected a lot of features, and on an open source platform, which provides many modules from the open source community.

The results are a first step towards the technical definition of a security baseline for IaaS cloud back-ends, which can be used for Industry 4.0 environments. Furthermore, we have identified a lack of support of IR and of DLCM in IaaS platforms, suggesting a need for research in these topics. To confirm this, future work will investigate current cloud research projects, to find out how well the research community supports this Industry 4.0 related issues. Additionally, we will include a wider set of guidelines to create a more comprehensive catalog.

#### ACKNOWLEDGMENT

This work has been performed as part of the project Power Semiconductor and Electronics Manufacturing 4.0 (SemI40), under grant agreement No 692466. The project is co-funded by grants from Austria, Germany, Italy, France, Portugal and - Electronic Component Systems for European Leadership Joint Undertaking (ECSEL JU).

#### REFERENCES

- [1] P. Mell, and T. Grance, "The NIST definition of cloud computing," 2011. [Online]. Available: <http://faculty.winthrop.edu/domanm/csci411/Handouts/NIST.pdf>. [Accessed 27 11 2017].
- [2] S. Ristov, M. Gusev and A. Donevski, "OpenStack Cloud Security Vulnerabilities from Inside and Outside," in CLOUD COMPUTING 2013: The Fourth International Conference on Cloud Computing, GRIDs, and Virtualization, Valencia, Spain, 2013.
- [3] M. Maurer, V. C. Emeakaroha, I. Brandic and J. Altmann, "Cost and Benefit of the SLA Mapping Approach for Defining Standardized Goods in Cloud Computing Markets," in International Conference on Advanced Computing (ICoAC 2010), Chennai, India, 2010.
- [4] N. Gonzalez, C. Miners, F. Redigolo, M. Simplicio, T. Carvalho, M. Näslund and M. Pourzandi, "A quantitative analysis of current security concerns and solutions for cloud computing," Journal of Cloud Computing: Advances, Systems and Applications 2012, 12 July 2012.
- [5] Cloud Standards Customer Council, "Practical Guide to Cloud Service Agreements V. 2.0," 4 2015. [Online]. [Accessed 20 8 2016].
- [6] M. Hoehl, "Proposal for standard Cloud Computing Security SLAs - Key Metrics for Safeguarding Confidential Data in the Cloud," 2015. [Online]. Available: <https://www.sans.org/reading-room/whitepapers/cloud/proposal-standard-cloud-computing-security-slas-key-metrics-safeguarding-confidential-dat-35872>. [Accessed 26 5 2015].
- [7] ITU, "FG Cloud TR," 2 2012. [Online]. Available: [https://www.itu.int/dms\\_pub/itu-t/opb/fg/T-FG-CLOUD-2012-P5-PDF-E.pdf](https://www.itu.int/dms_pub/itu-t/opb/fg/T-FG-CLOUD-2012-P5-PDF-E.pdf). [Accessed 13 8 2016].
- [8] SANS (Dan Rathbun), "Gathering Security Metrics and Reaping the Rewards," 7 10 2009. [Online]. Available: <https://www.sans.org/reading-room/whitepapers/leadership/gathering-security-metrics-reaping-rewards-33234>. [Accessed 20 8 2016].
- [9] Cloud Security Alliance, "Cloud Controls Matrix v3.0.1," 6 6 2016. [Online]. Available: <https://cloudsecurityalliance.org/download/cloud-controls-matrix-v3-0-1/>. [Accessed 13 8 2016].
- [10] G. Hogben and M. Dekker, "Procure Secure: A guide to monitoring of security service levels in cloud contracts," ENISA, 2012.
- [11] VMware, "VMWARE vCLOUD SUITE," 2017. [Online]. Available: <http://www.vmware.com/files/pdf/products/vCloud/VMware-vCloud-Suite-Datasheet.pdf>. [Accessed 27 11 2017].
- [12] V.V. Fomin, H. Vries, and Y. Barlette, "ISO/IEC27001 information systems security management standard: Exploring the reasons for low adoption," in Proceedings of the third European conference on Management of Technology (EuroMOT), 2013.
- [13] NIST, S., "Recommended Security Controls for Federal Information Systems," 800-53. 2003.
- [14] W. Jansen, and T. Grance, "Guidelines on security and privacy in public cloud computing,". 800-144. 2011.
- [15] "Cloud Service Level Agreement Standardisation Guidelines," 26 June 2014 . [Online]. Available: <https://ec.europa.eu/digital-single-market/en/news/cloud-service-level-agreement-standardisation-guidelines>. [Accessed 31 May 2016].
- [16] DMTF, "Software Defined Data Center (SDDC) Definition: A White Paper from the OSDDC Incubator," 25 June 2014. [Online]. Available: [https://www.dmtf.org/sites/default/files/standards/documents/DSP-ISO501\\_1.0.1a.pdf](https://www.dmtf.org/sites/default/files/standards/documents/DSP-ISO501_1.0.1a.pdf). [Accessed 27 11 2017].
- [17] T. J. Bittman, P. Dawson and M. Warrilow, "Gartner's Magic Quadrant for x86 Server Virtualization Infrastructure," 14 July 2015. [Online]. Available: Retrieved from Gartner database. [Accessed 31 5 2016].
- [18] A. Standard, "ISO/IEC27002," Informationtechnology- security techniques-code of practice for information security controls, (AS ISO/IEC 27002: 2015), Standards Australia, 2015.

# Secure E-Mail Communication – Comparison and Selection of Encryption Solutions Using an Utility Value Analysis Approach

D. Fischer, B. Markscheffel, and K. Scherr

Technische Universität Ilmenau

Ilmenau, Germany

{daniel.fischer | bernd.markscheffel}@tu-ilmenau.de, kilian.scherr@sicher-im-www.de

**Abstract** — E-mail is one of the most frequently used means of communication. Confidentiality, integrity and authenticity are often indispensable in e-mail communication, especially in business use. However, these security objectives can only be guaranteed with the help of additional encryption solutions. Today, there are a variety of client, gateway and software-as-a-service solutions for e-mail encryption. Companies are faced with the challenge of finding the most suitable solution for them. Our research work presents findings from a utility value analysis which provides a comprehensive process for selecting an appropriate solution for securing e-mail traffic. We present the basic principles behind the utility value analysis and how it is used for the evaluation and selection process of e-mail encryption solutions. In addition, we document how our methodology was applied in a practice-based project to make a selection decision between six existing encryption solutions.

**Keywords** — e-mail; encryption solutions; selecting; comparing; utility value analysis

## I. INTRODUCTION

Since Ray Tomlinson's first e-mail in 1971 [1] this asynchronous communication issue has evolved to a mass phenomenon with strong effects to all personal and business spheres. The total number of business and consumer e-mails sent and received per day will reach in 2017 269 billion and is expected to continue to grow to 319.6 billion by the end of 2021 [2]. Especially the impact in business communication is enormous. But in this case, it is essential to ensure that e-mail content remains confidential between the sender and the receiver, especially if the e-mail includes sensitive information like HR data, payment related data, financial records or intellectual property [3]. The main solution is encrypting the e-mail. The reality in practice very often looks quite different. According to a study of Ponemon Institute, 61% of the respondents said that employees sent sensitive data unencrypted through e-mail [4]. Another survey conducted by Osterman Research indicates that e-mail encryption is used "extensively" by only 40 percent of organizations [5]. On the other hand, there is a wide world of e-mail encryption solutions available [6]. This variety of solutions shows that the main problem has changed from looking for a method of e-mail encryption to selecting the most suitable encryption solution.

The objective of our work is to derive a methodology for the selection of a suitable e-mail encryption solution based on the utility value approach. Furthermore, the applicability of the methodology should be presented in a practice-based project.

## II. E-MAIL ENCRYPTION SOLUTIONS AND RELATED WORK

E-mail encryption solutions in this paper are all products that use cryptographic methods to improve the confidentiality, integrity or authenticity of e-mail communications [7, 8]. Based on [9] we distinguish between three different types of products:

- client solutions: cryptographic methods are used directly on the clients of the senders and receivers of e-mails,
- gateway (or server-based) solutions: central, organization-internal systems provide the functions for securing e-mail communication,
- and software-as-a-service (or cloud-based) solutions: organization-external providers enable encryption, decryption, signing and verification of e-mails.

There are a large number of alternatives for all three types of e-mail encryption solutions. Schneider et al. identify a total of 89 e-mail encryption solutions as part of a worldwide survey on encryption hardware and software products [6]. Various market studies assume that the offered amount of solutions will continue to grow and predict a strong growth of the global e-mail encryption market over the next few years [10, 11].

There is certainly literature that contains criteria and recommendations for selecting e-mail encryption solutions. The National Institute of Standards and Technology (NIST) publishes guidelines on the security of electronic messages to help organizations develop, set up and operate their e-mail systems [12]. Requirements and comparison criteria for e-mail encryption solutions can be derived from this, for example. Moecke and Volkamer identify evaluation criteria covering security, usability, and interoperability aspects of e-mail security, and to apply them to existing solutions [13]. Fox describes requirements for e-mail security solutions, specifies important evaluation criteria and gives an overview of the most important standards for e-mail security [14]. On this basis he

develops exemplary decision-making aids and makes recommendations for the selection of suitable solutions. Without evaluating specific products, other authors also compare the different types of e-mail encryption solutions (client, gateway, SaaS), discuss their advantages and disadvantages and derive typical applications and recommendations from them [9]. A systematic and yet flexible methodology for selecting e-mail encryption solutions is currently missing.

### III. METHODOLOGICAL APPROACH AND RESULTS

We use the utility value analysis (UVA) as a basic methodical guideline for structuring and justifying the selection process. The UVA provides a set of complex action alternatives with the purpose of arranging the elements of this set according to the decision-maker's preferences regarding a multi-dimensional target system [15].

Generally, the UVA is structured in following five steps:

- i. Define the target system
- ii. Determine the target returns
- iii. Determine the target values
- iv. Determine the weighting of target criteria
- v. Perform value synthesis [16].

In the following sections we will explain, how we had applied this methodology to support the selection process for an appropriate encryption solution in a practice-based project. In this project, we evaluated six solutions (two client, two gateway, and two software-as-a-service solutions) that were preselected by our practice partner.

#### i.) Define the target system

There is no generally applicable approach for defining the target system. However, it makes sense to aim for a minimization of the targets with simultaneous completeness and redundancy free [8]. We performed a literature analysis to identify the main targets. We queried ACM Digital Library, AISel, EBSCOhost, IEEE Xplore, ScienceDirect, Springer-Link, Google Scholar and a previous selected set of specialized journals with the following terms: "e-mail, encryption, software, solution" combined with "criteria, compar\*, evaluation, select\*, choose". As a result, we were able to compile the functional and non-functional targets shown in Table 1.

TABLE I. COMPILED TARGET SYSTEM

Functional Targets	Non-Functional Targets
High confidentiality of message exchange (VN)	Low installation effort (IA)
High integrity assurance during message exchange (IN)	Low acquisition costs (AK)
High identity assurance of the communication partners (IK)	Low operating costs (BK)
Ensure high availability (VB)	Ease of operation (EB)
Large scope of services (LU)	Good support (SU)
	High trustworthiness of the manufacturer/provider (HV)
	High system/product maturity (SR)
	Good integrability (IB)

#### ii.) Determine the target returns

In the second step, the degree of target achievement must be determined for each target criterion and alternative. First, we derived the target criteria, with the help of the previous selected literature. Then we specified the target returns of the alternative solutions with regard to these target criteria and assembled it in the target returns matrix. Table 2 shows a cutout of the complex matrix for the first target and its target criteria. The complete table with the determination of all target criteria for all six analysed solutions can be found at [17].

TABLE II. CUTOUT OF THE TARGET RETURNS MATRIX

Target: High confidentiality of message exchange (VN)				
Target criteria:				
	Content encryption procedure (VN1)			
	Place of encryption (VN2)			
	Encrypting Mail Attachments (VN3)			
	Transport encryption procedure (VN4)			
	Enhancements to Transport Encryption (VN5)			
Alternative Solutions				
	Mailvelope	Z1 SecureMail Gateway	StartMail.com	...
VN1	OpenPGP	OpenPGP, S/MIME	OpenPGP	
VN2	Client-PC	Gateway (intern)	Startmail-Server (extern)	
VN3	No	Yes	No	
...				

#### iii.) Determine the target values

The relatively freely formulated alphanumeric target returns cannot be compared properly with each other. Therefore, they are transformed into non-dimensional quantities and represented in so-called target values. In our case, the target returns are transformed into target values according to a five-level nominal scale (1 – insufficient ... 5 – very good). In Table 3 we show the same matrix cutout with its target values. The full target value matrix can also be found at [17].

TABLE III. CUTOUT OF THE TARGET VALUE MATRIX

Alternative Solutions				
	Mailvelope	Z1 SecureMail Gateway	StartMail.com	...
VN1	3	5	3	
VN2	5	4	1	
VN3	1	5	1	
...				

#### iv.) Determine the weighting of target criteria

In order to take into account, the individual application context of the encryption solution, which expresses itself through individual requirements, a weighting of the criteria is necessary. For example, costs can have a different weighting depending on the amount of financial resources allocated. In our practice-based project, weights were used on the target level and not on the target criteria level. Furthermore, these weights were collected through a survey where security and IT specialists were interviewed with the help of questionnaires.

The weights of the targets have been passed on to their criteria. Table 4 shows the target weights for our practice-based project.

TABLE IV. AVERAGE WEIGHTED TARGETS

Functional Targets							
VN	IN	IK	VB	LU			
3.8	3.6	3.8	3.2	2.4			
Non-Functional Targets							
IA	AK	BK	EB	SU	HV	SR	IB
2.6	3.0	3.4	2.8	2.8	3.2	2.6	3.6

(4 – important ... 1 – not important; abbreviations see Table 1)

v.) Perform value synthesis

In the final step of the utility value analysis the target values are multiplied with the weights and summed up to the so-called total benefit for each alternative. A detailed documentation of the calculation of the total benefits can be found in [17]. According to the principle of utility value analysis, the alternative is considered to be the "best" with the greatest overall benefit. In Fig. 1 the final results are shown for our set of alternative solutions.

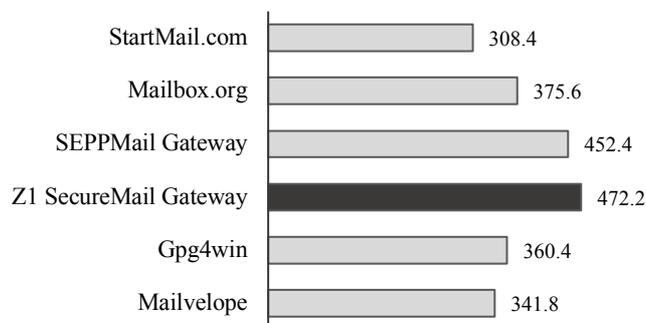


Figure 1. Total benefits of the alternative solutions

IV. CONCLUSIONS

In this paper, we have applied the utility value analysis (UVA) as a decision support concept to the selection process of encryption solutions for e-mail communication. In our practice-oriented project, Z1 SecureMail Gateway from Zertificon is the solution with the greatest total benefit. We were able to show that the selection process is made much more transparent by the UVA in particular because the individual requirements of

users can be taken into account directly in the selection process by weighting the target criteria.

REFERENCES

- [1] T. Campell, *The first email message*, <https://www.cs.umd.edu/class/spring2002/cmsc434-0101/MUIseum/applications/firstemail.html> (Access Date: 11 Sep, 2017).
- [2] The Radicati Group, *Email Statistics Report 2017-2021*, Feb 2017, <https://www.radicati.com/wp/wp-content/uploads/2017/01/Email-Statistics-Report-2017-2021-Executive-Summary.pdf> (Access Date: 11 Sep, 2017).
- [3] Thales, *Global encryption trends 2017*, Apr 2017, <https://gets.thalesecurity.com/pdf/ponemon-global-encryption-trends-study-infographic.pdf> (Access Date: 05 Jul, 2016).
- [4] Pnomenon, *The state of email encryption*, 2017, <https://www.laninfo.tech.com/email-security> (Access Date: 05 Jul, 2017).
- [5] Osterman Research, *Enterprise Encryption and Authentication Usage*, 2016, <https://www.echoworx.com/assets/Enterprise-Encryption-and-Authentication-Usage-A-Survey-Report.pdf> (Access Date: 05 Sep, 2017).
- [6] B. Schneier, K. Seidell, and S. Vijayakumar, *A Worldwide Survey of Encryption Products*, 2016, <https://www.schneier.com/academic/paperfiles/worldwide-survey-of-encryption-products.pdf> (Access Date: 26 Sep, 2017).
- [7] B. Schneier, *Secrets & Lies*, dpunkt, Heidelberg, 2004.
- [8] B. Schneier, *E-mail Security. How to Keep Your Electronic Messages Private*, John Wiley & Sons, New York, 1995.
- [9] DsN - Deutschland sicher im Netz, *Verschlüsselung von E-Mails – Leitfaden zur E-Mail-Sicherheit für Unternehmen*, Berlin, 2015, <https://www.sicher-im-netz.de/sites/default/files/download/leitfaden-e-mail-verschluesselung.pdf> (Access Date: 07 Aug, 2017).
- [10] Technavio, *Global E-mail Encryption Market 2017-2021*, Dec 2016, <https://www.technavio.com/report/global-it-security-global-e-mail-encryption-market-2017-2021> (Access Date: 10 Oct, 2017).
- [11] MarketsandMarkets, *Email Encryption Market by Deployment Type, Industry Vertical, and Region - Global Forecast to 2020*, Nov 2015, <http://www.marketsandmarkets.com/Market-Reports/email-encryption-market-182623205.html> (Access Date: 15 May, 2017).
- [12] National Institute of Standards and Technology (NIST), *Guidelines on Electronic Mail Security*, SP 800-45 Version 2, Feb 2007.
- [13] C. Moecke, M. Volkamer, "Usable secure E-Mail communications – criteria and evaluation of existing approaches," in *Information Management & Computer Security*, vol. 21, no. 1, 2013, pp. 41–52.
- [14] D. Fox, "E-Mail-Sicherheit: Kriterien, Standards und Lösungen," in *Datenschutz und Datensicherheit*, vol. 25, no. 8, 2001, pp. 452–458.
- [15] C. Zangemeister, *Nutzwertanalyse in der Systemtechnik*, Books on Demand, Norderstedt, 2014.
- [16] L. Heinrich, R. Riedl, D. Stelzer, *Informationsmanagement: Grundlagen, Aufgaben, Methoden*, Oldenbourg, München, 2014.
- [17] D. Fischer (Ed.), *Details on the Utility Value Analysis of six E-Mail Encryption Solutions*, 10/2017, [http://www.tu-ilmenau.de/fileadmin/public/iwm/uva\\_email-security.pdf](http://www.tu-ilmenau.de/fileadmin/public/iwm/uva_email-security.pdf) (Access Date: 10 Oct, 2017).

# Towards Comparing Programming Paradigms

Igor Ivkic, Markus G. Tauber

University of Applied Sciences Burgenland  
Eisenstadt, Austria

e-mail: {igor.ivkic,markus.tauber}@fh-burgenland.at

Alexander Wöhrer

University of Vienna  
Vienna, Austria

e-mail: alexander.woehrer@univie.ac.at

**Abstract**—Rapid technological progress in computer sciences finds solutions and at the same time creates ever more complex requirements. Due to an evolving complexity today's programming languages provide powerful frameworks which offer standard solutions for recurring tasks to assist the programmer and to avoid the re-invention of the wheel with so-called "out-of-the-box-features". In this paper, we propose a way of comparing different programming paradigms on a theoretical, technical and practical level. Furthermore, the paper presents the results of an initial comparison of two representative programming approaches, both in the closed SAP environment.

**Keywords**—programming; paradigms; comparison

## I. INTRODUCTION

Programming may be considered by some as an art form and/or by others as a craftsmanship, but it leaves little room for discussion about its incredibly fast development. Whereas in the past developers implemented classic desktop programs, today's applications require world-wide connectivity, web presence and mobile assistance. In many cases a technological leap is followed by a change in the method (or the way the technology is used). While in the past most problems were solved with a structured programming paradigm and a data-driven approach, the new frameworks require an object oriented (OO), generic and model-driven approach [1].

Although frameworks promise to standardize program code, save development time and costs they are often caught in the crossfire of criticism due to the obscure relationship between their pains and gains. The aim of this short paper is to present an approach of how to create a full-comparison of two programming paradigms, namely ABAP - (Advanced Business Application Programming) a 4th Generation Programming Language and BOPF - (Business Object Processing Framework) a modular framework which provides custom services and components, both in the closed SAP environment. In this regard, a comparison is defined as full, when theoretical, technical and practical differences of two programming approaches were identified. So, the idea behind splitting up the approach in these three areas was to look at two software development paradigms from many different angles to create a full-comparison and to, ultimately, paint the big picture.

There are three reasons for choosing two SAP approaches for an initial demonstration of the presented approach. First, SAP is a closed system which makes a technical and practical comparison easier. Second, in 2012 SAP released the Business Object Processing Framework (BOPF) which is an extension of the existing Advance Business Application Programming (ABAP) language. This means, that SAP built and released a new model-driven framework with powerful out-of-the-box

features - programmed with ABAP. Third, due to a lack of documentation and nonexistent literature there were many questions regarding the differences of these two programming approaches within the SAP programming community.

In this short paper, we present an approach for a full-comparison of two programming approaches (ABAP/BOPF) in a closed system (SAP) to identify theoretical, technical and practical differences. First, a literature research provides process- and program-theoretical differences of ABAP/BOPF. Next, for the technical comparison, a performance analysis measures the execution time of CREATE, READ, UPDATE, DELETE (CRUD) operations in an ABAP report compared to a BOPF report. Finally, a use case driven experimental study including a post-experiment User Experience Questionnaire (UEQ) [2] provides practical differences of ABAP and BOPF.

The rest of the paper is organized as follows: Section II summarizes related work in the field, followed by a technical performance analysis in Section III. Section IV deals with an empirical user experience study and a Post-Experiment Survey.

## II. RELATED WORK

The comparison of products, services, programming languages, etc. was the subject of many papers. Most of these papers i.e.: compare a specific aspect of a product and present which product was better, but none compares them on a theoretical, technical and practical level. In example, in [3] the BOPF was firstly introduced to the public and compared to ABAP just on a technical level in a "How-To"-manner. Another comparison in [4] used the UEQ to measure UX in interactive products. Related work in [5] comes the closest, but the research field was restricted to mobile development.

## III. PERFORMANCE ANALYSIS BETWEEN APPROACHES

Even though BOPF was programmed with ABAP the approaches could not be more different. While developing with ABAP goes hand in hand with classic models (i.e.: waterfall model, V-model) the BOPF welcomes agile development. Following the agile principles, the BOPF provides out-of-the-box tools for modelling, developing, testing and finalizing the logic of an application while involving the customer in the development process from the outset. Technically speaking, using ABAP for programming means starting from the greenfield and building an application from scratch. This greenfield-approach is a gift and a curse at the same time. For this reason, the BOPF requires first the creation of a model of the Business Object (BO) and then uses an OO Application Programming Interface (API) to control it programmatically. To be able to understand how the API works it is necessary to compare the CRUD operations.

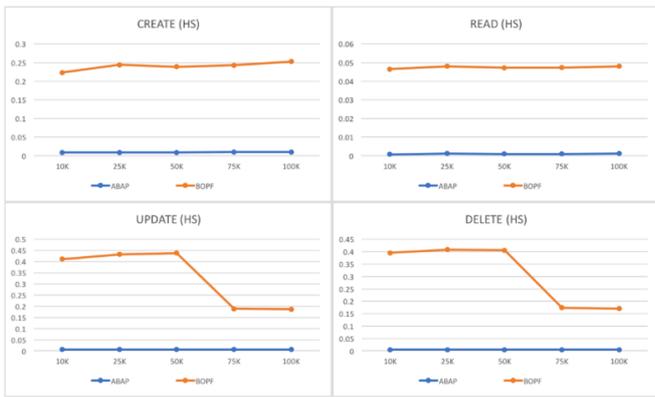


Figure 1. Performance (execution time in ms) comparison of CRUD operations (ABAP vs. BOPF) with increasing data block sizes

The main difference is that each CRUD operation has its own command in ABAP while the BOPF uses two different methods for the same purpose. The MODIFY method is used to CREATE, UPDATE and DELETE records, while the QUERY method reads the records from a database. In conjunction with applications where performance plays a big role the question arises if building it with the new generic framework is the right decision. To make a direct comparison of CRUD operations and their performance on a HANA system (HS) a performance analysis was conducted. The idea was to implement eight applications (4x ABAP, 4x BOPF) where each executed one of the CRUD operations. These applications were executed one after another with five different data amounts (n = 10K, 25K, 50K, 75K, 100K). So, by choosing this approach it was guaranteed that neither the two approaches (ABAP/BOPF) nor their CRUD operations were mixed during a performance analytical run. Fig. 1 compares the results of the CRUD operations and shows the performance advantages of ABAP vs. BOPF on the HS.

#### IV. DEVELOPER EXPERIENCE EXPERIMENT AND SURVEY

An experimental study was designed in two different phases which were decomposed into individual results. In the first phase eight test persons (TP) were randomly divided into two groups and confronted with a simple clear-cut programming task. The sole limitation in this scenario was that one group uses ABAP while the other one must only use the BOPF for the implementation. Building on that a survey was conducted to evaluate the user experience (UX) during the experiment. The TP answered a three-parted UEQ questioning the UX in general and how well the development environment of ABAP/BOPF supported them in solving their task. Fig. 2 shows how much time it took the TP for data modelling, programming, testing, and UI-design.

The experiment revealed that it took the BOPF group 40 minutes on average to complete 100% of the required tasks, while the ABAP group failed to implement all requirements (the TOP 2 TP finished approx. 50% of requirements after one hour; the experiment was terminated after one hour due to the superior results of the BOPF group). Finally, the post-experiment survey showed that the BOPF was rated better than ABAP by the TP. They criticized among others that

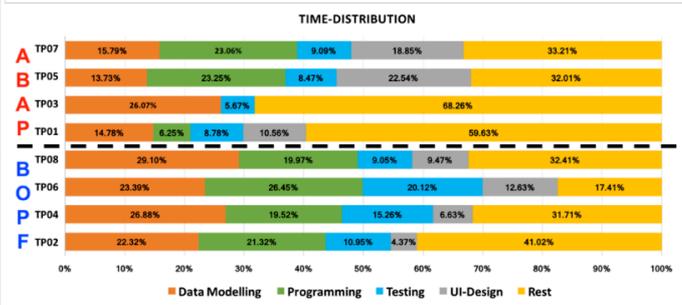


Figure 2. Time-Distribution of the Development Process

programming with ABAP was inefficient and complicated in comparison to the BOPF.

#### V. CONCLUSION AND FUTURE WORK

The main goal of this short paper was to present an approach which enables a systematic comparison of two software development approaches covering two aspects. In the first aspect, the comparison is done on a technical level where the performance of the program code of the two approaches was measured. The second aspect dealt with analyzing the differences of two approaches by applying them in a close to reality, use case driven experimental study.

For future work, we will develop a single metric to reflect and combine the results of all three comparisons. We will further investigate whether the results of the performance analysis and the experimental study would change when larger data amounts (n > 100K) and more than eight TP are used. Another idea for future work would be to execute the performance analysis using applications with more complex program logic (more than CRUD operations). As for the experimental study, it would also be great to find out, if the outcome of the experiment would be the same, when a more complex application is given for the implementation. Finally, it would be great to apply this approach on many different programming languages and frameworks to point out significant differences in the results.

#### REFERENCES

- [1] France, R., & Rumpe, B., "Model-driven development of complex software: A research roadmap", 2007, Future of Software Engineering, pp. 37-54.
- [2] Schrepp, M., Hinderks, A., & Thomaschewski, J., "Applying the user experience questionnaire (UEQ) in different evaluation scenarios", 2014, International Conference of Design, User Experience, and Usability, pp. 383-392.
- [3] Hardy, P., "ABAP to the future", 2015, SAP PRESS.
- [4] Rauschenberger, M., Schrepp, M., Cota, M. P., Olschner, S., & Thomaschewski, J., "Efficient measurement of the user experience of interactive products. How to use the user experience questionnaire (ueq). example: spanish language version", 2013, IJIMAI, 2(1), pp. 39-45.
- [5] Nitze, A., & Schmietendorf, A., "Qualitative und quantitative Bewertungsaspekte bei der agilen Softwareentwicklung plattformübergreifender mobiler Applikationen", 2014, Logos Verlag Berlin GmbH.

## **Session 19: Intelligent Control**

Title: Applying Dynamic Verification Tagging to the k-Anonymity Model  
(Authors: Ahmad Bennakhi, Mohamed A. Jeragh)

Title: A Study of Factors Effecting Thailand Talent Mobility Programme: Case of University and Food Technology Industry  
(Authors: Kritika Kongsoontornkijkul, Rath Pichyangkura, Pakpachong Vadhanasindhu, Kanlaya Vanichbuncha)

Title: shinySDM: Point and Click Species Distribution Modeling  
(Authors: Thomas Nash, Aspen Olmsted)

Title: Efficient Retrieval of Information from Hierarchical REST Requests  
(Authors: Seth Stoudenmier, Aspen Olmsted)

# Applying Dynamic Verification Tagging to the $k$ -Anonymity Model

Ahmad Bennakhi

Kuwait University

Kuwait University, Kuwait City, Kuwait

Email: a\_bennakhi@hotmail.com

Mohamed A. Jeragh

Kuwait University

Kuwait University, Kuwait City, Kuwait

Email: mohammad.jeragh@ku.edu.kw

**Abstract**—This paper proposes a new way of verifying obscured data ( $k$ -anonymity model) when data mining. This is achieved by hashing the unique key along with other values by using a one-way hash function so that the attribute is referenced by the resulting value of this hash operation. The unique key or ID is hashed using multiple values; one from each attribute. When the data miner requests multiple attributes for the purpose of data mining, the new reference/verification key is derived using the attributes of the table requested in order for the data fetched to be verified by a trusted third party. The proposed solution is compatible with all extensions of the  $k$ -anonymity method, especially  $l$ -diversity.

## I. INTRODUCTION

Data mining has attracted more attention in recent years, since big data solutions are becoming ubiquitous. Data mining is the process of discovering interesting patterns and knowledge from large amounts of data. As a highly application-driven discipline, data mining has been successfully applied to many domains, such as business intelligence, web search operations, scientific discovery, and digital libraries [1].

During recent years, the progress of hardware technology has made it easy to store and process large amounts of transactional information. Even simple transactions of everyday life such as using the phone or credit-cards are recorded today in an automated way. Such information is often private and should not be open to the public. Depending on the nature of the information, users may not be willing to reveal sensitive information about themselves. Data mining techniques are considered a challenge to privacy preservation due to their natural tendency to use sensitive information about individuals. This has led to a considerable amount of focus on privacy preserving data collection and mining methods. An innovative approach for privacy preserving data mining was proposed that relies on two facts [2]:

- 1) Users are not equally protective of all values in their records. Thus, users may be willing to provide modified values of certain fields by using a (publicly known) perturbing random distribution. This modified value may be generated using a custom code or a browser plug-in.
- 2) Data mining problems do not necessarily require individual records, but only distributions. Since the perturbing distribution is known, it can be used to reconstruct

aggregate distributions, i.e. the probability distribution of the data set. In many cases, data mining algorithms can be developed which use the probability distributions rather than individual records.

A popular emerging approach to preserve privacy while data mining is the  $k$ -anonymity model. This model was only relatively recently patented in 2002[3] and has already been the subject of many studies. The  $k$ -anonymity model works by applying suppression and generalization on data attributes. The objective of this paper is to develop a practical and secure way of verifying/authenticating data that has been obscured for the purpose of data mining, while also leaving duplicate markers such that the data miner could spot the duplicates himself in case the data provider did not intercept them or they were duplicated during the transmission phase.

The rest of section 1 will describe the concepts in which our proposed idea is associated with. While section 2 will go on and shed some light on the problem that our concept is trying to solve. Section 3 will mainly be about the proposed concept solution. Section 4 will be about the concepts application and the calculations involved. Section 5 will focus on the work related to our topic and the other proposed solutions that are related. Finally, the paper will end in section 6 with a conclusion.

### A. Types of Attributes in Tables

Different attributes do not just have different levels of sensitivity, but they also have different levels of disclosure. A non-sensitive attribute can pin-point the identity of the user if not properly concealed, especially if the attacker has background knowledge helping him reveal the identity of the target. This is called a background knowledge attack. When it comes to privacy in databases, attributes can be classified into 4 types:

- Identifier (ID): This attribute is normally used as the primary key in the tables, but it can also be the attribute that would be used to pinpoint the entity. Examples: social security number, passport ID, or mobile number.
- Quasi-Identifier (QID): Any attribute that can be used to trace and identify the entrees on tables. Examples: age, sex, address, and other similar attributes.

This work was not supported by any organization

- Sensitive Attributes (SA): Attributes that people do not want others to see. Publicly linking these attributes to the entities related would result in breach of privacy.
- Non-Sensitive Attribute (NSA): Any other attribute that is not mentioned above and does not compromise privacy whatsoever.

B. *k*-Anonymity Model

The *k*-anonymity model is very successful due to both of its simplicity and flexibility to hide/obscure data. The whole goal of the *k*-anonymity model is to suppress and generalize the values of the quasi-identifiers such that every tuple will look identical to at most  $k - 1$  tuples. Hence the *k* in this model stands for the least number of tuples that are identical within a given table.

If we apply the *k*-anonymity model to Table 1, where  $k = 2$  it would look like Table 2. As you can see the last two entries are a match with the exception of the the non-sensitive attributes. With *k* attributes matching, this will give us a probability of no more than  $P = 1/k$  of the tuple being identified.

Generally, suppression is applied to identifiers and to a lesser extent on some sensitive attributes. Suppression is where an asterisk replaces all or some of the values in a given attribute. While modifying the quasi-identifiers the generalization usually replaces all quasi-identifiers. Generalization is when a given number is replaced by a more general value. For example, if an entry represents a 27 year old person, his age will only appear as a person aged 20 to 29.

TABLE I. Regular table

Name	Age	Sex	Disease
John	5	M	Eczema
Mohammad	7	M	Flu
Lucy	10	F	Celiac
Sabina	13	F	Bronchitis
Daichi	15	M	Cancer
Grace	16	F	Arthritis
Yu	19	M	HIV
Olga	28	F	Hepatitis
Christina	29	F	HIV

TABLE II. After applying *k*-anonymity(sorted by age range)

Name	Age	Sex	Disease
*	[1,10]	NA	Eczema
*	[1,10]	NA	Flu
*	[1,10]	NA	Celiac
*	[11,20]	NA	Bronchitis
*	[11,20]	NA	Cancer
*	[11,20]	NA	Arthritis
*	[11,20]	NA	HIV
*	[21,30]	NA	Hepatitis
*	[21,30]	NA	HIV

C. Anatomization

Anatomization is the operation that does not modify the quasi-identifier or the sensitive attributes, instead it disassociates the relationship between two attributes [1]. This is done by sending two or more separate tables instead of one by splitting

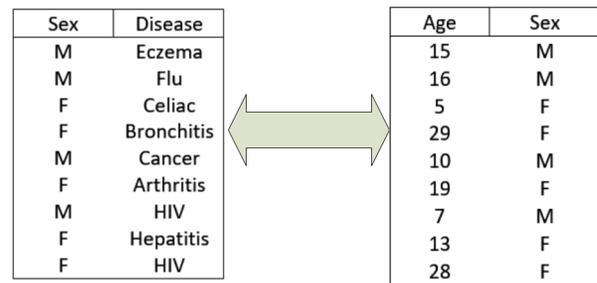


Figure 1. The two resulting tables after applying anatomization on Table 1

them apart. Some of the attributes may be redundant in the tables split to match the data miner’s goal.

If Table 1 would be anatomized, then the resulting two tables could be Tables 3 left and right. This is not the only variation of how Table 1 could be anatomized, hence our proposed solution has to be flexible enough to take this concept into account. Models such as the *l*-diversity use anatomization to make it even harder for attackers to identify the tuples [4].

II. THE OPTIMAL *k*-ANONYMITY PROBLEM

Now let us keep in mind that the higher *k* is, the less identifiable and more privacy the entity gets. However, if *k* is a large number then there will be insufficient data that the data miner could use for his research or analysis. There has been lots of research done to solve the optimal *k*-anonymity problem, but so far it is still an NP-hard problem [5][6]. The best that researchers could do is come up with an approximation and heuristics algorithms to come up with a near optimal *k* [7][8]. The other problem with a miscalculated high *k* is that duplicates can not be detected on the data miner’s side since he may have multiple tuples that look indistinguishable as seen in Table 3.

Traditional tagging of tuples can solve the problem of establishing identity while presumably hiding the ID, but does not maintain the privacy of the tuples in case an untrusted data miner with very high *k* gets his hand on the data of a trusted data miner with a very low *k* since he can just cross reference the tag. The same data miner can also exploit tuple tagging, especially if anatomization is used as a privacy safeguard without suppression or generalization. The data miner can simply request multiple tables then merge them using the tuple tag number.

TABLE III. After applying *k*-anonymity(sorted by age range)

Name	Age	Sex	Disease
*	[0,9]	*	Eczema
*	[0,9]	*	Flu
*	[10,30]	*	Celiac
*	[10,30]	*	Bronchitis
*	[10,30]	*	Cancer
*	[10,30]	*	Arthritis
*	[10,30]	*	HIV
*	[10,30]	*	Hepatitis
*	[10,30]	*	HIV



Figure 2. The general order of events of our proposed dynamic tagging process.

Changing the tag number or randomizing it for each request is a valid solution for the verification of tuples in a high  $k$ -anonymity problem. However, this would require a large amount of data (all tags and their corresponding IDs or keys) stored and possible overheads since documenting every request is a tedious matter especially when big data is concerned. While many papers did indeed recognize the problem of duplicates in the  $k$ -anonymity model [9][10], they either provide a solution by a hefty union operation or just completely keep duplicates out of their assumption model. A straight forward, dynamic, and robust method of tagging must be developed in order to satisfy all of the problems faced so far. The method should be compatible with anatomization since the  $l$ -diversity model mainly relies on it[4].

### III. APPLYING DYNAMIC TAGGING ON THE $K$ -ANONYMITY MODEL

Not every table is transferred with all of their attributes, therefore attributes themselves make good variables since they are likely to change with each request. This makes them suitable variables in our equation along with the ID (primary key) to hash out an adequate dynamic tag. The hash function would look like something like this:

$$H(ID, a_1 \times a_2 \times a_3 \times \dots a_n) = D_{tag} \quad (1)$$

Where “ID” represents the unique identifier or the primary key. The  $a$  is a number that represents the attributes that are present in the table sent to the data miner. The product of all the attributes has to be unique, such that:

$$a_1 \times a_2 \neq \forall a_x \times a_y \quad (2)$$

Hence each attribute has to be assigned a prime number. Due to prime numbers unique product property there should not be any collisions between different attribute products. Hence the transformation will look like something like the Figure 2.

If a table is queried and asks for a particular set of attributes then the tuples will have different tags from another table which has a different combination of attributes, even if both tables came from the same original table. Figure 3 demonstrates the effects of the dynamic tagging technique. Applying dynamic tagging to these tuples will allow the data miner to identify the duplicates (if they existed) no matter how high the  $k$  is, since each one of the duplicates has a different tag. If it occurs that a duplicate exists then it will have the same tag, hence allowing the data miner to detect it and get more accurate results.

To optimize this solution, the attributes and their corresponding prime numbers can be stored in a cache storage or any other fast data retrieval storage mechanism. If an attacker manages to change the  $k$  in the anonymity model while requesting a table with the same attributes, that could expose certain data and with enough requests he could get all of the tuple’s data with the exception of the ID. This is still risky since the tuple can be identified with enough quasi-identifiers (QID) being uncovered. When  $k$  is set on a given number, the implementation of the  $k$ -anonymity algorithm may decide to hide a particular set of quasi-identifiers. If  $k$  decreases, some of the hidden quasi-identifiers may be partially exposed while others will still be hidden, depending on the implementation of the  $k$ -anonymity model.

#### A. Breach of Privacy by Comparing Tags from Different Tables with Background Knowledge

Data miners can be assigned a different  $k$  depending on the data intended to be mined. Both tables 4 and 5 extracted data from the same data collector and both requested the same attributes. There are similarities and differences regarding the tuples since they are requested in different time periods and different constraints. Suppose that an attacker already has table 5, if he gets table 4 then he could supplement some of the data that he has using the tagging process that we suggested. Leaving such tags in tables would no doubt be exposing the data in the tables for a background knowledge attack.

TABLE IV. With  $k=2$

Tag	Name	Sex	Age	Children	Illness
xy8z	D*	Male	[30,39]	No	Diabetes
abc6	D*	Male	[50,59]	Yes	Diabetes
5plm	C*	Male	[30,39]	No	AIDS
78rv	B*	Female	[10,19]	No	Stroke
pd2l	A*	Male	[20,29]	Yes	Migraine
124j	A*	Male	[20,29]	Yes	Ulcer
bd32	Z*	Female	[20,29]	Yes	Migraine

TABLE V. With  $k=3$

Tag	Name	Gender	Age	Children	Illness
zobq	*	*	[60,69]	5	Stroke
abc6	*	*	[50,59]	2	Diabetes
tpf4	*	*	[40,49]	0	Fracture
78rv	*	*	[10,19]	0	Stroke
pd2l	*	*	[20,29]	3	Migraine
kv43	*	*	[20,29]	3	Warts
bd32	*	*	[20,29]	3	Migraine

#### B. Dynamic Tagging Compatible with a Variable $k$

The problem described in the previous section can be easily solved by including the  $k$  in the hashing formula so that its final form will look like:

$$H(ID, \prod_{i=1}^n a_i \times k) = D_{tag} \quad (3)$$

The attribute’s cumulative product multiplied with the  $k$  assigned will only be calculated once per table, while hashing

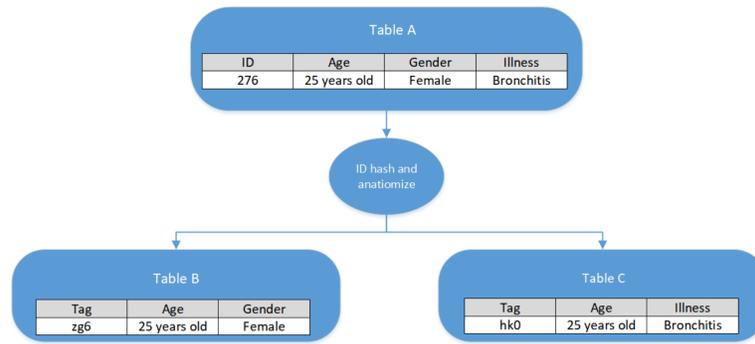


Figure 3. A demonstration of how the proposed process can enhance data privacy while applying anatomization

will be performed for each tuple once. This will make tables with different  $K$ s to have completely different tags, hence identifying the data from different tables will be hard. The  $k$  does not have to be saved for verification as a randomized tagging process would require, it can be calculated with a simple query if the corresponding prime numbers for the  $k$  and attributes are present. Thus, the complexity of applying this tagging technique will hence be:

$$\sum D_{tag} = O(n) \quad (4)$$

By using this formula, the tags will be different with the application of a differing  $k$ , hence only tables with the same  $k$  and same attributes will have the same tags for the tuples. One of a hash function's main properties is that they are supposed to be very fast [7], hence the application of such a tagging system will not take much time in exchange for the additional property of verification. In case there happens to be a collision, the penalty will be the possibility of uncovering of an ID. That means that a hashing technique with good collision resistance would greatly complement our proposed concept. Another weak point of applying this concept would be that if the hashing method and key is known, then extracting data with few attributes would make it easier for attackers to brute force their way into uncovering the ID since the  $k$  is not hidden.

### C. Dynamic Tagging Applications

The process of dynamic tagging concerns only 2 sides: the data collector and the data miner. The data collector is the entity responsible for maintaining the privacy of his data and is also in charge of providing the data to the data miner. It is worth noting that this tagging concept is also applicable to the  $l$ -diversity model. It works even better with the  $l$ -diversity since the attributes are separated, hence giving the entries more privacy by giving them a totally different tag. The data miner is supposed to get as much data as possible without compromising the privacy of the data given to him. The process starts in the following way:

- 1) The data miner requests a particular set of data from the data collector.

- 2) The data collector recognizes the data miner and decides to give him tables obscured by a certain  $k$ -anonymity value.
- 3) Before the tables are sent, the product of the attribute's prime numbers and the  $k$  are calculated.
- 4) Each ID is hashed using the product calculated in step 3, then joined to their corresponding tuples in the tables.
- 5) The data collector finally sends the tagged tables to the data miner.
- 6) The tables are received by the data miner and could be scanned for duplicates. The table could be also sent to a trusted third party for the verification of the data's authenticity.

Figure 4 depicts the steps that each side performs. The prime numbers of each attribute could also be given to a highly trusted third party so that they could perform the verification of the data in case the data miner was suspicious wanted some closure regarding the authenticity of the data. In case the data miner suspects that the tables that he is receiving do not have legitimate data or that he just needed verification from a trusted third party, the required procedure will be as follows. First, the data miner will send his sanitized tables (whether by the  $k$ -anonymity model or any of its extensions) to the third party. The third party will send the data collector a request for the verification of the received table. The data collector will recognize the tables and send the originals to the third party, which will then compare it to the sanitized tables to check if they match up. Finally, after the comparison, the third party will either verify a total match or a mismatch.

## IV. RELATED WORK

The field of privacy persevering data was recognized as an increasingly developing one, that is why survey studies such as [11] was written to summarize the concept behind the techniques that were being used. Privacy preserving publishing has been a subject of recent research, especially in the social media sector [8]. A study done in 2014 tackles the problem of how privacy is breached when cross referencing and how tag name suppression could/could not solve the problem. The paper explains the general idea of the individual's privacy concerns and the approaches in which individual's privacy



# A Study of Factors Effecting Thailand Talent Mobility Programme: Case of University and Food Technology Industry

Kritika Kongsoontornkijkul<sup>1</sup>, Rath Pichyangkura<sup>2</sup>, Pakpachong Vadhanasindhu<sup>1</sup>, and Kanlaya Vanichbuncha<sup>3</sup>

<sup>1</sup>Technopreneurship and Innovation Management Program, Graduate School, Chulalongkorn University, Bangkok, Thailand  
E-mail: kkritika@hotmail.com, pak.v@acc.chula.ac.th

<sup>2</sup>Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok, Thailand  
Email: prath@chula.ac.th

<sup>3</sup>Department of Statistics, Faculty of Commerce and Accountancy, Chulalongkorn University, Bangkok, Thailand  
Email: kanlaya@cbs.chula.ac.th

**Abstract**—The main purpose of this study is to explore factors which have high impact on national research university researchers' decision making to work with food technology industries in order to strengthen the competitiveness of technological innovation in Thailand. Semi-structured and in-depth interviews were conducted for this study in order to understanding factors of success and barriers to successful personnel sharing. It also identifies recommendations and solutions for personnel sharing mechanism in Thailand.

**Keywords**—*University-Industry Linkage (UIL); University-Industry Collaboration (UIC); Talent Mobility (TM); National Research University (NRU); Personnel Sharing*

## I. INTRODUCTION

Science, Technology, and Innovation currently play a crucial role in determining the direction of globalization, increasingly influencing social and economic in terms of improving the quality of human life, and also being an important economic driver in which creating wealth, competitive advantage, and sustainability.

Due to this importance, experts in science, technology, and innovation are a valuable and important resource for strengthening the economy by bringing their knowledge and research into the industry to grow efficiently.

The Thai governmental policies from the National Science and Technology Strategic Plan (2004-2013) to the National Economic and Social Development Plan (2017-2021), and the current model of Thailand 4.0 focus on human resources and labor development in science and technology leading to innovation. These mentioned policies have been driven to propel the Thai businesses to be competitive in the international arena [19].

Most researchers in Thailand, however, are personnel in universities and government research institutes rather than in the private sector, especially the industrial operations, which are the main driving force of the country's economy. Experts in production, product development, and the use of

technology and innovation are much essential for the competitiveness and affect the macroeconomic situation in Thailand.

The National Science Technology and Innovation Policy Office (STI) has recognized these problems and initiated the project to encouraging scientific, technological, and innovative personnel from public and higher education institutions to work in industries in order to strengthen their competitiveness in the manufacturing and service sectors called "Talent Mobility" (TM). From this supporting policy, industries are able to develop and upgrade their research capability and research potential to reinforce their productivity. The objective of the TM programme is to encourage researchers in educational and governmental research agencies to work with private industries in order to enhance the competitiveness of technological innovation of Thai industry [17].

The Talent Mobility project, however, has not received much attention and response from government researchers, especially research universities. Even though the Thai Governmental Talent Mobility Programme has been initiated for a few years, only limited universities and companies have participated in this program. Reference [2] describes the TM policy evaluation of linkage between universities/research institutes and private companies around the country in general.

This research, therefore, focuses on the efficient and sustained sharing of scientific, technological, and innovative personnel between universities and industries emphasizing on the case of National Research University (NRU) and food technology industry.

The main purpose of this study is to explore crucial factors effecting university researchers to work with industries in order to strengthen the competitiveness of technological innovation in Thailand. Semi-structured and in-depth interviews were used to collect data. The target

samples were selected from professors in national research universities who were in the field of food science or food technology. This study contributes to deeper understanding of the problems encountered with Talent Mobility Programme and to identify solutions for personnel sharing mechanism.

## II. LITERATURE REVIEWS

### A. Models of Talent Mobility

World Economic Forum (WEF) defines Talent Mobility as "the physical mobility of talent within or across organizations and industries as well as the professional movement of workers across occupations or skill sets" [22]. It is a mechanism which contributes to the creation and diffusion of both codified and tacit knowledge. Talent Mobility is especially relevant for the transfer of tacit knowledge, any form of knowledge that cannot be codified and transmitted as information through documentation, academic papers, lectures, conferences, or other communication channels. The transfer of this form of knowledge is more effective through interactions among individuals with a common social context and physical proximity.

Talent Mobility is not a goal in itself but is often linked with sustainable economic growth objectives. A study by WEF describes Talent Mobility as a policy instrument to achieve balance within global human capital markets and to stimulate national economic growth [22]. Thus, Talent Mobility can promote research and innovation as well as increase employability and promote career development for researchers.

WEF categorized Talent Mobility practices into four groups according to their levels of collaboration as summarized in the Table I.

TABLE I. TYPES OF TALENT MOBILITY PRACTICE

Level of Collaboration	Description	Examples of Key Practice
<b>1. Collaboration within the organization</b>	Collaboration across functions, units and geographies within an organization to develop employees, close information gaps, and better balance internal supply and demand.	<ul style="list-style-type: none"> <li>- Forecasting the supply and demand of critical talent;</li> <li>- Career and leadership development focusing on building critical skills;</li> <li>- Integrated diversity and inclusion strategy;</li> <li>- Global mobility philosophy aligned with talent development strategy;</li> <li>- Strategic succession planning;</li> <li>- Promoting internal mobility across business units and job functions</li> </ul>
<b>2. Collaboration across organizations within a country</b>	Collaboration among different organizations to source and develop talent locally.	<ul style="list-style-type: none"> <li>- Sending employees to other organizations;</li> <li>- Partnerships between companies, governments or educators on training, developing, and deploying talent;</li> <li>- Public sector initiatives on sharing information on labour supply and</li> </ul>

Level of Collaboration	Description	Examples of Key Practice
		demand; - Programmes led by the public sector to mitigate brain drain and facilitate immigration
<b>3. Collaboration on an industry or regional level</b>	Public-Private Partnerships designed to foster talent mobility and skill development, as well as industry associations working closely with the public sector to attract and develop talent.	<ul style="list-style-type: none"> <li>- Strategic talent assessment, development and deployment on an industry level;</li> <li>- Matching supply and demand through job fairs, job portals, and university visits;</li> <li>- Shaping academic curricula through participation in university advisory councils;</li> <li>- Subsidized internship programmes;</li> <li>- Industry specific training programmes and workshops</li> </ul>
<b>4. Collaboration on a global or multi-stakeholder level</b>	Sectors, governments, international organizations and academia across multiple countries and regions work closely together to solve complex talent mobility issues.	<ul style="list-style-type: none"> <li>- Private companies talent sourcing for educational institutions, governments, and NGOs in multiple countries;</li> <li>- International development initiatives in skill development and trade agreements</li> </ul>

Source: World Economic Forum, 2012. [22]

The Talent Mobility Programme in Thailand by The National Science Technology and Innovation Policy Office (STI) has been officially launched since 2013 dividing into four phases. The details of implementation models are described in Table II.

TABLE II. MODELS OF THE TALENT MOBILITY PROGRAMME

Model	Implementing Bodies	Financial Support
1	National Science and Technology Development Agency (NSTDA)	Up to 70% of the project budget but not exceeding 400,000 THB
2	Clearing Houses (CH)	1) 1.5 times salary of researchers compensated to universities 2) 8,000-12,000 THB monthly allowances for research assistants
3	Clearing Houses (CH) and other universities	1) 1.5 times salary of researchers compensated to universities 2) 8,000-12,000 THB monthly allowances for research assistants
4	Office of the Higher Education Commission (OHEC)	1) 400,000 THB based on FTE of researchers 2) 200,000 THB for testing and materials The project can receive financial support from both OHEC and STI

Source: Talent Mobility Project, STI, 2016. [2]

Remarks:

Clearing House (CH) means a facilitating center where matching the industries with researchers from the Talent Mobility database, and administrating contracts and disbursements of the programme.

Full-time Equivalent (FTE) means a unit that indicates the workload of an employed person (or student) in a way that makes workloads or class loads comparable across various contexts. FTE is often used to measure a worker's or student's involvement in a project, or to track cost reductions in an organization.

*B. University and Industry Linkage (UIL)*

Key factors in the innovation process involve interaction and collaboration among three players [6]: university, industry, and government [11]. Successful university–industry linkage (UIL) can be measured by efficiency, continuous marketing sales growth, understanding between university and corporate cultures, understanding of industrial problems, continuity of technology transfer to industry, and communication [8].

For Thailand, the Thai government has initiated the National Research University (NRU) project in 2009 under the direction of the Office of the Higher Education Commission (OHEC). The goal is to develop academic capacity to promote research activities in the country and to build links between universities and industries (UIL) for Thailand's competitiveness. Nine outstanding research universities were selected including Chulalongkorn University; Kasetsart University; Khonkaen University; Chiang Mai University; Thammasat University; Mahidol University; Prince of Songkla University; King Mongkut's University of Technology Thonburi; and Suranaree University of Technology [16].

In general, there are six official forms of personnel sharing between universities and industries in Thailand: Consultancy and Reach-Out; Licensing; Contract Research; Collaborative Research; Talent Mobility; and Academic Entrepreneurship [17].

In Thailand, however, UIL remains weak because of linkage limitations, low levels of innovation consulting, inflexible structures, complicated supports, lack of funds and unsupportive government policy [4].

*C. Key Success Factors of Thailand Talent Mobility Programme*

Factors affecting the success of TM programme include four aspects; organization management, regulations, networks, and industrial development in the regions [2]. First, organization management which clearing house staff members play an important role of handling the mobility services and organizing the overall project. Second, flexible regulations of an autonomous unit can effectively accelerate the programme implementation. Many success cases came from the autonomous unit that operates flexibly and efficiently allocating resources to the programme. Third, cooperation networks of clearing houses strongly link universities and industries. Some members have relevant experience in the Innovation and Technology Assistance Programme (iTAP) which is beneficial to the TM project.

Lastly, industry-oriented perception among the university executives, especially in small-sized technology universities participating in the project, is another key success factor contributing to the quality of the TM programme.

*D. Main Obstacles of Thailand Talent Mobility Programme*

The main obstacle for the success of Talent Mobility Programme is the career path of university professors which does not take collaboration with industries into account [2]. Some universities have a vision to become a national research university; therefore, they focus more on increasing the publications and improving their academic research performance in order to have a higher global ranking.

Workload, moreover, is another important problem since contribution from the academic service can account for only ten percent of their work so researchers in universities become less motivated to work with the industries.

Based on the Talent Mobility Programme evaluation, it was found that the project has been partially successful; not much attention and response received from researchers in national research universities. Most of the participating universities were small universities while the numbers of national research university, with high potential researchers, were limited; which is the research problem of this study.

This research, therefore, will explore why national research universities were less involved in the programme, or even joining but in different channels. This study mainly analyzes crucial factors that affect the participation of the efficient and sustained sharing of scientific, technological, and innovative personnel between universities and industries focusing on the case of National Research University (NRU) and food technology industry.

**III. METHODOLOGY**

Exploratory research was conducted to analyze data. This will lead to clarify the research questions for further empirical research. Semi-structured interviews were used to provide the information for the conceptual framework analysis [13].

This study emphasized on food science based academic researchers. Nine academic food science researchers, who have been successful in their research area, from national research universities were selected from the list. Six of them accepted the in-depth interviews.

Three food science researchers/professors came from Chulalongkorn University; two from Kasetsart University; and one from King Mongkut's Institute of Technology, Thonburi (KMUTT). Although some researchers were in the same faculty/department, they were experts in different areas of research. The details are shown in Table III.

The sampling design of this study, therefore, was from national research university and food industry dividing into 2 groups; participant and non-participant in Talent Mobility Programme of The National Science Technology and Innovation Policy Office (STI).

TABLE III. LIST OF INTERVIEWEES

Researcher	Faculty	Department	Expertise and Research Area
Prof. A	Science	Food Technology	Food Processing and Engineering
Prof. B	Science	Food Technology	Food Packaging
Prof. C	Science	Food Technology	Food Microbiology / Biotechnology
Prof. D	Agro-Industry	Science & Food Tech	Protein Chemistry
Prof. E	Agro-Industry	Product Development	Antimicrobial Packaging
Prof. F	Engineering	Food Engineering	Food Technology

#### IV. RESULTS AND DISCUSSIONS

##### A. Individual Motivations of Researcher

Most interviewees did agree that the individual motivations of researcher influence their decision to participate in the Talent Mobility Programme.

Three of academic professors insisted that the major reason for collaboration with the industry was to be able to accompany their students with them. Otherwise, they would not be interested in joining the TM project. Professor A said *“Joining Talent Mobility Programme is a good opportunity to bring our students to see the real world and to increase their knowledge, skill, and experience in industry environment. It is better than just learning in the classroom.”*

Three of six interviewees said that their motivations were from the right and interesting research topics agreed with industries. Other factors of industry side are also important such as company’s reputation, structure, and vision of the management team.

The results show that individual motivations of researcher are influential factors for their decision making to engage in the Talent Mobility Programme. Increasing the participation rate by inspiring and communicating the values of their motivations will raise the researchers’ interest in participating in the project.

##### B. Personnel Sharing: Linkage between University and Industry

From six official forms of personnel sharing between universities and industries in Thailand; Consultancy and Reach-Out, Licensing, Contract Research, Collaborative Research, Talent Mobility, and Academic Entrepreneurship, interviewees in this research preferred the Consultancy and Reach-Out via academic service center of their universities. The most popular practice, however, was using researcher’s personal connections/networks. This was due to the complication of university’s policy and process.

All interviewees did agree that university’s organizational and process factors are difficulties especially the permission and disbursement. Three professors, therefore, were not interested in Talent Mobility due to these

issues while the rest of interviewees, three researchers, still decided to join TM but via different channels.

Most professors commented that there should be effective intermediary persons in relevant agencies, STI office, OHEC, and university to support all processes and documentary issue. Professor D gave an opinion that *“This will much help, facilitate, and attract researchers to join the TM programme especially new participants.”*

##### C. Forms of Personnel Sharing

The interview results reveal that actually many researchers are interested in working with industries, for some specific reasons. Some of them, however, participated in the TM programme while some did not because of barriers described above. Forms of personnel sharing, moreover, are different.

The most common form of personnel sharing, regardless of participation of the TM programme, is through researcher’s personal connections/networks. This was mainly because of trustworthiness and good relationship of each other. Mindset and understanding are also factors affecting forms of personnel sharing. The details of personnel sharing forms are described in Table IV.

TABLE IV. FORMS OF PERSONNEL SHARING

Researcher	Participation in Talent Mobility Programme	Form of Personnel Sharing
Prof. A	No	Via Personal Connection/Network
Prof. B	No	Via Personal Connection/Network
Prof. C	No	Via Personal Connection/Network
Prof. D	Yes	Talent Mobility via Academic Service Center
Prof. E	Yes	Talent Mobility via Academic Service Center
Prof. F	Yes	Direct to Talent Mobility Programme

##### D. Main Supporting Factors to Successful Personnel Sharing

In this study of personnel sharing between national research university and food technology industry, the main supporting factors are:

1) Researcher’s personal characteristics and motivations: These were the most important factors effecting the successful programme. Professor D said

*“Researcher’s personal interest is the motivation to working with industries.”*

All interviewees insisted that their influential motivation of working with companies is to gain experience and networking.

2) Industry’s research topic: The second success factor was that the research topic must be useful and having high impact as Professor B expressed that

*“A useful research topic is not beneficial only to the industry and researcher but also to our students to have a great experience on it which we could not find in the classroom or a lab.”*

3) Relationship between researchers and industries: Professor A did give an important factor on networking with industries by saying that

*“We, industry and researcher, must have the right and common research topic. We, moreover, should be open-minded and keep trustworthiness, fairness, and good mindset to each other.”*

#### E. Barriers to Personnel Sharing

In this study, barriers to personnel sharing between national research university and food technology industry are:

1) University’s policy, rules, and regulations: All six interviewees did concern about this issue including internal conflicts and problems as Professor C illustrated that

*“Researchers themselves should keep a strong motivation to do the advanced research with industries regardless of concerning the internal conflicts. My case was a good example of having good relationship with a big food firm and could bring the budget to build a lab at our department. This finally benefits our university and students.”*

Professor C also described

*“Researchers are good materials as input. We just need the appropriate process for effective output or results of the programme.”*

The recommendation for policy implementation is that university has to respond and to deploy the policy by having appropriate policy application, and also to build the supporting ecosystem. In some cases, conditions might be adjusted and regulations would be customized properly and flexibly. There should be, moreover, a strong management team in order to deploy the policy effectively.

2) Inappropriate university’s workload, KPI, and evaluation system: All professors concerned about the workload issue. Professor E said

*“There is not only teaching but also doing research and student affairs including extra activities. We could implement the TM programme effectively only if our workload is based on reasonable assignments. Promotion criteria, moreover, still focus on publications which can be easier evaluated than joining extra projects. These all reflect less successful results.”*

Solutions are having reasonable university’s workloads, identifying clear objectives, and aligning KPIs to the evaluation system.

3) Lack of facilitator/intermediary person between university and researcher: As known that all TM processes, such as application and permission, including documentary issue are taking time. There should be a facilitator to assist on these as Professor F commented shortly that

*“Middle unit is crucial to accelerate the success of the Talent Mobility Programme.”*

Suggestion is providing support unit/system like international practice as Professor F gave an example of initiating the Knowledge Exchange for Innovation Center (KX) in three campuses of his university in Thailand.

#### F. Additional Findings

Individual characteristics of academic researchers or professors in general are arrogant and not proactive. Professor E expressed regarding researchers’ characteristic of self-esteem or ego that

*“Many professors have a big ego and cannot feel lose face or discreditable when being unsuccessful. Some of them are good at academic but not in practice; this is the reason why they do not want to take a risk of going out with industries.”*

Professor C also straightforwardly said

*“Most professors addict to and are in comfort zones where they can arrange their time flexibly. They do not like to be controlled or being demanded if working with companies. These are basic characteristics of Thai professors. The university or even management, however, does not have a strong intention or support policy to encourage their professors to work with industries so everyone can ignore without any penalty; they are still secure. Personal motivations of researchers are also different and specific individually; this may make the TM programme difficult to be successful.”*

Two of six interviewees, did agree that TM participants should not be young researchers since the seniors could deliver more effective performance to the industry than juniors. Professor F said

*“Experienced researchers also have their research team including assistants and students to work for companies.”*

Most researchers, however, encouraged new professors to attend the programme as Professor A described that

*“Joining Talent Mobility Programme can increase skill, expertise, and experience for researchers. Knowledge, moreover, from real experience in the industry environment is better than teaching in the classroom. This is good for new/young professors.”*

Finally, all professors gave weight on researcher factors as the most important factors effecting Thailand Talent Mobility Programme. The followings are industry and university factors respectively. Professor B said

*“Industry must be well-established and having good system. Importantly, there also is a strong vision from the management team.”*

Professor D gave an opinion regarding university factors that

*“University is just a support agency for the process and administration.”*

## V. CONCLUSIONS AND RECOMMENDATIONS

This exploratory study explored factors effecting successful Talent Mobility Programme, including other forms of personnel sharing, and barriers that they face. It has focused on the case of national research university and food technology industry in Thailand through in-depth interviews with six researchers from three universities. The results of this study disclosed five principle issues. First, individual motivations of researcher are influential factors towards the success of TM programme. Second, linkage between university and industry includes six official forms of personnel sharing; Consultancy and Reach-Out, Licensing, Contract Research, Collaborative Research, Talent Mobility, and Academic Entrepreneurship. Third, the most common form of personnel sharing, regardless of participation of the TM programme, is through researcher's personal connections/networks. Fourth, three successful factors effecting personnel sharing are researcher's personal characteristics and motivations; industry's research topic; and relationship between researchers and industries. Lastly, three barriers were identified causing difficulties for personnel sharing are university's policy, rules, and regulations; inappropriate university's workload, KPI, and evaluation system; and lack of facilitator/intermediary person between university and researcher.

To improve the efficiency and effectiveness of Talent Mobility Programme, it is important to understand the factors of success, the nature of the barriers to successful personnel sharing, and ways to overcome the barriers. The recommendations are building a strong collaboration and linkage between university and industry, and avoiding bringing the TM concept from other countries without considering the different contexts of Thailand that workload is enormously different. The TM preparation training courses are crucial both in business aspect and mindset. Sharing experience from former TM participants is also important.

This study only discussed on factors effecting Thailand Talent Mobility Programme in the case of national research university and food technology industry. Future research, therefore, should be studied in other areas of industry or other university characteristics in order to compare key success factors or barriers to personnel sharing mechanism that would reveal different results.

**Acknowledgement**—Prof. Emeritus Achara Chandrachai, Ph.D., the Professor Emeritus of Technopreneurship and Innovation Management Program, Graduate School, Chulalongkorn University; who gave valuable advice and great recommendations on interview topics and questionnaire for this research study.

## REFERENCES

- [1] Abdullah Alshehri, Saud A. Gutub, Mostafa A.-B. Ebrahim, Hani Shafeek, Mohamed F. Soliman, and M.H. Abdel-Aziz, "Integration between industry and university: Case study, Faculty of Engineering at Rabigh, Saudi Arabia". 2016.
- [2] Arum Kitipongwatana, Kittisak Kaweevijmanee, Oraphan Wiarachai and Poolsak Koseeyaporn, "An Empirical Study of Policy Implementation of Thailand Talent Mobility Programme". 2016. Presented at the 13<sup>th</sup> International Conference ASIALICS 2016, Bangkok, Thailand.
- [3] Augustine ObeleaguAgu, and Benedict O. Emunemu, "Entrepreneurial University Paradigm for Sustainable Social and Economic Development in Nigeria". 2011.
- [4] BRIMBLE, P., and DONER, R.F., "University-Industry Linkages and Economic Development: The Case of Thailand". World Development, vol. 35(6), 2007 pp. 1021-1036.
- [5] Edward B. Acworth, "University-industry engagement: The formation of the Knowledge Integration Community (KIC) model at the Cambridge-MIT Institute". 2008.
- [6] Etzkowitz, H., and Leydesdorff, L., "The dynamics of innovation: from National Systems and "Mode2" to a Triple Helix of university-industry-government relations". Research Policy 29, 2000, 109-123.
- [7] Etzkowitz, H., "The Second Academic Revolution and the Rise of Entrepreneurial Science". IEEE Technology and Society Magazine summer 2001.
- [8] Geisler, E., Furino, A., and Kiresuk, T.J., "Factors in the Success or Failure of Industry-University Cooperative Research Centers". Interfaces 20(6), 1990, pp. 99 - 109.
- [9] Graham S. King, and Cary R. Cameron, "An Enhanced Model for University-Industry Collaboration for Innovation in Trinidad and Tobago". 2013.
- [10] Intarakumnerd, P., and Schiller, D., "University-Industry Linkages in Thailand: Successes, Failures and Lessons Learned for other Developing Countries". The IV Globelics Conference Mexico City, 2008.
- [11] Inzelt, A., "The evolution of university-industry-government relationships during transition". Research Policy 33, 2004, pp. 975-995.
- [12] Isabel M. Bodas Freitas, Aldo Geuna, and Federica Rossi, "The governance of formal university-industry interactions: understanding the rationales for alternative models". 2012.
- [13] Kanyakit Keerati-angkoon, Rath Pichyangkura, and Achara Chandrachai, "Science Based Research Commercialization from Universities in Thailand: The Perspective of Successful Academic Researchers". IEEE Computer Society, 2011, pp. 508-513.
- [14] MikeWright, Bart Clarysse, Andy Lockett, and Mirjam Knockaert, "Mid-range universities' linkages with industry: Knowledge types and the role of intermediaries". 2008.
- [15] Mohammad S Khorsheed, and Mohammad A Al-Fawzan, "Fostering university-industry collaboration in Saudi Arabia through technology innovation centers". 2014.
- [16] Naparat Siripitakchai and Kumiko, "University-Industry Linkages (ULIs) and Research Collaborations: Case of Thailand's National Research Universities (NRUs)". Industrial Engineering, Management Science and Applications, 2015. pp.189.
- [17] National Science and Technology Development Agency, *Survey Report: Technical and Engineering Problems and R&D Needs in Private Sector*. 2015.
- [18] National Science Technology and Innovation Policy Office, *Research Development and Innovation Survey 2015*. 2015.
- [19] National Science Technology and Innovation Policy Office, *Thailand Science and Technology Indicators 2014*. 2015.
- [20] Timothy M. Thamae, Retselisitsoe I. Thamae, and Leboli Z. Thamae, "A Process Model for University-Industry Cooperation in Sub-Saharan Africa: Lessons from Lesotho". 2016.
- [21] Wong, P. K., Ho, Y.-P., and Singh, A., "Towards an entrepreneurial university model to support knowledge-based economic development: The case of the National University of Singapore". World Development 35, 2007, pp. 941-958.
- [22] World Economic Forum, *Talent Mobility Good Practices - Collaboration at the Core of Driving Economic Growth*. 2012.

# shinySDM: Point and Click Species Distribution Modeling

Thomas Nash

Department of Computer Science  
College of Charleston  
Charleston, SC 29401

Aspen Olmsted

Department of Computer Science  
College of Charleston  
Charleston, SC 29401

**Abstract**— The focus of this research work is to address the difficulties involved in creating visualizations for species distribution modeling. We focus on two aspects of this problem: running models for predicting the likelihood of outbreak locations and testing the significance of the models generated. To improve this process, this work develops a web application which allows researchers to upload their data, create informative and interactive visualizations, and run desired models in addition to testing their significance. Such an application empowers researchers without any programming experience to both generate complex models and interpret results quickly and effectively. This paper will focus on maximum entropy modeling as the example modeling technique by providing an example run using data on vaccine-preventable diseases.

**Keywords**- data visualization; web application; R; shiny; maxent

## I. INTRODUCTION

Modeling of disease outbreaks plays an integral role in public health as it can assist in identifying at-risk areas for both the spread of current epidemics and future outbreaks. The data contains geographic coordinates of where known outbreaks have occurred, but there often may be supplementary data providing details of each outbreak. Using species distribution modeling (SDM), researchers can generate likelihood distributions of where potential outbreaks may occur by studying the attributes of the locations where outbreaks are known to have occurred. This analysis can be done using a variety of software, but one of the most popular is the open-source statistical computing language R due to the availability of complex modeling functions. Current methods to create such models can be cumbersome and difficult to learn for those not well-versed in programming, though. The creation of a web application that can run these models and facilitate data exploration will present an opportunity to explore and visualize data to a broader audience. The organization of the paper is as follows. Section 2 describes the related work and the implementations of current methods. In Section 3 we give a motivating example where our application is useful. Section 4 details the underlying framework as well as the advantages of the design and the feature set provided compared to a user running the same processes through an R shell. Section 5 provides a walkthrough of the

application using an example dataset. We conclude and discuss future work in Section 6.

## II. RELATED WORK

Philips et al. [1] introduced the maxent model which is the only supported modeling technique for the application, currently. Maxent is one of many SDM algorithms which can be adapted for disease outbreak modeling but is ranked as one of the most popular. The algorithm works to model geographic distributions using presence only data combined with environmental measures of the presence sites. Provided in the application are 19 Bioclim variables used to quantify features of these locations. Hijmans et al. [2] author the dismo package in R which contains a maxent implementation in addition to that of other SDM algorithms. This package is used in our application to run the model on the provided data using the provided features. Further work will include using residual values of the generated distribution to test model fit. Chang et al [3] developed the Shiny web application framework for R. This framework allows developers to integrate R code with a user interface that users can run the inside the RStudio IDE for R or in their browser. This allows for remote access on a web server where the application is hosted or local instances. The added benefit is that users can download the application and extend or modify the existing code with ease.

## III. MOTIVATING EXAMPLE

Our contribution in this work is to develop an application which allows users to perform SDM with no programming experience. Our hypothesis was that we could develop an application which allowed for data exploration, data modeling and significance testing. We demonstrate the application capabilities using data provided by the Council on Foreign Relations' (CFR) data on vaccine-preventable diseases [4]. This data is comprised of hundreds of reports of disease outbreaks with features ranging from geographic coordinates to size of the outbreak. The CFR provides an interactive visualization for the data, but it is not easy to navigate nor does it provide capability for further data processing.

IV. MODELING APPLICATION

To achieve the goal of creating a robust and user-friendly tool for complex data analysis and exploration, we developed our application using the Shiny framework provided in R. The GUI contains a main panel for data upload and visualization which is done by overlaying the data points and generated model plots on a world map. In addition, ROC and variable significance figures are presented after models are run and selected. Further functionality exists through reverse geocoding which takes the input geographic coordinates and querying the Google Maps API to retrieve information regarding city, state and/or country. A second tab presents the input data in tabular format which allows for sorting and searching, but also the ability to jump to data points on the map. Currently, only maxent modeling is supported, but the Shiny framework allows for the expansion of new modeling methods by adding the appropriate code to the server function.

For a comparison of the application to current methods, we identify several significant attributes a powerful visualization should contain. The two baselines are the map presented in [5] of the outbreak data and the standard plot generated in R by plotting the maxent model with the 'plot()' function. While the CFR plot created using Tableau provides interactivity, it lacks integration for any further analysis. The standard plot output in R can be more informative of the model itself, but lacks further insight into the data and can require intensive coding expertise. The Shiny application shines in the ability to bring more information to the forefront, compare multiple models at once and reach a larger audience. These factors are what makes it a strong tool for researchers in the field.

TABLE I. COMPARING FUNCTIONALITY OF DIFFERENT METHODS

Feature	CFR map	Standard plots	shinySDM
Map zoom	X		X
Modeling capability, incl. significance, feature selection		X	X
Data exploration	X		X
Raw data access	X	X	X
Reverse geocoding			X
Data independent		X	X
Customizable view		X	X
No coding required	X		X

V. WORKED EXAMPLE

Use of the application aims to be as simple and intuitive as possible. Users upload data, in the form of tab-separated or CSV latitude and longitude points, using the labeled button. From there, one can zoom and pan the map which contains overlaid data points log scaled to the size of the outbreak. If the user chooses to provide additional information about each

data point, e.g. outbreak description, this information is displayed when a data point is clicked. Further exploration is provided in the "Data explorer" tab where data is presented in a searchable table and points on the map can be accessed with a click. After data is uploaded, all models are run with the button press to support ensemble models. For this example, all predictors are chosen. The selected model map is then combined with the world map to create the visualization shown in Fig. 1. The figure presents data uploaded from [5] and the maxent model distribution overlaid. Additional plots indicating the chosen model's accuracy and predictor significance are presented on the left. Further details are provided in the documentation found on the package website <https://github.com/TomNash/shinySDM>.



Figure 1. Application user interface with data and model overlay

VI. CONCLUSIONS AND FUTURE WORK

In this work, we provide a web application capable of running and displaying SDMs through interactive visualizations. We show that it is possible to run such models using the Shiny framework and users with no coding experience can utilize the R libraries and packages necessary to create informative plots. The Shiny application contains features beyond those of standard plots generated through R functions, thus allowing for more in-depth understanding of both the data and the associated models. Future work will include the addition of more modeling techniques, support for ensemble methods and a more refined measure of model accuracy.

REFERENCES

- [1] S. J. Phillips *et al.*, "Maximum entropy modeling of species geographic distributions," *Ecological Modelling*, vol. 190, no. 3-4, pp. 231-259, 2006. R. J.
- [2] Hijmans *et al.*, *dismo: Species Distribution Modeling*, R package version 1.1-1, 2016.
- [3] A. M. Samy *et al.*, "Mapping the global geographic potential of Zika virus spread," *Memórias do Instituto Oswaldo Cruz*, vol. 111, no. 9, pp. 559-560, 2016.
- [4] Council on Foreign Relations, "Map: Vaccine-Preventable Outbreaks," Council on Foreign Relations, 2014.[Online].Available:[http://www.cfr.org/interactives/GH\\_Vaccine\\_Map/index.html#map](http://www.cfr.org/interactives/GH_Vaccine_Map/index.html#map). [Accessed: 03-Mar-2017].

# Efficient Retrieval of Information From Hierarchical REST Requests

Seth Stoudenmier, Aspen Olmsted  
Computer Science Department  
College of Charleston  
Charleston, SC, USA  
stoudenmiersh@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—REST requests are utilized by developers across many fields to access public data that they would not normally have access to. Typically, efficiency is not a concern when making a REST request to a server since only a handful are performed at once. With the increasing popularity of gaming, more and more web services have emerged that aim to present the data taken from these APIs. API Keys and other methods of authorization used are associated with an account that has a maximum number of requests that are allowed per a specific unit of time. The problem is that these databases are relational and result in multiple REST requests to traverse the hierarchical structure and retrieve the needed information, leading to a long runtime before the application is loaded. As a result, many of these web services have downtime at specific intervals, such as the release of a new patch, so that all necessary information can be retrieved, stored on a local server, and then presented to users. This paper presents a solution to increase the efficiency of retrieving data from a REST API such that a web service that presents this information may minimize its downtime.

**Keywords**—REST; API; API keys; big data; League of Legends

## I. INTRODUCTION

The gaming industry is continually growing as time progresses, resulting in a rapid increase of web services aimed at improving the experience and community surrounding these games. To implement these web services, developers must make use of APIs made available by different companies throughout the industry. In some cases, only a few of the REST requests are made at a time to the API which causes no conflict. On the other hand, sometimes a large amount of data is desired, and this results in many subsequent REST requests that would surpass the maximum allowed in a specified unit of time. The solution to this problem is to pull back all data that would be needed for a specific web service and store it on a local server. This still leaves one concern, which is the amount of time that it takes to pull back copious quantities of data. A proposed solution to this problem is the use of multiple API keys on a rotating basis allowing for the downtime between calls to be minimal.

Of course, this solution does have its restraints and limitations. In this paper, we will talk about a simple implementation that uses multiple API keys to retrieve information from a hierarchical database. Additionally, we will

discuss some future work for this project that will result in an increase in efficiency.

## II. RELATED WORK

Background research on the topic of efficient REST requests did not result in any articles that try to solve a similar problem as what this paper aims to do. However, many papers were found that were used and will be used to improve the presented solution.

Heiland et al. [1] talk about the process of deciding what type of authorization the API for their application, Science Gateway Platform, should utilize. In doing so, they discuss briefly some of the other applications that utilize API keys and other authorization techniques. While API keys are not the most secure method of authentication, they are the most user-friendly and are used widely across multiple applications. The solution presented in this paper focuses primarily on API keys since they are so widely used, but it can be adapted to other forms of authorization.

Li and Halfond [2] discuss HTTP request in relation to power consumption on Android. They state that network usage is to blame for the majority of battery usage since HTTP requests have multiple steps within them, resulting in a large amount of processing per request. Their solution to this problem involves the capturing of sequential requests so that they may be sent as a single request to a proxy that handles the rest. Li and Halfond's paper results in the realization that for enormous quantities of data, these HTTP requests would be run on a local server instead of client side, resulting in the scope of the problem for this paper being altered to address the exponential growth of time that is needed for hierarchical databases.

Schreier [3] addresses the issue of how to model RESTful applications. She talks of how the behavioral model "offers the possibility to describe the behavior with state machines." A similar idea can be used in the mapping of REST requests of an API to create a hierarchy. The run time to traverse all branches can be estimated if a rough idea of the number of nodes at each level is known. From here, it can be easily seen how utilizing multiple API keys can reduce the runtime by half, minus the necessary overhead.

### III. MOTIVATING EXAMPLE

The motivation for this idea comes from the gaming industry. The industry is growing rapidly and the number of developers with it. Many of these developers utilize some sort of an API that companies in the industry make public. From these APIs, many different web services have been developed to present statistics, raw information, and much more to the community that shares a similar interest. The game that was used for proof of concept in this paper is League of Legends by Riot Games [4]. They have an outward-facing API that provides a multitude of information; however, like any other API, there is a maximum to the number of REST requests that may be made per some unit of time. Some web services that utilize this API are required to have down time that surpasses twenty-four hours so that they can pull all necessary data to a local server when a new patch is released. The solution proposed in this paper hopes to minimize this downtime for the previously mentioned web services and for developers that would like to test different combinations of data.

### IV. IMPLEMENTATION

The implementation used for the paper is a simple proof of concept. Fig. 1 shows the top portion of a function that is used to handle the pulling and relating to data from the API. The key component of the solution can be found in the segments “interval = ((maxTime \* 60.0) / maxRequests) / keyCount” and “[requestCount % keyCount].” First is the interval between REST requests. This is determined by the number of API keys and the request limitations that an API has. Between each request, the sleep function is called so that the threshold is not reached, which would result in an API key being blacklisted and no longer able to be used. Next is the index into the list that contains the API keys. A count of the requests made is maintained so that, through modular division, the solution can cycle over the keys one at a time. Paired with the interval, this maximizes the number of requests that are allowed per key. A summary of the hierarchy that the data was retrieved from can be found in Fig. 2. The result of each run was a JSON pairing each Account ID to its corresponding Match IDs and Match Info.

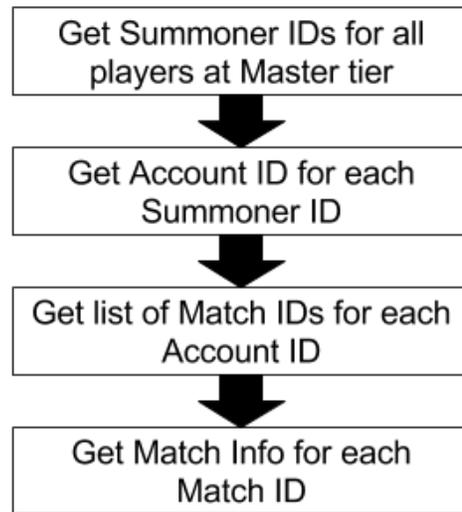


Figure 2. Example hierarchy of REST requests

### V. RESULTS

The results for the solution were as predicted. The addition of a second API key decreased the amount of time necessary to retrieve the same information by 33%. When used in a small sample size, such as the data presented in Fig. 3, the difference is only a matter of minutes. However, this difference becomes significant as the number of requests increases. For the testing of the implementation, a range of 1 to 20 players’ data was pulled out of hundreds of players, possibly thousands depending on the dataset.

### VI. FUTURE WORK

To improve on this proof of concept, two key areas will be considered: security and modularity. Security is a major concern for this solution since it handles API keys that are unique to each user. For this solution to be used, multiple people would have to be willing to offer their unique key. A possible solution to this would be to implement public-key cryptography. This would allow people in these gaming communities to use their favorite web services with greater safety.

```

def fetchData(keyList, maxRequests, maxTime):
    # variable that keeps track of the total request count
    requestCount = 0
    # beginning of function execution
    startTime = time.time()
    # number of apikeys
    keyCount = len(keyList)
    # calculate the amount of time needed between each REST call
    interval = ((maxTime * 60.0) / maxRequests) / keyCount
    print(interval)
    # list containing RiotAPI objects
    apiList = [RiotAPI(key) for key in keyList]
    # get the summoner id list for master tier
    summonerIdList = apiList[requestCount % keyCount].getSummonerListFromMasterTier()
    requestCount += 1
    time.sleep(interval)
    # get the account id for each summoner id
    accountIdList = []
    count = 0
    
```

Figure 1. Small portion of implementation for proof of concept

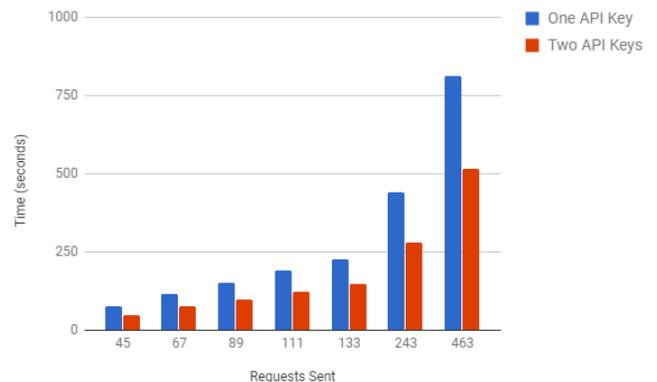


Figure 3. Results showing number of requests against the time in seconds

Next is the goal to make this solution more modular since, currently, it is nothing more than a theory that has been proven. Implementing a way for developers to utilize this with any API seamlessly would allow for efficient data retrieval and collaboration between communities. Of course, this concept of shared keys can be applied to any web service with an API that requires a substantial number of REST requests to retrieve all necessary data.

#### REFERENCES

- [1] R. Heiland et al. "Authentication and Authorization Considerations for a Multi-tenant Service," in Proc. SCREAM, 2015, pp. 29-35.
- [2] D. Li and W. G. J. Halfond. "Optimizing Energy of HTTP Requests in Android Applications," in Proc. DeMobile, 2015, pp. 25-28.
- [3] S. Schreier. "Modeling RESTful applications," in Proc. WS-REST, 2011, pp. 15-21.
- [4] *League of Legends*. [PC]. United States: Riot Games, 2017.

## **Session 20: Information Security and Intelligent Control**

Title: Multiple Assignment Secret Sharing Scheme Using Hierarchical Threshold Scheme  
(Author: Kouya Tochikubo)

Title: Bot or Not  
(Authors: Husna Siddiqui, Elizabeth Healy, Aspen Olmsted)

Title: Intelligent Laboratory Management System Based on Internet of Things  
(Authors: Yichen Ma, Fuyao Wang, Zhuozheng Wang)

# Multiple Assignment Secret Sharing Scheme Using Hierarchical Threshold Scheme

Kouya Tochikubo

Department of Mathematical Information Engineering,  
College of Industrial Technology, Nihon University,  
Chiba, Japan  
tochikubo.kouya@nihon-u.ac.jp

**Abstract**—We propose a new secret sharing scheme realizing general access structures, which is based on unauthorized subsets. In the proposed scheme, shares are generated by Tassa's  $(k, n)$ -hierarchical threshold scheme. The proposed scheme can reduce the number of shares distributed to each participant.

**Keywords**— secret sharing scheme; general access structure;  $(k, n)$ -hierarchical threshold scheme; key management

## I. INTRODUCTION

In 1979, Blakley and Shamir independently introduced the concept of secret sharing [1], [2]. In Shamir's  $(k, n)$ -threshold scheme [1], every group of  $k$  participants can recover the secret  $K$ , but no group of less than  $k$  participants can get any information about the secret from their shares. The collection of all authorized subsets of participants is called the access structure. A  $(k, n)$ -threshold scheme can only realize particular access structures that contain all subsets of  $k$  or more participants.

Secret sharing schemes realizing more general access structures than that of a threshold scheme were studied by numerous authors. Koyama proposed secret sharing schemes for multi-groups [3]. Simmons studied secret sharing schemes realizing multilevel access structures [4], [5]. Subsequently, Tassa proposed a hierarchical threshold scheme using polynomial derivatives [6]. Farràs and Padró formalized the concept of hierarchical access structure [7]. Secret sharing schemes based on graph access structures were also proposed [8]–[10]. These schemes obtain the optimal information rates for some access structures, but these schemes cannot be applied to many access structures.

On the other hand, Ito, Saito and Nishizeki proposed a multiple assignment secret sharing scheme for general access structures and showed an explicit share assignment algorithm for any access structure [11]. Their scheme can realize an arbitrary access structure by assigning one or more shares to each participant. Benaloh and Leichter proposed a secret sharing scheme for general access structures based on a monotone-circuit [12]. Secret sharing schemes which have an explicit assignment algorithm for any access structure are categorized by two types. One type is schemes based on unauthorized subsets [11], [13], [14]. Another type is schemes based on authorized subsets [12], [15], [16].

In the implementation of secret sharing schemes for general access structures, an important issue is the number of shares distributed to each participant. Obviously, a scheme constructed of small shares is desirable. However, in general, the proposed secret sharing schemes for general access structures are impractical in this respect when the size of the access structure is very large.

In this paper, we modify the scheme I of T04 [13] and propose a new secret sharing scheme realizing general access structures, which is based on unauthorized subsets. In the proposed scheme, shares are generated by Tassa's  $(k, n)$ -hierarchical threshold scheme instead of Shamir's  $(k, n)$ -threshold scheme. Thus, the proposed scheme can reduce the number of shares distributed to each participant. Furthermore, we show that the proposed scheme is more efficient than or equal to the scheme I of T04 [13] from the viewpoint of the number of shares distributed to each participant.

## II. PRELIMINARIES

### A. Secret Sharing Scheme

Let  $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$  be a set of  $n$  participants. Let  $D(\notin \mathcal{P})$  denote a dealer who selects a secret and distributes a share to each participant. Let  $\mathcal{K}$  and  $\mathcal{S}$  denote a secret set and a share set, respectively. For sets  $A$  and  $B$ , we denote a difference set by  $A - B$ . The access structure  $\Gamma(\subset 2^{\mathcal{P}})$  is the family of subsets of  $\mathcal{P}$  which contains the sets of participants qualified to recover the secret. For any authorized subset  $A \in \Gamma$ , any superset of  $A$  is also an authorized subset. Hence, the access structure should satisfy the monotone property:

$$A \in \Gamma, A \subset A' \subset \mathcal{P} \Rightarrow A' \in \Gamma.$$

Let  $\Gamma_0$  be a family of the minimal sets in  $\Gamma$ , called the minimal access structure.  $\Gamma_0$  is denoted by

$$\Gamma_0 = \{A \in \Gamma : A' \not\subset A \text{ for all } A' \in \Gamma - \{A\}\}.$$

For any access structure  $\Gamma$ , there is a family of sets  $\bar{\Gamma} = 2^{\mathcal{P}} - \Gamma$ .  $\bar{\Gamma}$  contains the sets of participants unqualified to recover the secret. The family of maximal sets in  $\bar{\Gamma}$  is denoted by  $\bar{\Gamma}_1$ . That is,

$$\bar{\Gamma}_1 = \{B \in \bar{\Gamma} : B \not\subset B' \text{ for all } B' \in \bar{\Gamma} - \{B\}\}.$$

Let  $p_{\mathcal{K}}$  be a probability distribution on  $\mathcal{K}$ . Let  $p_{\mathcal{S}(A)}$  be a probability distribution on the shares  $\mathcal{S}(A)$  given to a subset

$A \subset \mathcal{P}$ . Usually a secret  $K$  is chosen from  $\mathcal{K}$  with the uniform distribution. A secret sharing scheme is perfect if

$$H(K|A) = \begin{cases} 0 & (\text{if } A \in \Gamma) \\ H(K) & (\text{if } A \notin \Gamma), \end{cases}$$

where  $H(K)$  and  $H(K|A)$  denote the entropy of  $p_{\mathcal{K}}$  and the conditional entropy defined by the joint probability distribution  $p_{\mathcal{K} \times \mathcal{S}(A)}$ , respectively.

In general, the efficiency of a perfect secret sharing scheme is measured by the information rate  $\rho$  [17] defined as

$$\rho = \min\{\rho_i : 1 \leq i \leq n\},$$

$$\rho_i = \frac{\log |\mathcal{K}|}{\log |\mathcal{S}(P_i)|}$$

where  $\mathcal{S}(P_i)$  denotes the set of possible shares that  $P_i$  might receive. Obviously, a high information rate is desirable. A perfect secret sharing scheme is ideal if  $\rho = 1$ . Throughout the paper,  $p$  is a large prime, and let  $Z_p$  be a finite field with  $p$  elements. In this paper, we assume  $\mathcal{K} = \mathcal{S} = Z_p$ .

### B. Shamir's Threshold Scheme

Shamir's  $(k, n)$ -threshold scheme is described as follows [1]:

- 1) A dealer  $D$  chooses  $n$  distinct nonzero elements of  $Z_p$ , denoted by  $x_1, x_2, \dots, x_n$ . The values  $x_i$  are public.
- 2) Suppose  $D$  wants to share a secret  $K \in Z_p$ ,  $D$  chooses  $k - 1$  elements  $a_1, a_2, \dots, a_{k-1}$  from  $Z_p$  independently with the uniform distribution.
- 3)  $D$  distributes the share  $s_i = f(x_i)$  to  $P_i$  ( $1 \leq i \leq n$ ), where

$$f(x) = K + a_1x + a_2x^2 + \dots + a_{k-1}x^{k-1}$$

is a polynomial over  $Z_p$ .

It is known that Shamir's  $(k, n)$ -threshold scheme is perfect and ideal [17], [18]. This implies that every  $k$  participants can recover the secret  $K$ , but no group of less than  $k$  participants can get any information about the secret.

The access structure of  $(k, n)$ -threshold scheme is described as follows:

$$\Gamma = \{A \in 2^{\mathcal{P}} : |A| \geq k\}.$$

### C. Tassa's Hierarchical Threshold Scheme

Let  $\mathcal{P}$  be a set of  $n$  participants and assume that  $\mathcal{P}$  is divided into  $m + 1$  disjoint subsets  $\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_m$ , i.e.

$$\mathcal{P} = \bigcup_{i=0}^m \mathcal{U}_i \text{ and } \mathcal{U}_i \cap \mathcal{U}_j = \emptyset \text{ for all } 0 \leq i < j \leq m.$$

Let  $\mathbf{k} = \{k_i\}_{i=0}^m$  be a monotonically increasing sequence of integers  $0 < k_0 < \dots < k_m$ . We set  $k = k_m$ . Tassa's  $(\mathbf{k}, n)$ -hierarchical threshold scheme is described as follows [6]:

- 1) Suppose A dealer  $D$  wants to share a secret  $K \in Z_p$ ,  $D$  chooses  $k - 1$  elements  $a_1, a_2, \dots, a_{k-1}$  from  $Z_p$  independently with the uniform distribution and defines a polynomial over  $Z_p$

$$f(x) = K + a_1x + a_2x^2 + \dots + a_{k-1}x^{k-1}.$$

- 2)  $D$  identifies each participant  $P \in \mathcal{P}$  with a field element. For simplicity, the field element that corresponds to  $P_r \in \mathcal{P}$  will be also denoted by  $r$  ( $1 \leq r \leq n$ ).
- 3)  $D$  distributes the shares to all participants in the following manner: Each participant of  $i$  th level in the hierarchy  $P_r \in \mathcal{U}_i$  receives the share  $f^{(k_{i-1})}(r)$  where  $f^{(k_{i-1})}(r)$  denotes the  $(k_{i-1})$  th derivative of  $f(x)$  at  $x = r$  and  $k_{-1} = 0$ .

The access structure of Tassa's  $(\mathbf{k}, n)$ -hierarchical threshold scheme is described as follows:

$$\Gamma = \left\{ \mathcal{V} \subset \mathcal{P} : \left| \mathcal{V} \cap \left( \bigcup_{j=0}^i \mathcal{U}_j \right) \right| \geq k_i, \forall i \in \{0, 1, \dots, m\} \right\}.$$

It is known that Tassa's  $(\mathbf{k}, n)$ -hierarchical threshold scheme is perfect and ideal [6].

*Example 1:* Let  $\mathbf{k} = (k_0, k_1, k_2) = (1, 3, 4)$ ,  $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5, P_6\}$  and

$$\begin{aligned} \mathcal{U}_0 &= \{P_1\}, \\ \mathcal{U}_1 &= \{P_2, P_3, P_4\}, \\ \mathcal{U}_2 &= \{P_5, P_6\}. \end{aligned}$$

In this case, the access structure  $\Gamma$  and the minimal access structure  $\Gamma_0$  of Tassa's  $((1, 3, 4), 6)$ -hierarchical threshold scheme are described by

$$\Gamma = \left\{ \mathcal{V} \subset \mathcal{P} : \left| \mathcal{V} \cap \left( \bigcup_{j=0}^i \mathcal{U}_j \right) \right| \geq k_i, \forall i \in \{0, 1, 2\} \right\}$$

and

$$\Gamma_0 = \{ \{P_1, P_2, P_3, P_4\}, \{P_1, P_2, P_3, P_5\}, \{P_1, P_2, P_3, P_6\}, \{P_1, P_2, P_4, P_5\}, \{P_1, P_2, P_4, P_6\}, \{P_1, P_3, P_4, P_5\}, \{P_1, P_3, P_4, P_6\} \},$$

respectively. Here, we shall realize this access structure by Tassa's scheme.

- 1)  $D$  selects a random polynomial

$$f(x) = K + a_1x + a_2x^2 + a_3x^3.$$

- 2)  $D$  distributes the share  $s_1 = f(1)$  to  $P_1$ .
- 3)  $D$  distributes the share  $s_r = f'(r)$  to  $P_r$  ( $2 \leq r \leq 4$ ), where

$$f'(x) = a_1 + 2a_2x + 3a_3x^2.$$

- 4)  $D$  distributes the share  $s_r = f^{(3)}(r)$  to  $P_r$  ( $5 \leq r \leq 6$ ), where

$$f^{(3)}(x) = 6a_3.$$

D. Secret Sharing Schemes Based on Unauthorized Subsets

Ito, Saito and Nishizeki's multiple assignment secret sharing scheme realizes general access structures by assigning one or more shares to each participant. Shares are generated by a  $(k, n)$ -threshold scheme with  $k = n$ . For  $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$ ,  $K \in \mathcal{K}$  and  $\Gamma$ , Ito, Saito and Nishizeki's scheme [11] is described as follows.

*Ito, Saito and Nishizeki's scheme:*

- 1) Let  $\bar{\Gamma}_1 = \{B_1, B_2, \dots, B_t\}$ . Compute  $t (= |\bar{\Gamma}_1|)$  shares

$$S = \{w_1, w_2, \dots, w_t\}$$

for the secret  $K$  by using Shamir's  $(t, t)$ -threshold scheme.

- 2) Distribute shares to  $P_i \in \mathcal{P}$  ( $1 \leq i \leq n$ ) according to the function  $g_{ISN} : \mathcal{P} \rightarrow 2^S$  defined as

$$\begin{aligned} g_{ISN}(P_i) &= \{w_j : P_i \notin B_j \in \bar{\Gamma}_1, 1 \leq j \leq t\} \\ &= \bigcup_{\substack{1 \leq j \leq t \\ P_i \notin B_j}} \{w_j\}. \end{aligned} \quad (1)$$

*Example 2:* For  $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5, P_6\}$ , we consider the following access structure

$$\begin{aligned} \Gamma_0 = & \{\{P_1, P_2, P_3, P_6\}, \{P_1, P_2, P_4, P_6\}, \{P_1, P_2, P_5\}, \\ & \{P_1, P_3, P_5\}, \{P_2, P_3, P_5\}, \{P_1, P_4, P_5\}, \\ & \{P_3, P_4, P_5\}, \{P_3, P_4, P_6\}, \{P_1, P_5, P_6\}, \\ & \{P_2, P_5, P_6\}, \{P_3, P_5, P_6\}, \{P_4, P_5, P_6\}\}. \end{aligned}$$

We shall realize this access structure by Ito, Saito and Nishizeki's scheme. In this case,  $\bar{\Gamma}_1$  is given by

$$\bar{\Gamma}_1 = \{B_1, B_2, \dots, B_{10}\},$$

where

$$\begin{aligned} B_1 &= \{P_1, P_2, P_3, P_4\}, \\ B_2 &= \{P_2, P_4, P_5\}, \\ B_3 &= \{P_1, P_2, P_6\}, \\ B_4 &= \{P_1, P_3, P_6\}, \\ B_5 &= \{P_2, P_3, P_6\}, \\ B_6 &= \{P_1, P_4, P_6\}, \\ B_7 &= \{P_2, P_4, P_6\}, \\ B_8 &= \{P_1, P_5\}, \\ B_9 &= \{P_3, P_5\}, \\ B_{10} &= \{P_5, P_6\}. \end{aligned}$$

- 1) Since  $|\bar{\Gamma}_1| = 10$ , compute 10 shares

$$w_1, w_2, \dots, w_{10}$$

by using a  $(10, 10)$ -threshold scheme for the secret  $K$ .

- 2) According to the function  $g_{ISN}$ , distribute shares as follows:

$$\begin{aligned} g_{ISN}(P_1) &= \{w_2, w_5, w_7, w_9, w_{10}\}, \\ g_{ISN}(P_2) &= \{w_4, w_6, w_8, w_9, w_{10}\}, \\ g_{ISN}(P_3) &= \{w_2, w_3, w_6, w_7, w_8, w_{10}\}, \\ g_{ISN}(P_4) &= \{w_3, w_4, w_5, w_8, w_9, w_{10}\}, \\ g_{ISN}(P_5) &= \{w_1, w_3, w_4, w_5, w_6, w_7\}, \\ g_{ISN}(P_6) &= \{w_1, w_2, w_8, w_9\}. \end{aligned}$$

In this scheme, to recover the secret a group  $X \subset \mathcal{P}$  need to collect all shares. If  $X \subset B_j \in \bar{\Gamma}_1$ ,  $X$  cannot collect the share  $w_j$ . On the other hand, If  $X \in \Gamma$ , then there exists  $P \in X$  such that  $P \in X - B_j$  for all  $B_j (1 \leq j \leq t)$ . Thus,  $X$  can collect  $w_1, \dots, w_t$  and recover the secret.

In this example,  $16/3$  shares are distributed on average. A disadvantage of this scheme is that the number of shares distributed to each participant becomes large as the size of  $\bar{\Gamma}_1$  gets large. This scheme needs one share for each maximal unauthorized subset. Thus this scheme needs  $|\bar{\Gamma}_1|$  shares in total. Since the number of shares distributed to each participant depends on the number of maximal unauthorized subsets, in the worst case, each of the  $n$  participants may have to hold as many as  $\binom{n-1}{k-1}$  shares.

Next, we describe the scheme I of T04 [13], which was obtained by modifying Ito, Saito and Nishizeki's scheme. Though Ito, Saito and Nishizeki's scheme is based on a special  $(k, n)$ -threshold scheme with  $k = n$ , this scheme is based on general  $(k, n)$ -threshold schemes, removing the restriction  $k = n$  and reduce the number of shares distributed to each participant, where  $k$  is the smallest size of authorized subsets. Furthermore, this scheme includes Shamir's  $(k, n)$ -threshold schemes as a special case.

*Scheme I of T04:*

- 1) Let  $\bar{\Gamma}_{1+} = \{B \in \bar{\Gamma}_1 : |B| \geq l\}$ , where  $l = \min_{A \in \Gamma} |A|$  and represent it as

$$\bar{\Gamma}_{1+} = \{B_1, B_2, \dots, B_d\}$$

with  $d = |\bar{\Gamma}_{1+}|$ .

- 2) Let  $t = \sum_{i=1}^d (|B_i| - l + 1)$  (if  $\bar{\Gamma}_{1+} = \phi$  then  $t = 0$ ). Compute  $n + t$  shares

$$S = \{s_1, s_2, \dots, s_{n+t}\}$$

for the secret  $K$  by using Shamir's  $(l+t, n+t)$ -threshold scheme.

- 3) If  $t > 0$ , partition the set  $\{s_{n+1}, s_{n+2}, \dots, s_{n+t}\}$  into subsets  $S_1, \dots, S_d$  such that

$$|S_i| = |B_i| - l + 1 \quad (1 \leq i \leq d)$$

and

$$\bigcup_{i=1}^d S_i = \{s_{n+1}, s_{n+2}, \dots, s_{n+t}\}.$$

- 4) Distribute shares to  $P_i \in \mathcal{P}$  ( $1 \leq i \leq n$ ) according to the function  $g_{T04_I} : \mathcal{P} \rightarrow 2^S$  defined as

$$g_{T04_I}(P_i) = \{s_i\} \cup \left( \bigcup_{\substack{1 \leq j \leq d \\ P_i \notin B_j}} S_j \right). \quad (2)$$

*Example 3:* We shall realize the access structure of Example 2 by the scheme I of T04 [13].

- 1) For this access structure, we have  $l = 3$  and  $\bar{\Gamma}_{1+}$  is given by

$$\bar{\Gamma}_{1+} = \{B_1, B_2, \dots, B_7\}.$$

- 2) Since  $n = 6$ ,  $l = 3$  and  $t = 8$ , compute 14 shares

$$S = \{s_1, s_2, \dots, s_{14}\}$$

for the secret  $K$  by using Shamir's (11, 14)-threshold scheme.

- 3) Since  $t > 0$ , choose  $S_i$  ( $1 \leq i \leq 7$ ) as

$$\begin{aligned} S_1 &= \{s_7, s_8\}, \\ S_2 &= \{s_9\}, \\ S_3 &= \{s_{10}\}, \\ S_4 &= \{s_{11}\}, \\ S_5 &= \{s_{12}\}, \\ S_6 &= \{s_{13}\}, \\ S_7 &= \{s_{14}\}. \end{aligned}$$

- 4) According to the function  $g_{T04_I}$ , distribute shares as follows:

$$\begin{aligned} g_{T04_I}(P_1) &= \{s_1, s_9, s_{12}, s_{14}\}, \\ g_{T04_I}(P_2) &= \{s_2, s_{11}, s_{13}\}, \\ g_{T04_I}(P_3) &= \{s_3, s_9, s_{10}, s_{13}, s_{14}\}, \\ g_{T04_I}(P_4) &= \{s_4, s_{10}, s_{11}, s_{12}\}, \\ g_{T04_I}(P_5) &= \{s_5, s_7, s_8, s_{10}, s_{11}, s_{12}, s_{13}, s_{14}\}, \\ g_{T04_I}(P_6) &= \{s_6, s_7, s_8, s_9\}. \end{aligned}$$

In this example, the scheme I of T04 requires an additional share for each participant and one more share for  $B_1$ . Hence, the scheme does not require shares corresponding to unauthorized subsets

$$\{P_1, P_5\}, \{P_3, P_5\}, \{P_5, P_6\}.$$

In this example, 14/3 shares are distributed on average. Thus this scheme is more efficient than Ito, Saito and Nishizeki's scheme from the viewpoint of the number of shares distributed to each participant. Furthermore, this scheme is optimal whenever the access structure is equal to that of a  $(k, n)$ -threshold scheme. However, there are some access structures for which this scheme is not more efficient than Ito, Saito and Nishizeki's scheme. In other word, this scheme is not always more efficient than Ito, Saito and Nishizeki's scheme.

### III. PROPOSED SCHEME

Here, we modify the scheme I of T04 [13] and propose a new secret sharing scheme realizing general access structures, which is based on unauthorized subsets. In the proposed scheme, shares are generated by Tassa's  $(\mathbf{k}, n)$ -hierarchical threshold scheme instead of Shamir's  $(k, n)$ -threshold scheme. The proposed scheme is more efficient than the scheme I of T04 [13] from the viewpoint of the number of shares distributed to each participant. For  $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$ ,  $K \in \mathcal{K}$  and  $\Gamma$ , the proposed scheme is described as follows.

*Proposed Scheme:*

- 1) Let  $\bar{\Gamma}_{1+} = \{B \in \bar{\Gamma}_1 : |B| \geq l\}$ , where  $l = \min_{A \in \Gamma} |A|$  and represent it as

$$\bar{\Gamma}_{1+} = \{B_1, B_2, \dots, B_d\}, \quad (3)$$

with  $d = |\bar{\Gamma}_{1+}|$ .

- 2) By using Tassa's  $((d, l+d), n+d)$ -hierarchical threshold scheme, compute  $n+d$  shares

$$S = \{s_1, s_2, \dots, s_{n+d}\}$$

for the secret  $K$  as follows:

$$s_j = \begin{cases} f(j) & (n+1 \leq j \leq n+d) \\ f^{(d)}(j) & (1 \leq j \leq n). \end{cases}$$

- 3) Distribute shares to  $P_i \in \mathcal{P}$  ( $1 \leq i \leq n$ ) according to the function  $g : \mathcal{P} \rightarrow 2^S$  defined as

$$g(P_i) = \{s_i\} \cup \left( \bigcup_{\substack{1 \leq j \leq d \\ P_i \notin B_j}} \{s_{n+j}\} \right). \quad (4)$$

*Example 4:* We shall realize the access structure of Example 2 by the proposed scheme.

- 1) For this access structure, we have  $l = 3$  and  $\bar{\Gamma}_{1+}$  is given by

$$\bar{\Gamma}_{1+} = \{B_1, B_2, \dots, B_7\}.$$

- 2) Since  $n = 6$ ,  $l = 3$  and  $d = 7$ , compute 13 shares

$$S = \{s_1, s_2, \dots, s_{13}\}$$

for the secret  $K$  by using Tassa's  $((7, 10), 13)$ -hierarchical threshold scheme, i.e.

$$s_j = f(j) \quad (7 \leq j \leq 13)$$

and

$$s_j = f^{(7)}(j) \quad (1 \leq j \leq 6).$$

- 3) According to the function  $g$ , distribute shares as follows:

$$\begin{aligned} g(P_1) &= \{s_1, s_8, s_{11}, s_{13}\}, \\ g(P_2) &= \{s_2, s_{10}, s_{12}\}, \\ g(P_3) &= \{s_3, s_8, s_9, s_{12}, s_{13}\}, \\ g(P_4) &= \{s_4, s_9, s_{10}, s_{11}\}, \\ g(P_5) &= \{s_5, s_7, s_9, s_{10}, s_{11}, s_{12}, s_{13}\}, \\ g(P_6) &= \{s_6, s_7, s_8\}. \end{aligned}$$

In this example, the proposed scheme requires an additional share for each participant, though the scheme I of T04 requires an additional share for each participant and one more share for  $B_1$ . The proposed scheme I does not also require shares corresponding to unauthorized subsets

$$\{P_1, P_5\}, \{P_3, P_5\}, \{P_5, P_6\}.$$

Hence, for the access structure of Example 2, the proposed scheme is more efficient than Ito, Saito and Nishizeki's scheme and the scheme I of T04 from the viewpoint of the number of shares distributed to each participant. Actually, the proposed scheme distributes 13/3 shares on average, which is smaller than 16/3 and 14/3 achieved by Ito, Saito and Nishizeki's scheme and the scheme I of T04, respectively. Furthermore, the number of shares distributed to  $P_i$  is smaller than or equal to that of the scheme I of T04 for any  $P_i \in \mathcal{P}$ .

The next theorem shows that the proposed scheme is perfect.

*Theorem 1:* For  $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$  and any access structure  $\Gamma (\subset 2^{\mathcal{P}})$ , distribute shares for a secret  $K$  by using the proposed scheme. Then, for any subset  $X \subset \mathcal{P}$ ,

- (a)  $X \in \Gamma \Rightarrow H(K|X) = 0$ ,
- (b)  $X \notin \Gamma \Rightarrow H(K|X) = H(K)$ .

*Proof:* First, we prove (a). According to the definition of  $l$  and (4), for any  $A \in \Gamma$ , we immediately have

$$\left| \{s_1, s_2, \dots, s_n\} \cap \left( \bigcup_{P \in A} g(P) \right) \right| \geq |A| \geq l. \quad (5)$$

Since  $\bar{\Gamma}_1$  is the family of maximal sets in  $\bar{\Gamma}$ , for any  $B \in \bar{\Gamma}_{1+}$  and any  $A \in \Gamma$ , there exists  $P \in A$  such that  $P \notin B$ . According to (4), we have

$$\{s_{n+1}, s_{n+2}, \dots, s_{n+d}\} \subset \bigcup_{P \in A} g(P). \quad (6)$$

Combining (5) and (6), we have

$$\left| \bigcup_{P \in A} g(P) \right| \geq l + d. \quad (7)$$

Since every shares are computed by using Tassa's  $((d, l+d), n+d)$ -hierarchical threshold scheme, we obtain  $H(K|A) = 0$  from (6) and (7).

Next we prove (b). From the definition of  $\bar{\Gamma}_1$ , for any  $B \in \bar{\Gamma}$ , there exists  $X \in \bar{\Gamma}_1$  such that  $B \subset X$ . Here, we divide the case.

(Case i)  $|X| < l$ : In this case, from (4), we have

$$\left| \{s_1, s_2, \dots, s_n\} \cap \left( \bigcup_{P \in B} g(P) \right) \right| \leq |X| < l.$$

This implies that

$$\begin{aligned} & \left| \bigcup_{P \in B} g(P) \right| \\ &= \left| \{s_1, s_2, \dots, s_n\} \cap \left( \bigcup_{P \in B} g(P) \right) \right| \\ & \quad + \left| \{s_{n+1}, s_{n+2}, \dots, s_{n+d}\} \cap \left( \bigcup_{P \in B} g(P) \right) \right| \\ & \leq d + l - 1. \end{aligned} \quad (8)$$

(Case ii)  $|X| \geq l$ : In this case,  $X \in \bar{\Gamma}_{1+}$ . Then, there exists  $B_i \in \bar{\Gamma}_{1+}$  such that  $X = B_i$ . From (4), we have

$$\{s_i\} \cap \left( \bigcup_{P \in B} g(P) \right) = \phi.$$

According to (4), we have

$$\begin{aligned} & \left| \{s_{n+1}, s_{n+2}, \dots, s_{n+d}\} \cap \left( \bigcup_{P \in B} g(P) \right) \right| \leq d - |\{s_i\}| \\ & = d - 1. \end{aligned} \quad (9)$$

Since Tassa's  $((d, l+d), n+d)$ -hierarchical threshold scheme is perfect, we obtain  $H(K|B) = H(K)$  from (8) and (9). ■

The next theorem shows that the proposed scheme includes Shamir's  $(k, n)$ -threshold schemes as a special case.

*Theorem 2:* Let  $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$ . If  $\Gamma = \{A \in 2^{\mathcal{P}} : |A| \geq k\}$ , then the proposed scheme coincides with Shamir's  $(k, n)$ -threshold scheme.

*Proof:* For any  $X \subset \mathcal{P}$ , if  $|X| \geq k$ , then  $X \in \Gamma$ . Hence, we have  $l = k$ ,  $\bar{\Gamma}_{1+} = \phi$  and  $d = 0$ . In this case, the function  $g$  can be reduced to

$$g(P_i) = \{s_i\} \quad (1 \leq i \leq n),$$

where  $s_1, s_2, \dots, s_n$  are obtained by using Tassa's  $((0, k), n)$ -hierarchical threshold scheme. ■

Let  $l = \min_{A \in \Gamma} |A|$ ,  $m = \max_{B \in \bar{\Gamma}_1} |B|$  and

$$\bar{\Gamma}_{1,i} = \{B \in \bar{\Gamma}_1 : |B| = i\} \quad (l-1 \leq i \leq m).$$

Since  $\bar{\Gamma}_1$  is the family of maximal sets in  $\bar{\Gamma}$ ,  $\bar{\Gamma}_1$  can be represented as

$$\bar{\Gamma}_1 = \bigcup_{i=l-1}^m \bar{\Gamma}_{1,i}. \quad (10)$$

Here, we evaluate the efficiency of the proposed scheme.

*Theorem 3:* For any  $P \in \mathcal{P}$ , we have

$$\begin{aligned} |g(P)| &= |g_{T04_I}(P)| \\ & \quad - \sum_{l+1 \leq i \leq m} (i-l) |\{X \in \bar{\Gamma}_{1,i} : P \notin X\}| \end{aligned} \quad (11)$$

and

$$|g(P)| = |g_{T04_I}(P)| - |\{X \in \bar{\Gamma}_{1,l-1} : P \notin X\}| + 1. \quad (12)$$

*Proof:* According to (1) and (10), we have

$$\begin{aligned} |g_{ISN}(P)| &= |\{X \in \bar{\Gamma}_1 \mid P \notin X\}| \\ &= \sum_{l-1 \leq i \leq m} |\{X \in \bar{\Gamma}_{1,i} : P \notin X\}|. \end{aligned} \quad (13)$$

On the other hand, according to (2) and (10), we also have

$$|g_{T04_l}(P)| = 1 + \sum_{l \leq i \leq m} (i - l + 1) |\{X \in \bar{\Gamma}_{1,i} : P \notin X\}|. \quad (14)$$

Similarly, we have

$$|g(P)| = 1 + \sum_{l \leq i \leq m} |\{X \in \bar{\Gamma}_{1,i} : P \notin X\}| \quad (15)$$

from (4) and (10). Equations (11) and (12) are easily obtained by (13) - (15). ■

Equation (11) shows that the proposed scheme is more efficient than or equal to the scheme I of T04 from the viewpoint of the number of shares distributed to each participant.

Here, we show some access structures in which the proposed scheme is more efficient than Ito, Saito and Nishizeki's scheme.

*Example 5:* For  $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$ , we consider the access structure such that

$$\Gamma_0 = \{X \in 2^{\mathcal{P}} : |X| = k\}.$$

This access structure can be realized by a  $(k, n)$ -threshold scheme. The number of shares distributed to each participant is evaluated as follows:

$$\begin{aligned} |g_{ISN}(P)| &= \binom{n-1}{k-1}, \\ |g(P)| &= 1. \end{aligned}$$

*Example 6:* For  $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$ , we consider the access structure such that

$$\Gamma_0 = \{X \in 2^{\mathcal{P}} : |X| = k\} - \{\{P_1, P_2, \dots, P_k\}\}.$$

This access structure cannot be realized by a  $(k, n)$ -threshold scheme. In this example, the number of shares distributed to each participant is evaluated as follows:

$$\begin{aligned} \binom{n-1}{k-1} - k + 1 &\leq |g_{ISN}(P)| \leq \binom{n-1}{k-1} - 1, \\ 1 &\leq |g(P)| \leq 2. \end{aligned}$$

The scheme I of T04 [13] distributes at most two shares to each participant, though Ito, Saito and Nishizeki's scheme distributes at least  $\binom{n-1}{k-1} - k + 1$  shares to each participant.

From (12) we have the following lemma.

*Lemma 1:* If an access structure  $\Gamma$  satisfies

$$|\bar{\Gamma}_{1,l-1}| \geq \frac{n}{n-l+1}$$

then we have

$$\sum_{P \in \mathcal{P}} |g(P)| \leq \sum_{P \in \mathcal{P}} |g_{ISN}(P)|.$$

As mentioned above, the proposed scheme is not always more efficient than Ito, Saito and Nishizeki's scheme, though the proposed scheme is always more efficient than or equal to the scheme I of T04. This lemma characterizes the access structure in which the proposed schemes is more efficient than Ito, Saito and Nishizeki's scheme from the viewpoint of the number of shares distributed to each participant on average.

#### IV. CONCLUSION

We have proposed a new secret sharing scheme realizing general access structures. In the proposed scheme, shares are generated by Tassa's  $(\mathbf{k}, n)$ -hierarchical threshold scheme instead of Shamir's  $(k, n)$ -threshold scheme. Thus, the proposed scheme can reduce the number of shares distributed to each participant. The proposed scheme is more efficient than or equal to the scheme I of T04 [13] from the viewpoint of the number of shares distributed to each participant.

#### ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 15K00192.

#### REFERENCES

- [1] A. Shamir, "How to share a secret," *Comm. ACM*, Vol.22, No.11, pp.612-613, 1979.
- [2] G. Blakley, "Safeguarding cryptographic keys," *Proceedings of AFIPS*, 48, pp. 313-317, 1979.
- [3] K. Koyama, "Cryptographic key sharing methods for multi-groups and security analysis," *Trans. of the IECE*, Vol. E66, No. 1, pp.13-20, 1983.
- [4] G. Simmons, "How to (really) share a secret," *Proc. of CRYPTO '88*, pp. 390-448, 1988.
- [5] G. Simmons, "Prepositioned shared secret and/or shared control schemes," *Proc. of EUROCRYPT '89*, pp. 436-467, 1989.
- [6] T. Tassa, "Hierarchical threshold secret sharing," *Journal of Cryptology* Vol. 20, pp. 237-264, 2007.
- [7] O. Farràs and C. Padró, "Ideal hierarchical secret sharing schemes," *IEEE Trans. on IT* Vol.58, No.5, pp. 3273-3286, 2012.
- [8] H. Sun, H. Wang, B. Ku and J. Pieprzyk, "Decomposition construction for secret sharing schemes with graph access structures in polynomial time," *SIAM Journal on Discrete Mathematics*, Vol. 24, No.2, pp.617-638, 2010.
- [9] L. Csirmaz, P. Ligeti and G. Tardos, "Erdős-Pyber theorem for hypergraphs and secret sharing," *Graphs and Combinatorics* Vol. 31, No.5, pp.1335-1346, 2015.
- [10] A. Beimel, Y. Mintz and O. Farràs, "Secret-sharing schemes for very dense graphs," *Journal of Cryptology* Vol. 29, Issue 2, pp. 336-362, 2016.
- [11] M. Ito, A. Saito and T. Nishizeki, "Secret sharing scheme realizing general access structure," *Proc. IEEE Globecom '87*, pp.99-102, 1987.
- [12] J. Benaloh and J. Leichter, "Generalized secret sharing and monotone functions," *Proc. of CRYPTO '88*, pp.27-35, 1988.
- [13] K. Tochikubo, "Efficient secret sharing schemes realizing general access structures," *IEICE Trans. Fundamentals*, Vol. E87-A, No.7, pp. 1788-1797, 2004.
- [14] K. Tochikubo, "Efficient secret sharing schemes based on unauthorized subsets," *IEICE Trans. Fundamentals*, Vol.E91-A, No.10, pp.2860-2867, 2008.
- [15] K. Tochikubo, T. Uyematsu. and R. Matsumoto, "Efficient secret sharing schemes based on authorized subsets," *IEICE Trans. Fundamentals*, Vol.E88-A, No.1, pp.322-326, 2005.
- [16] K. Tochikubo, "New construction methods of secret sharing schemes based on authorized subsets," *Journal of Information Processing*, vol.21, No.4, pp.590-598, 2013.
- [17] D. R. Stinson, *Cryptography: theory and practice* 3rd edition, CRC Press, 2005.
- [18] E. D. Karnin, J. W. Greene and M. E. Hellman, "On secret sharing systems," *IEEE Trans. on IT* Vol.29, No.1, pp.35-41, 1983.

# Bot or Not

Husna Siddiqui, Elizabeth Healy, Aspen Olmsted

Department of Computer Science

College of Charleston

Charleston, SC 29401

siddiquih@g.cofc.edu, healye@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—In recent years, Twitter, a social networking website, has been affected by a steady rise in spam on its network. Hijacking of social media accounts has become a modern-day danger. Motivations for this can range from attempts in identity theft to simply skewing the perception of an audience. In this paper, we extend our previous work, *Engineering Your Social Network to Detect Fraudulent Profiles*, by doing an investigation of spam bots on Twitter. We propose an algorithm that will distinguish a spam bot, from a genuine user account by using a JavaScript testing framework that consumes Twitter's REST API. We ran a dataset of 700 Twitter accounts through our algorithm and identified that roughly 11% of the dataset were bots.

**Keywords**-spam; social network;bot; API; Twitter.

## I. INTRODUCTION

Social media is changing the world. Networking websites such as Twitter, Facebook, and Instagram have become some of the most popular activities. For many, it is an important aspect of daily life. These websites influence the news that users consume on a day to day basis. Spam bots are automated accounts that mimic real users. In some cases, these bots can be harmless. For example, there is a prime numbers bot that systematically tweets prime numbers. However, some bots can be malicious and are intended to tamper with commonly tracked statistics by posting automated and false information.

In this paper, we extend our previous results [1] to determine the validity of a user account on Facebook. Here we will investigate the detection of bots on Twitter. To do this, we use Twitter's REST API, application program interface, to collect and analyze user data. The organization of this paper is as follows: Section II will describe the works that are related to our research area. Section III further explains the motivation for our research. Section IV explains our implementation method and data gathering techniques. Section V gives an overview of our results. Finally, Section VI will give a conclusion of our work.

## II. LITERATURE REVIEW

The most important document we will consider is Twitter's developer API. The API will allow us to integrate with Twitter and programmatically access Twitter data, so we can gather a database of records to run through our algorithm.

In our previous work [1], we proposed an improvement to the calculation of a validity score for a Facebook user account to determine if the profile is real or fake. We used empirical analysis, to define the attributes that impact a user's validity

score. Each of these attributes was given a weighted score, which was then aggregated to determine a validity score.

In his paper, *Detecting Spam Bots in Online Social Networking Sites*, Wang extracts Twitter data to discover spam bots. Wang analyzes three graph based features which are the number of friends, the number of followers, and the follower ratio. He also looks at three content-based features, which are the number of duplicate tweets, the number of HTTP links, and the number of replies from the user's 20 most recent tweets. [2] This data is run through various classification methods to determine if an account is a spam bot or not. We will use findings from Wang's research as a basis of our knowledge when identifying which features are most significant in identifying a spam bot.

In *Securing e-Loyalty Currencies*, Olmsted analyzed user activity on various social networks to determine the validity of a user's social network rewards. To validate the authenticity of a user account, particularly for Twitter, the following account data was used: the presence of an account, the number of people following this account, number of accounts this account follows and the number of times the account has tweeted a microblog. [3] Each attribute has a maximum cap to ensure that one feature does not contribute too much weight to the validity of a user.

Freitas, Benevenuto, Ghosh and Veloso designed an experiment to understand infiltration strategies of socialbots on Twitter. They created 120 socialbot accounts with different characteristics and strategies to investigate what constitutes a successful bot. [4] Their study reveals findings that are key for the detection and counter measurement approaches for bots on Twitter.

## III. KNOWLEDGE GAP

Twitter's appeal is in the use of "tweets." A tweet is a status message that is limited to 140 characters. This warrants that a tweet is concise and can be easily scanned. If you see a tweet that you find interesting, you can "like" the tweet to let the user know that you enjoy their content. Additionally, you can "follow" the corresponding account to view and get updates on their account activity. Twitter also has functionality that allows users to "retweet" a tweet which instantly shares and spreads information. Popular topics on Twitter include politics and news, sports, fashion, and pop culture. Users, can skim through tweets and read the trending topics to get a quick update on what is happening around the world.

TABLE 2. Sample Range of Attribute values and Scores

Twitter Attribute	User Range	Score
ratio	0, <=0.2, >= 0.3	10, 5 0
default profile image	False or true	0, 15
keywords in tweets	Zero, 1-2, >= 3	0, 3, 15
url patterns in statuses	Zero, 1-2, >= 3	0, 3, 15
Statuses-count (frequency)	0 -3.0, 3.1-5.0, 5.1-10.0, > 10.0	0, 5, 10,15
verified	False or true	0, 10

Currently, Twitter has 319 million monthly active users. [5] Based on research from the University of Southern California and Indiana University, up to 15% of these are bots. This means that roughly 48 million accounts are bots, not humans. We strive to analyze Twitter data to expose these bots and uncover commonalities that may relate them.

Spam bots can be used for malicious acts such as the spread of fake news, cyber-stalking, spread of malware and clickjacking. We propose a way to detect bots and educate users so that they are aware of these issues.

IV. RESEARCH AND IMPEMENTATION

Our first step was to identify characteristics of a successful Twitter bot. By utilizing knowledge from the research community, we determined the following attributes to be the most significant: ratio of number of followers and number of following, profile image, keywords in tweets, url patterns in tweets, whether it is a verified account and the status count. Each of these attributes provides a different weight towards the calculation of a score to uncover a spam bot. TABLE 1 gives sample extended rules that are used to assign a score. Each attribute has a specific maximum score. This max score is based on the predictive significance of each metric. For example, the verification of an account contributes a smaller weight than the ratio of followers versus following.

The next step is to gather a dataset by using the Twitter API. We use the followers/list.json method to acquire details on 700 random twitter accounts. This endpoint returns a JSON object that contains fields such as username, id, verification-status, followers, friends, status count and profile image. We extract relevant values from the JSON object. After we gather our dataset, we build an array of the usernames that were returned. We use these usernames to make an API call to the search/tweets.json endpoint which requires the username and a query string that contains the keywords we want to search for. We target the following keywords in a user’s status to flag it is a bot: offer, free, click, prize, debt, deal, credit, and sex. The last API call hits the search/tweets.json endpoint, and this time we look for url patterns in a user’s tweet.

After we record the results and have fully iterated through the array of users, we run a script to iterate through the array

to assign a score based on each attribute value. TABLE 2 details the process of assigning a spam bot score. If an attribute falls within a range, the user is given a score for that attribute. Finally, the user is assigned a total score by adding up the points for each attribute. An overall score of 40 or more signifies that the account is a spam bot.

TABLE 1. SAMPLE USER ATTRIBUTE POINTS

Sample Attribute	Score	Max
ratio	10	10
default profile image	5	15
keywords in tweets	15	15
url patterns in statuses	2	15
Statuses-count	15	15
verified	15	10
<b>Total Points</b>	<b>77</b>	<b>80</b>

V. RESULTS

Fig. 1 shows the results of our implementation. We ran a sample database of 700 Twitter accounts through our algorithm and identified 79 accounts as bots, which is roughly 11% of our dataset.

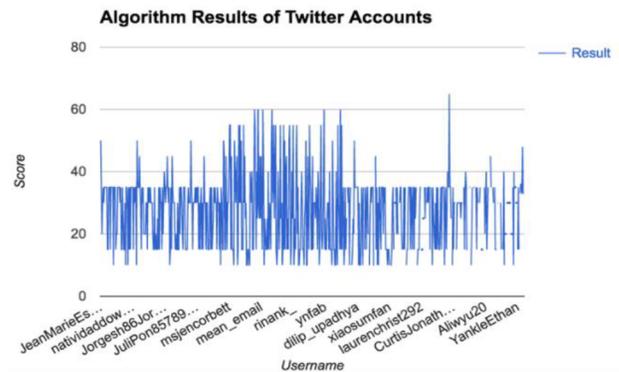


Figure 1. Implementation Results

VI. CONCLUSION

In this paper, we propose a method to identify spam bots on Twitter. Our solution is based on empirical analysis and a JavaScript testing framework that uses Twitter’s REST API. We identified roughly 11% of our dataset to be spam bots.

REFERENCES

- [1] A. O. Z. D. Husna Siddiqui, "Engineering your social network to detect fraudulent profiles," in *Information Society (i-Society)*.
- [2] A. H. Wang, "Detecting Spam Bots in Online Social Networking Sites: A Machine Learning Approach."
- [3] A. Olmsted, "Securing e-Loyalty Currencies," *Journal of Internet Technology and Secured Transactions*.
- [4] F. B. S. G. a. A. V. Carlos Freitas, "Reverse Engineering Socialbot Infiltration Strategies in Twitter".
- [5] M. Newberg, "CNBC," 20 03 2017. [Online]. Available: <http://www.cnbc.com/2017/03/10/nearly-48-million-twitter-accounts-could-be-bots-says-study.html>. [Accessed 20 04 2017].

# Intelligent Laboratory Management System Based on Internet of Things

Yichen Ma, Fuyao Wang, Zhuozheng Wang  
Beijing University of Technology  
Beijing, China  
Zjz2013b@126.com

**Abstract**—For the purpose of optimizing laboratory management, a set of intelligent laboratory management system based on Internet of things is described in this paper. The hardware platform of this system is STM32 micro-controller, adopts WIFI intelligent power module, RFID-RC522 card reader, using Android/Java language to develop raspberry 3. When the students get to the laboratory and put their student cards on, the system would read the student information in the cloud database and find the student's course information, recorded in the STM32 micro-controller and display the information on the raspberry3. According to the information, the system would sign the data in the cloud database. Achieved an efficient laboratory intelligent information management system based on Internet of things.

**Keywords**-Laboratory; Internet of things; Android; Raspberry Pi 3; Cloud Database

## I. INTRODUCTION

With the rapid development of the Internet in recent decades, all walks of life have been inseparable from the application of the Internet. The Internet helps people improve the quality of life and work efficiency. In order to meet the growing market demand, and constantly meet the process of social development, intelligent and convenient Internet of things on the basis of the Internet came into being. The study and development of Internet of Things applications, web and mobile, is on the increase. Applications, working with data obtained from different areas such as transportation, smart homes, health care, public services, industry and many others.[1] Internet of Things is a platform where every day devices become smarter, every day processing become intelligent, and every day communication become informative. While the Internet of Things is still seeking its own shape, its effects have already started in making incredible strides as a universal solution media for the connected scenario.[2] With the transformation of the construction and management of the gradual progress of the school, the laboratory construction, maintenance and application management put forward higher and higher requirements, which urgently need to use advanced technology to standardize and strengthen the laboratory management. Promotion an open intelligent classroom management software in classroom is necessary.[3] To this end, we have designed and developed a laboratory intelligent management system based on the Internet of Things technology. The system is designed to integrate the computer technology, database technology and Internet technology

including RFID technology and sensing technology. Making laboratory management more humane, standardized, and helping teachers to manage the laboratory easily.

## II. DESIGN PRINCIPLES

In order to save energy and to implement the intelligent of the classroom management.[4] The management of the entire laboratory should follow the following principles:

- Advanced, integrated principle. On the basis of the Internet of Things, to achieve the purpose of making relation between machines, between people and machines, and between machines and networks, to improve the overall system of the advanced, practical. It would greatly improve the efficiency of the entire system. At the same time, this form greatly improved the level of intelligent management, to achieve the concept of laboratory group management.
- The principle of extensibility. For the current system management, we should make the management open and make cross-platform applications. So as to improve the overall control of diversity, greatly improve the control capacity.
- The principle of safety. School management system is also a very important internal system. Greatly improve the safety performance and improve reliability is also necessary.
- Practical principle. The design of the system should be corresponding to the needs of teachers and students. The specific circumstances should be analyzed separately. The main purpose of the system should be teaching better, so don't forget to pay more attention to practical.

## III. SCHEME DESIGN

Figure 1 shows the structure of the entire Internet of things laboratory management system. The figure could be divided into three parts: STM32, Raspberry Pi 3 and the cloud database. STM32 is responsible for the storage and transmission of most of the data. Raspberry Pi 3 is mainly responsible for human-computer interaction and display cards information, the cloud database is mainly responsible for storing large amounts of data uploaded and for the system query.



A. *RFID Reader Functions*

When the reader reads a plurality of student cards at the same time, it is necessary to introduce an anti-collision mechanism. If there is no anti-collision mechanism, RFID-RC522 can only read and write one student card, in this case if there are more than two student cards in the range of read and write will lead to read and write errors.

The most important core of these functions is about finding cards, anti-collision mechanism, selecting cards and reading cards.

B. *Communication Protocol*

As the Json packet efficiency is relatively high, so developing a communication protocol for Json is necessary.

TABLE I. PACKET DEFINITION

Start	Tag	Length	Data	Verification	End
1 byte	1 byte	2 byte	N	1 byte	1 byte

Start: Only data packets beginning with 0x55 are valid.

Length: 2 bytes indicates the "data" length.

Data: Based on the length of the transmitted data. Is the student information or electricity information data.

Tag: 0x01 is the student information, 0x02 is the power information.

Verification: Data bit xor (hexadecimal same bit is 0, different bit is 1). Generates a data having the packet feature, and compare.

End: Only packets ending with 0x56 are valid.

Write a program to compress the data. The format of the packet is written to the buffer in bytes and returns the length of the data for inspection. After the database is created, the data can be obtained by STM32 and transmitted stably to the Raspberry Pi 3.

C. *Android Program Development*

Use the open source serial port android-serialport-api from Google to connect the Cubieboard Android platform to the microcontroller and write the program.

In the Android application, the first step is to analyze the Json packet came from STM32, otherwise it will make the serial number shown as messy code. Write a Java class import Gson jar package to remove all the messy code.

The next step is to convert RF card information into json format and then json data containing the RF card information would be converted into class information.

The system flow chart is as follows:

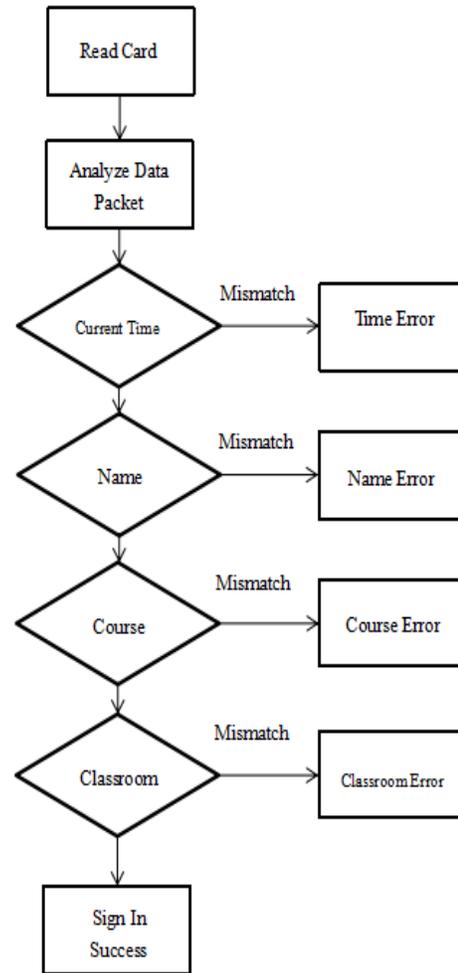


Figure 4. System Flow Chart

The system uses JDBC to connect Baidu cloud RDS Mysql5.6 database. The android project should be imported the mysql jar package. In the program, the cloud database address and account and password also should be written, to establish a connection. After the student card serial number are read by RFID reader, use sql statement in the database to find the corresponding students and their corresponding class. Through the class number and the current system time to query the current course information. Check the course number and then check the classroom of this course. If the information is correct, the certification is successful. Use the update statement to sign in the database. If any of the above matches fails, an error message would be displayed and the student is instructed to attend the right classroom

VI. EXPERIMENT RESULTS AND ANALYSIS

In the experiment, the response time is stable and fast enough to meet the needs of intelligent laboratory management.

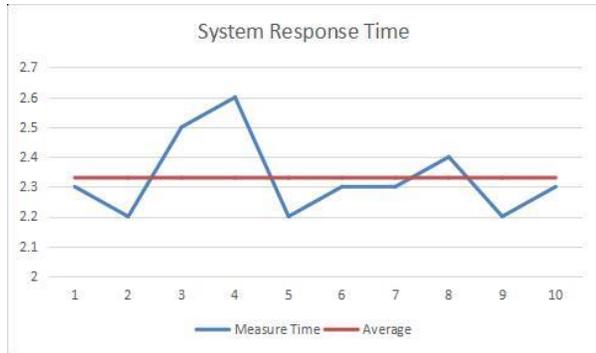


Figure 5. Response Time

Date	Classroom	Time	Total	Actual
04/04/2017	503	8:00-9:40	29	28
11/04/2017	503	8:00-9:40	29	27
18/04/2017	503	8:00-9:40	29	29
25/04/2017	503	8:00-9:40	29	28
02/05/2017	503	8:00-9:40	29	26

Figure 6 Check-in Display

VII. CONCLUSIONS

The system is based on the principle of Internet of things. The use of STM32 microcontroller, RFID-RC522 card reader, raspberry pi 3 is to achieve the intelligent management of the laboratory. In the practical application, the results are very good, the system greatly reduced the workload of staff. In the future development, the system could be developed with more features, such as student course information analysis, intelligent switch computer and other functions. The overall practicality, reliability of this system is great and it also has a great development prospects. Smart sign-in system will become the new style of teaching in the future.

ACKNOWLEDGMENT

This research was supported by our professor Zhuozheng Wang. The knowledge he had taught to us helps us a lot to finish this research.

REFERENCES

- [1] Mayra Samaniego, Ralph Deters. Management and Internet of Things[J]. Amsterdam:Procedia Computer Science. 2016.08.022
- [2] P.P. Ray.A Survey on Internet of Things Architectures[J]. Amsterdam:Journal of King Saud University. 2016.10.003
- [3] Yan Yu,Jian Hua Wang. Model Design of Universal Open Intelligent Classroom Teaching Software[J]. Adamstown:Advanced Materials Research.2012:207-212
- [4] Hong Mei Wang, Xu Ming Long, Hong Yao Cao. The Classroom Intelligent Control System Based on Wireless Communication[J].Adamstown: Advanced Materials Research:2014:288-292
- [5] Ahmed Jedda, Mazen G. Khair,Hussein T. Mouftah. Decentralized RFID Coverage Algorithms Using Writeable Tags[J]. Amsterdam: Computer Networks. 2016.03.003
- [6] Dhvani Shah, Vinayak haradi. IoT Based Biometrics Implementation on Raspberry Pi[J]. Amsterdam: Procedia Computer Science. 2016.03.043

## **Session 21: Infonomics and e-Technology**

Title: Spark framework for transcriptomic trimming algorithm reduces cost of reading multiple input files

(Authors: Walter Blair, Aspen Olmsted, Paul Anderson)

Title: Avoiding unnecessary deaths Drag-Back, a deadlock avoidance model

(Authors: Luis Mejía-Ricart, Aspen Olmsted)

Title: Role of Transformational Leadership on E-Government Switching: Multi-Channel and Digital Divide

(Authors: Khurram Mahmood, Zainab Nayyar, Hafiz Mushtaq)

Title: Shared Situational Awareness in Information Security Incident Management

(Authors: Keshnee Padayachee, Elias Worku)

Title: A Collaborative System for Corporate Performance Evaluation using Gamification and the Learning Vectors Model

(Authors: Michelle G. Cacaïs, Gilvandenys L. Sales)

# Spark framework for transcriptomic trimming algorithm reduces cost of reading multiple input files

Walter Blair, Aspen Olmsted, Paul Anderson

Department of Computer Science

College of Charleston, Charleston, SC

wmblair@cofc.edu, olmsteda@cofc.edu, andersonpe2@cofc.edu

**Abstract** - In this paper, we investigate the feasibility and performance improvement of adapting a common stand-alone bioinformatics trimming tool for in-memory processing on a distributed Spark framework. The rapid and continuous rise of genomics technologies and applications demands fast and efficient genomic data processing pipelines. ADAM has emerged as a successful framework for handling large scientific datasets, and efforts are ongoing to expand its functionality in the bioinformatics pipeline. We hypothesize that executing as much of the pipeline as possible within the ADAM framework will improve the pipeline's time and disk requirements. We compare Trimmomatic, one of the most common raw read trimming algorithms, to our own simple Scala trimmer and show that the distributed framework allows our trimmer to suffer less overhead from increasing the number of input files. We conclude that executing Trimmomatic in Spark will improve performance with multiple file inputs. Future work will investigate the performance benefit of passing the distributed dataset directly to ADAM in memory rather than writing out an intermediate file to disk.

**Keywords**-transcriptomics, Spark, ADAM

## I. INTRODUCTION

Whereas genomics is the study of an organism's genetic blueprint stored as stable DNA, transcriptomics is the study of an organism's genetic activity expressed as unstable RNA. DNA contains information on what an organism is genetically capable of doing throughout its life, but RNA provides information on which genes are being turned on and which genes are being shut off and thereby provides a snapshot of what various cells in the body are currently doing [1]. Due to the falling cost of gene sequencing, real-time RNA diagnostics will soon become a routine medical procedure. More and more data will need to be processed very quickly to inform medical decisions, so there is a pressing need for faster, and more standardized genomic data processing [2].

A relatively recent development in transcriptomics is the use of distributed processing via Apache Spark to speed up the data processing pipeline [3], [4]. One Spark-based genomics processing framework that has been particularly successful is ADAM, developed by the Big Data Genomics partnership between UC Berkeley and the Broad Institute.

ADAM tunes Spark for genomics dataset by combining its in-memory processing with custom columnar schemas and storage formats using Apache Avro and Apache Parquet, respectively, as well as modified coordinate joins that work more efficiently than Spark's out-of-the-box functionality [5]. There is a great deal of interest in extending the functionality of ADAM in an effort to offer some level of standardization, cohesion, and interoperability in the genomics processing pipeline. One possible strategy for building up the pipeline would be to plug existing bioinformatics tools directly into the ADAM framework. For example, raw sequence trimming is one of the first steps in the processing pipeline and is usually executed on a local machine before the dataset is passed to a distributed framework like ADAM [6]. If the trimming process could instead be distributed, then overall time, memory, and disk space requirements would be reduced. There are several promising research questions that arise from such efforts to incorporate more and more of the transcriptomics processing pipeline into the increasingly popular ADAM / Spark framework. In the current research, we investigate the distribution of raw read trimming as a sort of proof of concept that a number of other existing tools can also be distributed to achieve the same benefits demonstrated here.

The organization of this paper is as follows. Section II describes the related work and the current alternative solutions for solving the big data pipeline problem in bioinformatics. Section III states the guiding hypothesis for the current research. Sections IV documents the methods that have so far been implemented in pursuit of our hypothesis, and Section V presents our preliminary findings and current aims for the immediate future.

## II. RELATED WORK

Of particular interest is an alternative approach to solving the bioinformatics pipeline that does not use the ADAM / Spark framework. Diao and Bloom [7] present a full pipeline that uses the Hadoop distributed file system with a custom scheduler that optimizes I/O operations and allows users to incorporate existing stand-alone tools into the pipeline without additional modification. The authors argue against ADAM's approach because 1) a number of algorithms in the pipeline are inherently quite difficult to parallelize, and 2) the need to reimplement tools to fit into the ADAM framework is burdensome.

Deng *et al.* [8] recently presented HiGene, a genomics processing platform built on Spark and reported a dramatic reduction in execution time for the genomics pipeline from days to on the order of an hour. Fagerli [9] published a graduate thesis presenting COMBUSTI/O based on parallelized Spark workflows inspired by the ADAM project and demonstrated its utility in several bioinformatics use cases.

### III. HYPOTHESIS

Our current hypothesis is that performing as much of the pipeline as possible in memory via Spark is preferable to reading and writing files between steps in the pipeline. To test this hypothesis, we will write a simple Scala-based sequence-trimming program that is able to read in a raw text file, trim the sequences within a Spark dataset, and either 1) write out a text file or 2) pass the dataset directly to ADAM in memory. Our expected result is that writing out a “middleman” text file will require more time and disk space than passing the dataset directly to ADAM. This experiment will provide a clear demonstration that an in-memory processing pipeline will perform better than a pipeline that reads and writes intermediate files.

### IV. METHODS

We timed our Scala-based trimming algorithm on a directory containing either one, five, or ten raw fastq files, and it wrote out to several trimmed fastq files depending on the partition setting. The code for our Scala-based trimming algorithm is available and in progress on github (<https://github.com/waltermblair/ScalaTrim>). We compared our Scala trimmer's performance to one of the leading stand-alone trimmers Trimmomatic [10] (<http://www.usadellab.org/cms/?page=trimmomatic>) using the CLI options SE, LEADING:25, TRAILING:5. We timed Trimmomatic by calling it either one, five, or ten times on a single raw fastq file to write out one, five, or ten trimmed fastq files. All tests were executed on a local machine with a Intel® Xeon(R) CPU E3-1270 v5 @ 3.60GHz × 8 processors with 62.8GiB memory running ubuntu 16.04 LTS. The Spark environment for our Scala trimmer was a local cluster running Spark 2.1.0 and a local HDFS on Hadoop 2.7.3.

### V. PRELIMINARY RESULTS

Our results so far have compared the performance of our simple Scala-based in-memory sequence trimmer to one of the most widely used trimming programs Trimmomatic. Our trimmer is slower than Trimmomatic (TABLE 1), but because our Spark-enabled trimming algorithm is able to optimize the input of multiple sequence files at once, we do not suffer the same loss

TABLE 1. COMPARISON OF EXECUTION TIME

File Size	# Files	Scalatrims (s)	Trimmomatic (s)	File Size	# Files
100MB	1	15	1	100MB	1
100MB	5	25	3	100MB	5
100MB	10	31	5	100MB	10

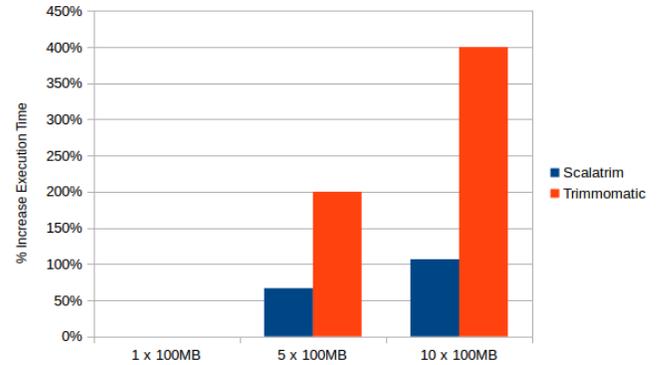


Figure 1. Execution Time Increase

of performance when trimming multiple files compared to Trimmomatic (FIGURE 1). This finding is important because it suggests that a high-performance algorithm like Trimmomatic would see enormous advantages if performed in a Spark framework. Moreover, the Scala trimmer used resources to place the input file into a Spark dataset and then write it back out to disk, but Trimmomatic performed no such conversion into a Spark-enabled context.

We hope in future research to be able to wrap or in some other way convert existing tools like Trimmomatic in such a way that they are made Spark-ready without any modification to the underlying algorithm. We will be able to test Hypothesis 1 after we are able to pipe our Scala-based trimmed dataset directly to ADAM. We will then compare performance to the current function of writing out to a text file. Our ability to plug our application directly into ADAM will go a long way toward answering our future questions around plugging other tools directly into ADAM.

### REFERENCES

- [1] T. Lappalainen *et al.*, “Transcriptome and genome sequencing uncovers functional variation in humans,” *Nature*, vol. 501, no. 7468, pp. 506–511, 2013.
- [2] R. Bao *et al.*, “Review of Current Methods, Applications, and Data Management for the Bioinformatics Analysis of Whole Exome Sequencing,” *Cancer Inform.*, vol. 13, no. S2, pp. 67–82, 2014.
- [3] M. S. Wiewiorka, A. Messina, A. Pacholewska, S. Maffioletti, P. Gawrysiak, and M. J. Okoniewski, “SparkSeq: Fast, scalable and cloud-ready tool for the interactive genomic data analysis with nucleotide precision,” *Bioinformatics*, vol. 30, no. 18, pp. 2652–2653, 2014.
- [4] S. Salloum, R. Dautov, X. Chen, P. X. Peng, and J. Z. Huang, “Big data analytics on Apache Spark,” *Int. J. Data Sci. Anal.*, vol. 1, no. 3, pp. 145–164, 2016.
- [5] F. Austin Nothhaft *et al.*, “Rethinking Data-Intensive Science Using Scalable Analytics Systems,” *Proc. 2015 ACM SIGMOD Int. Conf. Manag. Data*, pp. 631–646, 2015.
- [6] M. Szczerba, M. S. Wiewiorka, M. J. Okoniewski, and H. Rybiński, “Scalable Cloud-Based Data Analysis Software Systems for Big Data from Next Generation Sequencing,” in *Big Data Analysis: New Algorithms for a New Society*, vol. 16, 2016, pp. 263–268.

- [7] Y. Diao, A. Roy, and T. Bloom, "Building Highly-Optimized, Low-Latency Pipelines for Genomic Data Analysis," *Proc. 7th Bienn. Conf. Innov. Data Syst. Res.*, 2015.
- [8] Liqun Deng, Guowei Huang, Yuzheng Zhuang, Jiansheng Wei, and Youliang Yan, "HiGene: A high-performance platform for genomic data analysis," in *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2016, pp. 576–583.
- [9] J. Fagerli, "Combust I/O: Abstractions facilitating parallel execution of programs implementing common I/O patterns in a pipelined fashion as workflows in Spark."
- [10] A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: A flexible trimmer for Illumina sequence data," *Bioinformatics*, vol. 30, no. 15, pp. 2114–2120, 2014.

# Avoiding unnecessary deaths

## Drag-Back, a deadlock avoidance model

Luis Mejía-Ricart, Aspen Olmsted

Computer Science Department

College of Charleston

Charleston, SC, United States

mejriaricartlf@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—Real world distributed database systems must be ready to handle with thousands of concurrent transactions, if not more. Current concurrency control mechanisms such as Wait-Die and Wound-Wait can deal with deadlocks, but not always do so efficiently. In a system that uses serialization isolation, it is common to see these mechanisms rolling back whole transactions to prevent or avoid deadlocks. If transactions are very large, this potentially results in much wasted I/O operations. We propose a different approach to avoid deadlocks limiting the number of wasted resources by implementing Drag-Back, a model that introduces the concept of partial rollbacks of transactions without compromising their ACID properties.

**Keywords**—component; database; deadlocks; scheduling; concurrency;

### I. INTRODUCTION

Concurrency is an intrinsic challenge of shared, distributed systems. Distributed database systems are expected to handle multiple transactions in parallel at an acceptable rate without compromising the integrity of their data or their availability. In very large database systems, you may see the number of concurrent transactions rise to the millions. As Kaur et al. discussed in [1], not only is it expected to have better performance in a distributed system, but it is also crucial for distributed databases to serve incoming transactions without compromising their ACID properties. The inherent complexity of handling concurrency in a serializable manner adds another layer of management overhead to these systems but, thankfully, there are numerous algorithms that tackle the issue of concurrency control [1], [2], [3]. These concurrency control mechanisms are built using techniques such as resource locking, transaction logging, and transaction scheduling to prevent inconsistent database states and deadlocks. However, these mechanisms, albeit proven functional, are not most efficiently using computational resources.

One such example is seen in deadlock avoidance mechanisms Wait-Die and Wound-Wait, which cause transactions to die (or be wounded). These transactions are forced to be rolled back entirely when transactions in conflict with another transaction, regardless of how much overlap there is between them. These rollbacks, or wounds, may result in an unnecessarily high amount of high amount of I/O operations, not only wasting time but also, and this is especially true for solid-state drives, consuming the limited amount of disk writes, resulting in a lower lifespan of the drive. The amount of I/O operations caused by rollbacks can be reduced by leveraging the

functionality of compensation logs, a type of write-ahead log entry (WAL entry), to implement partial rollback functionality for conflicting transactions.

Given two concurrent, timestamped (TS), conflicting transactions  $T_1$  and  $T_2$ , where  $T_1$  is older than  $T_2$  ( $TS(T_1) < TS(T_2)$ ) but  $T_2$  acquired the lock to a resource  $R$  first, a partial rollback or a drag-back consists of rolling back the actions pertaining to transaction  $T_2$  up to the last action before  $T_2$ 's first action over the conflicting resource  $R$ . Normally,  $T_2$  would be undone entirely and restarted later. This contrasts with partial rollbacks where transactions need neither be killed nor restarted. In the rest of this work, we expand into the concept of partial rollbacks and our Drag-Back scheme to deadlock avoidance.

### II. RELATED WORK

Locking is the concurrency control strategy of choice [3]. However, one of the drawbacks of locking is the possibility of deadlocks happening. Concurrent systems must implement good strategies to prevent or avoid deadlocks. We present popular techniques used by systems to deal with deadlocks. We end this section with a glimpse into a logging strategy that helps provide transactions with isolated, serialized environments.

#### A. Deadlock Prevention

One very effective deadlock prevention mechanism consists of delaying a Transaction until it acquires all its required locks [1]. This is too optimistic an approach when faced with a database system that handles thousands of concurrent transactions. Transactions with overlapping resources may be unnecessarily delayed due to one *waiting* transaction holding the locks.

#### B. Deadlock Avoidance

Two popular schemes for avoiding deadlocks altogether are the timestamp-dependent Wait-Die (non-preemptive) and Wound-Wait (preemptive) algorithms [1] [2] [3].

Given a transaction  $T$  with timestamp  $TS(T)$ , when a transaction attempts to acquire a lock currently held by  $T$ , the wait-die and wound-wait algorithms do the following, respectively:

1) *Wait-Die*: Compares both transactions' timestamps. If the requesting transaction is younger than  $TS(T)$ , meaning its timestamp is greater than  $TS(T)$ , then it can wait until  $T$  releases the lock. If the requesting transaction is older than  $T$ , meaning

its timestamp is less than  $TS(T)$ , the requesting transaction dies and is restarted with a new timestamp.

2) *Wound-Wait*: Similar to Wait-Die, when a transaction attempts to acquire a lock currently held by T, Wound-Wait proceeds to compare their timestamps. It differs from Wait-Die in that the requesting transaction can wait until T releases the lock if the new transaction is younger than T. On the other hand, if the requesting transaction is older than T, it wounds T and acquires the lock. In this context, wounding means “to roll back entirely.” T is then restarted with its timestamp unchanged and waits for the lock to be available again.

3) *Log Sequence Numbers (LSN)*: Databases use logging techniques to recover from failures. In serializable isolation, transactions can only see a version of the database that corresponds to the state of the database at the time of arrival of said transaction. In short, if  $T_1$  arrived before  $T_2$ , then  $T_1$  will not see any changes made by  $T_2$  even if both transactions are happening concurrently. LSNs helps achieve serialization by versioning database states. Later on, we’ll explain how the implementation of undo actions in versioned logging strategies help achieve partial rollbacks.

### III. THE DRAG-BACK SCHEME

The Drag-Back scheme leverages current implementations of database engines to power partial transaction rollbacks that will save time and I/O’s when dealing with deadlocks.

#### A. Partial rollbacks

Before diving into partial rollbacks, it is important to note that partial rollbacks, in a way, challenge the notion that database transactions are atomic. In no way, however, does this contradict the principle of database consistency that stipulates that all transactions must leave the database in a consistent state.

A transaction may consist of any combinations of read/write operations. We use  $r(X)$  to denote a read operation on resource or element X and  $w(X)$  to denote a write on X. Using this notation, transaction T may look like this:

$$T = r(X), w(X), r(Y), r(Z), w(Z)$$

If two transactions  $T_1$  and  $T_2$  attempt to write the same resource, then one transaction must wait for the other to finish. In serializable isolation, a database engine must know which transaction to prioritize in lock acquisition. In this scenario,  $T_1$  either acquires the lock before  $T_2$  or after  $T_2$ . Assuming  $T_1$  is older than  $T_2$ , this scenario can only result in two possible outcomes: the *favorable outcome*, where  $T_1$  acquires the lock first; and the *unfavorable outcome*, where  $T_2$  acquires the lock first. In traditional Wait-Die and Wound-Wait algorithms,  $T_2$  either dies or is wounded by  $T_1$  whenever there is an unfavorable situation. This results in  $T_2$  being rolled back entirely, regardless of the number of actions. This may not pose a performance issue when both transactions consist of merely a few actions, but as the number of actions in each transaction

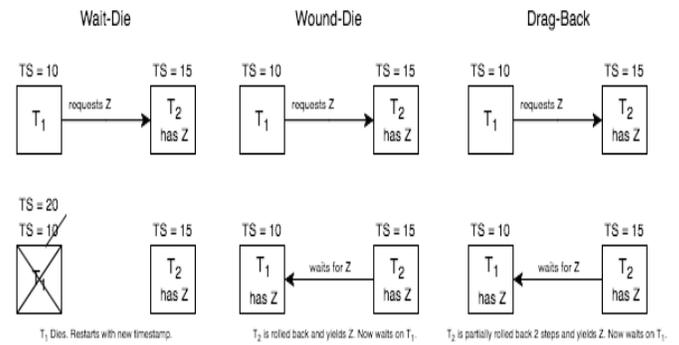


Figure 1. Wait-for graphs for Wait-Die, Wound-Die, and Drag-Back algorithms illustrating their behavior in the unfavorable outcome of Example 3.1

grows, so makes the impact of rolling back in full. Alas, it is not the same to roll back 5 actions as it is to roll back 20,000.

Using Drag-Back, this scenario would instead cause  $T_1$  to drag back  $T_2$ , rolling back some but not all the actions done by  $T_2$  up until its action prior to first writing on the shared element and surrendering the lock to  $T_1$ . In example 3.1, the transactions below conflict while writing element Z.

#### Example 3.1

$$T1 = r(X), w(X), r(Y), w(Y), w(Z) ; TS(T1) = 10$$

$$T2 = r(A), r(B), w(C), w(Z), w(D) ; TS(T2) = 15$$

Fig. 1 shows the Wait-For Graphs for Drag-Back, Wait-Die, and Wound-Wait to illustrate how each algorithm handles this situation given that  $T_2$  locked Z first in serializable isolation.

#### B. Implementation of Drag-Back

Implementing partial rollbacks in existing database engines is not easy. Existing databases have been designed under the premise that Transactions are atomic [#]. Certain logging strategies do not have enough transactional information to undo the actions of a transaction in serializable isolation. In the case of PostgreSQL, a transaction may explicitly request locks, but these cannot be released at will, as all locks can only be released by committing. Below, we address these two concerns:

##### 1) Undo actions in serialization isolation

Luckily, logging techniques that implement Log Sequence Numbers already strike undo logging in serialization isolation. It is the details of their implementation that allows for partial rollbacks. To make undoing possible in serialization isolation, every action  $a$  log entry has an associated undo action  $a^{-1}$ . An

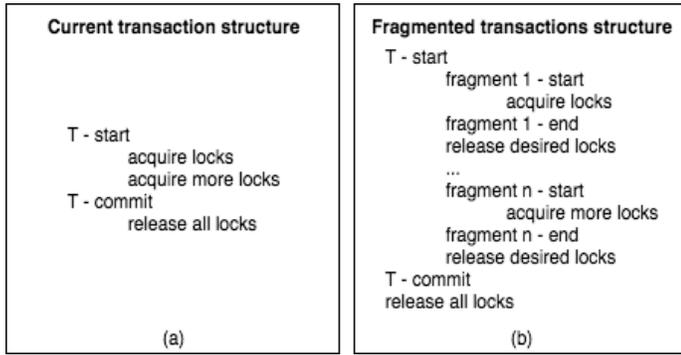


Figure 2. (a) Structure of current transaction implementations.  
 (b) Structure of proposed fragmented transactions.

action such as  $w(Z)$  would have its undo action  $w(Z)^{-1}$ , which undoes the change done to Z. A Drag-Back implementation can exploit these already-logged undo actions to perform a partial rollback.

2) *Overcoming transaction unlocking limitations*

Some database engines have imposed limitations in the way transactions unlock resources. Such is the case of PostgreSQL, which only unlocks resources after a transaction has been committed. To tackle these limitations, we introduce **fragmented transactions**. It is an abstraction of the current transactional stack. You may also think of this as a cluster of smaller transactions that make up a full transaction and share a common goal. The difference being, these transaction fragments share the same commit clause, ensuring data integrity persists between clusters. These fragmented transactions provide flexibility in lock management by also allowing unlocks in between fragments as opposed to only unlocking by committing. For logging purposes, the lifecycle of a transaction fragment consists of both a start and an end, and these are contained in a transaction.

Currently, log entries for a transaction T have a similar structure the one shown in Fig. 2 (a). The proposed structure for fragmented transactions is shown in Fig. 2 (b).

Furthermore, it is crucial for transactions to be fragmented on the fly. This is important, as it will allow for fragments to be

ended and yield the necessary locks when a dragback occurs. When the resources are available again, a new fragment is created, the locks are reacquired, and the transaction continues as intended. Fig. 3 depicts how a dragback works in practice when facing an unfavorable situation.

IV. CONCLUSION

We have shown our proposition to adopt a new deadlock avoidance scheme that can more intelligently invest time and I/O operations in serialized environments. This new scheme, which we have named Drag-Back, leverage already existing ideas in concurrency controls techniques to enable partial rollbacks. Implementing dragback functionality requires a level of flexibility in transaction processing that is not yet found in any mainstream database engine. Many engines have been designed in a way that would make this implementation especially difficult. Hence, we proposed a structure for transaction processing that would allow database engine PostgreSQL, and those similarly designed in terms of lock handling, to enable the dragback functionality.

In the future, we can expand this by adding more implementation details, as well as modeling the performance improvement of using Drag-Back over other deadlock avoidance or prevention mechanisms.

V. REFERENCES

[1] M. Kaur and H. Kaur, "Concurrency Control in Distributed Database System," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 7, 2013.  
 [2] J. Holliday and A. El Abbadi, "Distributed Deadlock Detection," *Encyclopedia of Distributed Computing*.  
 [3] R. Agrawal, M. J. Carey and L. W. Mcvoy, "The Performance of Alternative Strategies for Dealing with Deadlocks in Database Management Systems," *IEEE Transactions on Software Engineering*, Vols. SE-13, no. 12, pp. 1348-1363, December 1987.

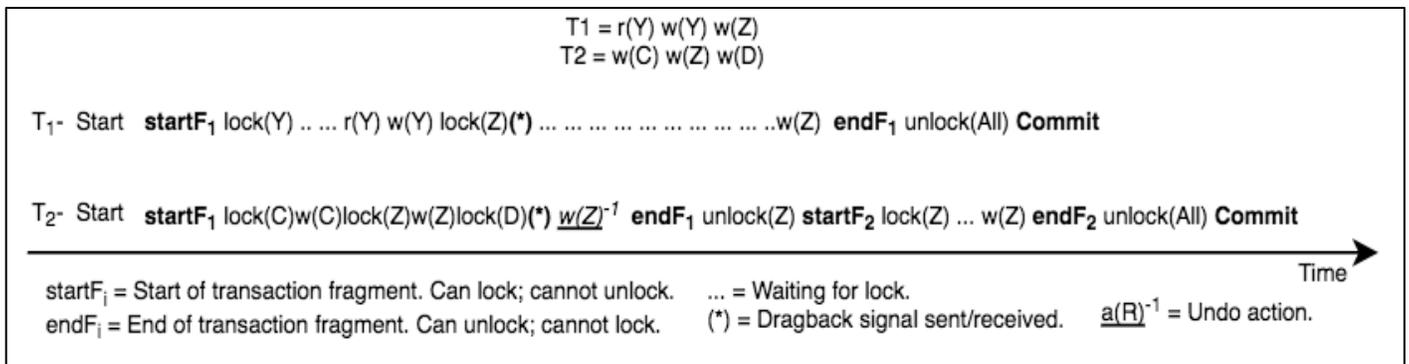


Figure 3. The timeline shows two concurrent transactions using the Drag-Back scheme.  
 T<sub>2</sub> is dragged back by T<sub>1</sub> and is fragmented as a result.

# Role of Transformational Leadership on E-Government Switching: Multi-Channel and Digital Divide

Khurram Mahmood<sup>1</sup>, Zainab Nayyar<sup>2</sup>, Hafiz Mushtaq<sup>3</sup>

Department of Software Engineering

Bahria University

Islamabad, Pakistan

km\_khuraam@yahoo.com<sup>1</sup>, zainabnayyar786@gmail.com<sup>2</sup>, h.mushtaqahmed@gmail.com<sup>3</sup>

**Abstract**— it has been observed from last 16 years that to switch from the manual government processes to Electronic government can put a huge impact in regulating the government processes. The purpose of electronic government is to use Information and communication technologies in government sector, combined with organizational change and new skills to improve government sector service delivery, democratic processes and capacity building to give strength and support the government policies. Previously government servants always work through traditional channels like front desks or telephones to fulfill their requirements. The use of traditional channels was time consuming and was only fulfill the requirements of one user at a time. These consequences lead towards a digital divide dilemma in which people moved far away from the electronic process. This study is based upon filling the gaps between E-Government and digital divide by utilizing the multi-channel services governed by transformational leaders. The study will reveal new ways of switching to electronic government through digital channels. Transformational leadership will influence the adapting phase of this concept. Data gathered through questionnaires has shown that transformational leaders along with digital channels have put a positive effect in switching from manual to electronic government processes. These aspects have increased the transparency, efficiency, accountability and security in government processes.

**Keywords**- electronic government; transformational leadership; multichannel; digital divide

## I. INTRODUCTION

The explosive entrance of technology in today's world has changed every aspect of life; way of interaction, study, work, and living. The government and its agencies are also getting effected by this drastic change and taking serious steps to make use of IT all over the world. Pakistan is part of global village and hence trying to cope with the technological change. By using IT techniques governments are reinventing themselves as they also want to transform and indulge them into e-businesses and certain e-processes for their economic fueling and better growth in the global world. The change in businesses at governmental level has provided an ease to up lift the social status of its citizens and hence that was an emergence of E-government [1]. The integrated use of

Information and Communication Technology in public sector organizations for the improvement in their processes is known as E-government. In the digital era it is a dire need of government to automate their traditional manual processes into the sophisticated automated systems to provide E-services to the citizens of the country through E-government applications.

Digital divide [2] [3] is simply the gap between people who have an easy access to the internet and those who have not. E-Government and digital divide are closely related as if government has implemented certain IT policies and defining and implementing the E-Government processes in the society but the citizens cannot make use of them efficiently then such technology fueling is of no use because people are not getting interacted with the political processes in meaningful ways.

The main purpose of the e-government is to use Information and communication technologies in government sector organizations combined with organizational change and new skills in order to improve government sector service delivery and democratic processes and capacity building to give strength and support to the government policies. The main focus of the e-government is on the delivery of better and efficient services to all citizens with transparency, accountability and in timely fashion at lower cost. It will increase the productivity of the government sector organizations and these citizen services may be accessed through a central gateway. The e-government demands to use the new multi- channels [4] i.e. internet, kiosks, mobile and Digital TV media, new infrastructure and technologies to enable transparency, accountability and openness of government organizations for the retrieval of specific information.

A composed leadership [5] is very important factor for the e-government processes implementation and coordination within and among the departments in a speedy way and reinforcement of governance objectives. The matter of fact is that e-government is a complex issue and having many dimensions. A strong leadership is required for each dimension which will well conversant with the technology to take decisions.

This study is based upon filling the gaps between E-Government and digital divide by utilizing the multi-channel services governed by transformational leaders. Multi-Channels include digital and non-digital channels; the primary focus on this study is on digital channels. The study will reveal new ways of switching to E-government by using the concept of digital channels. However transformational leadership will influence the adapting phase of this concept. The main focus of the e-services is:

- a. To ensure User Interaction with technology (Digital Divide).
- b. To make accessible and flexible services (Anytime, Anywhere).
- c. To analyze Multi-Channel Service delivery.
- d. To explore appropriate Channel Selection without limitations.
- e. To Manage Transformational Leadership approach for the implementation of e-government and to achieve transparency, accountability, effectiveness and Scalability.

## II. LITERATURE REVIEW

The authors studied [6] the software assessment model by using metrics products for e-Government in the government to business model. They found that the government to Business (G2B) model was not much used by the researchers as compared to G2C. In this regard, many software applications were developed to support G2B model but their utilization were very low due to noncompliance of needs of the users and stakeholders. It was important to develop the software application which may fulfil the requirements and expectations of the end users. The purpose of their study was to assess of this model by doing enhancement of metrics products from ISO 9126 for e-government to business model. The development was based on the assessment, functionality, reliability and effective usability with efficiency to develop quality software for the G2B model. The model not only connected the citizens but also other areas such as business, customer, and organizations or connection between public bodies. The authors were proposed the application and demonstrated through a case study. They implemented the model in local government bodies. They found that the security of the application was not up to the mark and the maturity of the software was low but functionality, operability and accuracy of the software was good enough.

A survey was conducted [7] on the impacts of E-government all over the world and identified that E-government has put a positive impact on political environment. It provided a lot of political benefits not only in the form of automation and transparency but also in operations, preventive and awareness detection by reporting and promoting ethical attitudes. She surveyed the due to E-governemnt implementation 90% reduction in corruption occurred by enhancing the security, reduction in bribery and legal logging.

The impact of e-government was examined [8] by reducing administrative corruption. The main purpose of study was to give answer of electronic government constraints on

managerial corruption. The study had found biasness between the group level and at individual level which were considered as the major cause of corruption. They found the factors which contributed in corruptions in the government namely monopoly power and discretionary power. The authors used to measure the corruption factor with the help of the following formula:

$$\text{Corruption} = \text{Monopoly} + \text{discretion} - \text{accountability}$$

The researcher prepared a questionnaire, for collecting the required data. Questionnaire stability was calculated by a reliability coefficient, where the reliability coefficient (93.5%) confirms the validity. The study sample consisted of 147 males (55.5%) and 118 females (44.5%). Based on the data collected through questionnaires it was resulted that the 88.7% population agreed that the use of e-government may help to reduce the corruption in public sector.

In [9] the e-government vision and its implementation strategy in Lebnon was studied. Futhermore they enlighten the importance of using ICT applications in Lebnon and emphasised the use of ICT in public administration, e-government solutions etc their use and implications.

Similarly [10] also conducted a survey to deal with the challenges faced by Indian Government for implementation of E-government. They suggested to implment interoperability by enforcing security and privacy policies, trust between deaprtments and common public and standardization of processes.

The relationship between E-government and curroption in developed and underdeveloped countries was examined [11]. While examining this relationship they investigated the impact of E-government on curroption and either the impact of e-government was higher in developed or underdeveloped countries. They developed the empirical model named as probit model to test and examine the results on the basis of relationships between two variables from different time periods. They took a data of seven years from (2003 to 2010) and observed the change in e-government development index with the change of curroption perception index by applying the causal method. The samples were deivided into 3 different intervals from 2003 to 2008 nd from 2008 to 2010. They found that curroption is less in those countries where e-government usage was increased. Secodly the e-government usage and its impacts were higher in developing countries. The results from probit method and regression method showed that 1% increase in e-government made 1.17% decrease in curroption. The limitations on their research were that research focused the intra country comparisons and this didn't effect each and every area of country. Secondly the conditions of curroptions affected from economic, political, historical and public sector policies of countries.

The implication of multichannel service delivery for sustainable electronic government was examined in [12]. It was important to provide multi-channel for expanding the services to the citizens especially in rural areas. They have further identified the operations carried out by the governmental bodies of China through multi-channel delivery.

From their study they concluded that the internet, mobile and electronic service channels were primarily involved in the reduction of digital divide. Based on that study the e-services like e-taxation services through multi-channel delivery were launched through establishing the call center at local administration body in China.

### III. FRAMEWORK

In this paper the conceptual frame work will represent the synthesis of literature review. It will show that how the certain variables used in the research is interrelated with each other. As [13] stated that the conceptual framework presents the specific research question that is associated with the problem statement. Problem statement highlights those issues that occurred during research. Fig 1 shows the research framework in which it is clearly seen that Transformational Leadership being an independent variable is influential on all other variables of the research.

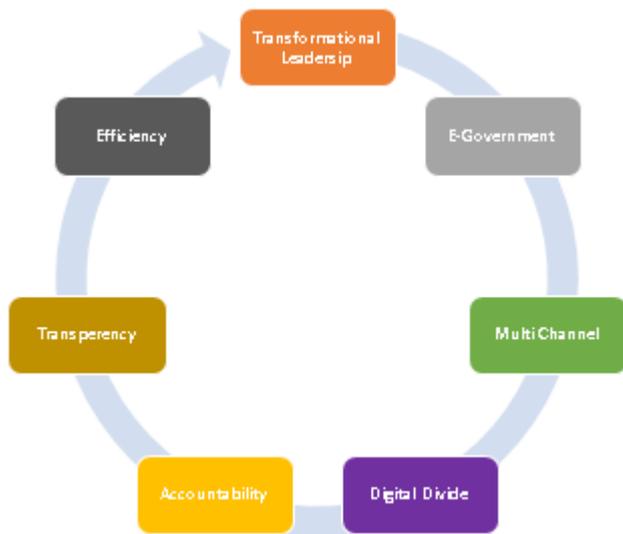


Figure 1. Research Framework

### IV. METHODOLOGY

After a thorough and extensive literature review it is inferred that although everyone has put his efforts for the successful implementation of electronic government and they are successful to some extent but no one has emphasized on the impact of transformational leadership on electronic government implementation with the utilization of digital channels and decreasing the digital divide. So the focus of our research is to proof that how and in what ways transformational leadership is making success in applying the e-government and reducing the digital divide through digital channels. For this purpose an inductive hypothesis approach was applied (Hume Fork a famous researcher called it an un justifieable approach even then he emphasized on the adoption of this approach because it always moves from specific to general and thus open more areas and dimentions of research in some specific domain) and a questionnaire was developed with a focus on four variables one independent, two dependent and one moderator. Transformational Leadership is taken as

independent variable and its direct impact is observed on e-government and an indiect effect is checked on digital divide with the help of a moderator, which is digital channel. The major focus of data gathering was from public institutions, government offices, educational institutes, hospitals, students and teachers of different age groups, posts, jobs and qualifications. Research stategy was descriptive and quantitative data is gathered and a sample size was 288. The sampling technique is probabily based which is discussed in next section.

### V. DATA ANALYSIS

The data gathered through questionnaires. Three hundred questionnaires were circulated in the market, including academia, public sector, private sector, semi government and unemployed professionals out of which 288 were recieved. The data was analysed using IBM SPSS version 24. There were five variables which were linked with the research directly or indirectly named as E-Government (EG), Accountability (A), Efficiency (E), Transperency (T) and Transformational Leadership (TL). After the computation of variables the reliablity, skewness and kurtosis of computed data is statistically analysed. According to [14] the benchmark of reliability is 0.7, skewness is  $\pm 1$  and kurtosis is  $\pm 3$ . Based on the above mentioned benchmark the reliability of data is 0.787 which is close to the benchmark value which means that the results are acceptable. Similarly the values obtained by computing the skewness and kurtosis are near to the benchmark values which validated the research data. Table 1 is showing the values of reliability, skewness and kurtosis that are obtained after the computation of data.

Table I. Data Analysis

Variables	Reliability Statistics	Skewness		Kurtosis	
	Cronbach's Alpha	Statistic	Std Error	Statistic	Std Error
All Variables	0.787	-	-	-	-
EG	-	-0.205	0.144	3.836	0.286
A	-	-1.495	0.144	4.607	0.286
E	-	-1.161	0.144	3.836	0.286
T	-	-0.855	0.144	3.393	0.286
TL	-	-0.181	0.144	-0.065	0.286

From the above mentioned figure it is clearly shown that 80% population believes that tranformational leadership is compulsory for the successful implementation of electronic government through digital channels. 60% population believes that the implementation of e-government through transformational leadership will bring transperency, 40% beleve that it will bring accountability into the system and 28% believe that efficiency in carrying out certain processes will be enhanced. Thus it is proved that tranformational leadership has a positive impact on E-government.

## VI. RESULTS

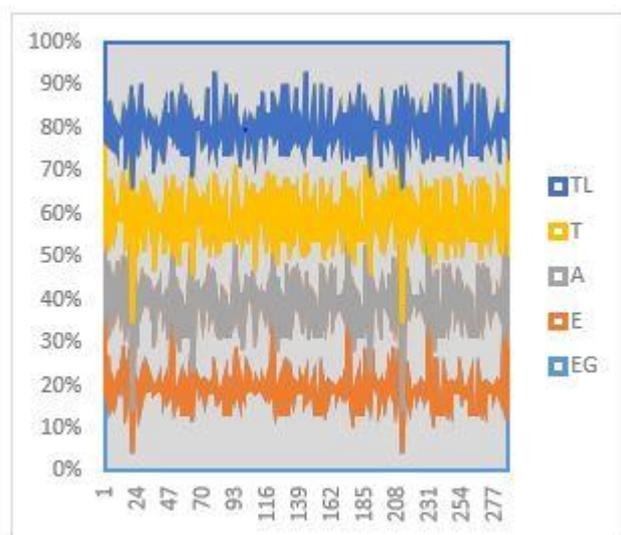


Figure 2. Graphical Results

## VII. CONCLUSION

Electronic government implementation is a very important aspect for the automation of various processes which take place among private, public and government sector. In this research the major focus was to implement electronic government through transformational leadership which will decrease the digital divide dilemma by increasing the usage of digital channels. The data analysis and results have shown that transformational leadership is a backbone to achieve transparency, accountability and efficiency by implementing electronic government system. This will bring a revolutionised change in Pakistan.

## REFERENCES

- [1] Zhu Y.Q, Kindarto.A, (2016), A garbage can model of government IT project failures in developing countries: The effects of leadership, decision structure and team competence. Published in Journal of Government Information Quarterly, Elsevier, No. 43, Sec. 4, pp 1-9.
- [2] Ebberts W.E, Jansen.M.GM, Alexander J.A.M., Deursen.V, (2016), Impact of the digital divide on e-government: Expanding from channel choice to channel usage, published in Journal of Government Information Quarterly, Elsevier, 2016, pp 1-9.
- [3] Jaeger.P.T, Thompson, K.M, (2016), E-government around the world: Lessons, challenges, and future directions. Published in Journal of Government Information Quarterly, Elsevier, Pp 389-394.
- [4] Helbig.N, Gil-García.J.R, Ferro.E, (2009) Understanding the complexity of electronic government: Implications from the digital divide literature. Published in Journal of Government Information Quarterly, Elsevier, Pp 89 -97.

[5] Elnaghi.M, Alshawi.S, (2007), A Leadership Model for e-Government Transformation. Published in Proceedings of European and Mediterranean Conference on Information Systems.

[6] Germanakos.P, Samaras.G, Christodoulou.E, (2007) Multi-channel Delivery of Services – the Road from e-Government to m-Government: Further Technological Challenges and Implications.

[7] Andrian.R, Hendradjaya.B & Sunindyo.W.D, 2016, Software Assessment Model Using Metrics Products for e-Government in The G2B Model, Fourth International Conference on Information and Communication Technologies (ICoICT), ISBN: 978-1-4673-9879-4 (c) 2016 IEEE.

[8] Athukorala.C, Perera.I & Meedeniya.D, 2016, the Impact of Transformational and Transactional Leadership Styles on Knowledge Creation in Sri Lankan Software Industry, 978-1-5090-0645-8/16/IEEE.

[9] Yin.L & Jiahong.Z, 2014, the influence of the digital divide to the electronic government affairs, International Conference on Management of e-Commerce and e-Government, 978-1-4799-6543-4.

[10] Wicking. S, 2013, Technological Innovations to identify and Reduce Corruption. Transparency International.

[11] AL-Hussaini.A, AL-Mutairi.N.N.M & Thuwaini.S.F, 2013, The Impact of Adopting E-Government on Reduce Administrative Corruption: Empirical Evidence from Kuwait's Public Sector, Academy of Contemporary Research Journal, volume 2, issue 2, pp. 31-43.

[12] Choueiri. E, Choueiri. G and Choueri. B, 2013, An Overview of E-government Strategy in Lebanon, International Arab Journal of E-Technology, volume 3, no. 1, 50-57.

[13] Hunnius.S & Schuppan.T, 2013, competency requirements for transformational e-government, 46th Hawaii International Conference on System Sciences.

[14] Chander. S and Sharmila, 2012, E-Governance: Interoperability Issues, International journal of Research in Economics and Social Sciences, volume 2, issue 7, pp. 22-37.

[15] Mistry. J and Jalal. A, 2012, An Empirical Analysis of the relationship between E-government and corruption, The International Journal of Digital Accounting and Research, volume 12, pp 145-176.

[16] Fu.Y & Xiao.K, 2012, the author's examined the Promoting Sustainable E-Government with Multichannel Service Delivery: A Case Study, Supported by the Doctoral Scientific Research Foundation of HUAT, 978-1-4673-2008-5/12/IEEE.

[17] Panda. P, Sahu. GP and Gupta. P, 2010, Promoting Transparency and Efficiency in Public Procurement: E-Procurement Initiatives by Government of India, 7th International Conference on E-government (ICEG).

[18] Kazmi S.N.A, 2010, the factors influencing e-governance Implementation: issues and challenges in Pakistan, 978-1-4244-7571-1/10/IEEE.

[19]Regoniel, P. (n.d.). CONCEPTUAL FRAMEWORK: A STEP BY STEP GUIDE ON HOW TO MAKE ONE. Retrieved January 5, 2015, from <http://simplyeducate.me/2015/01/05/conceptual-framework-guide/>.

# Shared Situational Awareness in Information Security Incident Management

Keshnee Padayachee

Institute for Science and Technology Education  
University of South Africa  
Pretoria, South Africa  
padayk@unisa.ac.za

Elias Worku

School of Computing  
University of South Africa  
Pretoria, South Africa  
elias.worku@aau.edu.et

**Abstract**—The role of participatory reportage and awareness-raising in the process of Information Security Incident Management (ISIM) has been identified as a critical approach towards proactive information security in organizations. Studies show that reporting, communication and awareness processes are fragmented and imprecise. This paper presents a review and discussion of the concept of Shared Situational Awareness towards addressing this concern. Based on this review of extant literature, this paper aims to propose a conceptual model combining the planned and participatory reporting and awareness mechanisms of Shared Situational Awareness in organizations, towards improving the process of ISIM.

**Keywords**-Situational Awareness; Information Security Incident Management

## I. INTRODUCTION

There are few guarantees that an organization will be impervious to information security incidents. Information Security Incident Management (ISIM) involves preparing for incidents, identification of resources required and preventing future incidents [1]. Hove, Tarnes, Line and Bernsmed [2] findings suggest that while organizations have incident management plans and procedures in place, their reporting procedures, in particular, are not well-established. As a result, managing information security incidents is a major challenge for organizations which requires effective protective measures not only by decision makers but also by end-users of organizations [3]. Ahmad, Hadgkiss and Ruighaver [4] found in their case study that incident response teams lack the ability to utilize organizational learning as stakeholders ignore opportunities to learn from low impact incidents. Further, they assert that the dissemination of information to stakeholders is poor as is their technical approach to reporting results in misconstructions. It is evident that organizations should be proactive in building the knowledge base of their stakeholders. Tøndel, Line and Jaatun [5] assert that learning from incidents is a way of reducing incidents in the future. It is clear that if reporting procedures are not well-established then learning from incidents will not be sufficient.

Several studies have highlighted the role of awareness as a critical factor towards preventing information security threats

[6, 7]. The extant literature depicts the disjointed information security incident management approaches – lack of standardized incident reporting and non-participatory role of stakeholders and end-users in the practice of ISIM processes [4, 5]. In a case study by He and Johnson [8] in a healthcare organization in China, they found that the distribution of ‘incident knowledge’ was uncoordinated and that incident reporting to stakeholders was challenging. It is clear that awareness plays a crucial part in incident management. Hence it is vital to create an awareness strategy that is more structured.

According to the United States Coast Guard [9], ‘Situational Awareness is the ability to identify, process, and comprehend the critical elements of information about what is happening to the team with regards to the mission. More simply, it is ‘knowing what is going on around you’. This concept appears to be a core component that is absent from information security incident management. Therefore ‘it is interesting to measure to what extent a human decision-maker is aware of the situation, i.e. has reached a certain level of situational awareness; and how well he/she manages to maintain and develop this awareness as time progresses’ [10].

Situational Awareness (SA) has been applied to various aspects of information security (see [10, 11, 12]). Webb [11] suggests using SA to aid the risk management process. As with incident management, risk assessment requires comprehensive information to support the decision-making process. Extant research on information security SA shows that there is more development within intrusion detection systems. However, the areas involving information exchange for information security SA are weak [10]. Clearly, within some areas of such incident management, the exchange of information is found wanting. The exchange of information is a complex construct as ‘it is important to communicate the right information to the right people’ [2]. For example, communicating too much information to the wrong individual can be counterproductive. It is evident that while ‘communication is a vital component of every step’ in information security incident response [13], all communication must be relevant.

Hawk [14] suggests that SA is suitable for the major domains of cyber security including ISIM. He views SA as

being a philosophical and dynamic entity unlike the typical measures in information security. Jajodia, Liu, Swarup and Wang [15] suggest that when an incident occurs there are three key questions: (1) ‘What has happened?’ (2) ‘Why did it happen?’ and (3) ‘What should I do?’ They suggest that the first two questions are the ‘core of Cyber SA’ and the last question is highly dependent on the ‘cyber SA capability’ of the organization. Bartnes, Moe and Heegaard [16] consider the value of cyber situation awareness as a means of enhancing the ISIM process. However, there are few comprehensive conceptual models to demonstrate how to achieve SA in ISIM. The Canadian Government presented a basic model, which demonstrated the need for SA in ISIM through the process of reporting and communication to ensure that all stakeholders are provided the SA ‘required to make decisions and maintain an understanding of potential impacts of incidents. Oyewole [20] also related situational awareness to incident response to the four stages: Detection, Analysis, Containment, and Improvement.

However, SA in a multi-actor activity like ISIM requires shared situation awareness. According to Endsley [17], Shared Situational Awareness (SSA) is further defined as ‘the degree to which team members have the same SA on shared SA requirements’. SSA which is more appropriate to organizational settings involves ‘a number of persons trying to form a common picture’ [18]. According to Kurapati, et al. [19], the research of SSA has ‘not dealt enough with the multi-stakeholder networks or organizations, specifically in socio-technical systems’. This study leverages this work, as ISIM is also a multi-stakeholder network, which is a socio-technical solution involving participation.

Consequently, the aim of this research is to explore the aspects of SA and promote proactive ISIM in organizations, which is the subject of Section 2. Section 3 provides a review of SA. Section 4 suggests a conceptual framework aimed at encouraging effective communication strategies with respect to incident management using SSA. The article concludes in section 5 with possible future research opportunities.

## II. BACKGROUND TO INFORMATION SECURITY INCIDENT MANAGEMENT

ISO/IEC 27035:2011, which is one of the contemporary International Organizations for Standardization of [20] information security incident management standards, is designed for large and medium-sized organizations. The standard document justifies that the efficacy of ISIM relies on the obligation to notify incidents, quality of notification, ease of use, speed and training [20]. It is clear that some of these factors relate to the quality of communication of end-users and in turn, this relates to their level of awareness. The new standard ISO/IEC 27035-1, together with ISO/IEC 27035-2, replaces ISO/IEC 27035:2011, which has been revised. The phases identified by ISO/IEC 27035-2 are described next. The first phase involves planning and preparing for incidents. The second phase involves detection and assessment of information security incidents. The third phase involves reporting the incidents so that they may be dealt with effectively. The fourth phase includes responding to the incidents to prevent, reduce and recover from incidents. The final phase is focussed on

learning from incidents and aims to institute preventative controls and make improvements to the approach. The conceptual model will leverage this approach as the standard is widely accepted in the industry.

The participation of stakeholders (managers, end-users, ICT, decision makers) is important not only in the sharing of information but also in analyzing and learning from incidents. Moreover, such studies have depicted that the impact of information security awareness among end-users is a significant factor in securing organizational informational assets [21]. Proactive information security can be achieved with increased stakeholder engagement, [22] however, this increases the complexity. Bartnes, et al. [16] found two challenges to incident management. Firstly, forming cross-function teams and then learning from past mistakes. Cross-function teams involve the collaboration of different perspectives.

In this paper, the notion of shared situational awareness is leveraged to increase stakeholder engagement.

## III. A REVIEW OF SITUATIONAL AWARENESS

Tadda and Salerno [23] define a situation as a ‘person’s worldview of a collection of activities that one is aware of at any moment in time’. Tadda and Salerno [23] argue that while a computer system may assist a person in developing and maintaining awareness, it cannot transfer this to situational awareness like a person who is the decision maker can.

Situational awareness as proposed by Endsley [24] works on three levels: **Perception** (i.e. sensitivity to environmental cues) **Comprehension** (i.e. sense-making of the information by combining, storing, interpreting and retaining) and **Projection** (i.e. forecasting about future incidents). The final level aids decision-making by anticipating future events and their implications.

According to Barford et al. [25], there are seven aspects of SA:

- Awareness of a current situation (situation perception) which includes situation recognition (knowing that an attack is occurring) and identification (i.e. type of attack) of both the source (who, what) and the target.
- Awareness of the impact of the attack (impact assessment, vulnerability analysis) includes current impact and future assessment.
- Situation tracking.
- Awareness of the adversary behaviour, trends and intent analysis.
- Awareness of how and why the current situation has been caused. This involves causality analysis (via backtracking) and forensics.
- Awareness of the trustworthiness of the collected situation awareness information and decisions. Metrics include truthfulness, completeness, and freshness.

- Predict future actions from the adversary and constrain the adversary in the future. The constraining involves understanding intent, opportunity, and capability.

SSA which is required in multi-actor networks is ‘a consensus view of a number of individual views about a specific activity or set of activities’ [23]. When a security incident occurs the individual should determine what has occurred and why it has occurred, however, the response is dependent on the shared situation awareness capability of the organization [15].

Nofi [18] suggests that building SSA involves the following criteria: First shape the individual SA within the framework of what needs to be accomplished by establishing common ground through training and experience. Secondly, establish roles of other members of the organization in order to appropriately share their awareness (mental models) using a communication protocol. Thirdly, integrate various individual mental models of the situation in order to develop a common understanding but not necessarily a single team model. This must sufficiently overlap to work towards a common goal.

There are various team processes involved e.g. communication, coordination, collaboration, etc. Most ‘attempts to understand team SA have centered on a ‘shared understanding’ of the same situation’ [26]. Hence Socio-technical implementations involving multi-actors forming a shared situation awareness are non-trivial solutions [19]. Further, there are factors, which can impede SSA. Nofi [18] comprehensively discussed the factors that degrade individual SA and SSA. The factors that impede individual SA include: fatigue, expectations, biases, stress, misperceptions, erroneous expectations, lack of experience, task overload, task underload, information shortage, information overload, information interruption, information irrelevancy, multitasking and singular foci (no concept of the ‘big picture’). The factors that impede SSA include false group mindset, peer pressure, poor communication skills, the perception of conflict, personal turbulence, insufficient training or varied skill levels and degraded operating conditions, poor communication and collaboration, computer glitches and system incompatibilities.

However, organizations can action policies and procedures to promote SSA. The United States Coast Guard [9] states that SA is a combination of the following as related to incident management:

- Identify when users deviate from procedures and policy.
- Monitor the performance of other users.
- Provide information related to incidents in advance.
- Identify potential or existing weaknesses in the system or operations.
- Continually assess and reassess the situation in relation to the information security.
- Clarify the expectations of users.

#### IV. A CONCEPTUAL MODEL OF SSA IN INCIDENT MANAGEMENT

In this section, the conceptual model of SSA of incident management is presented (see Figure 1). The phases intend to move cyclically from individual SA to SSA. The initial point for an ISIM is planning and preparation.

##### A. Plan and Prepare

The organization prepares by creating awareness of incidents and considering the type, source, and target of incidents. This includes defining security incidents, defining the responses and assigning responsibilities, training, awareness and implementing tools [5].

##### B. Individual Situational Awareness

This phase begins with the perception or detection of an incident. The individual will then have to comprehend the situation (enabled by skills, training, competence and culture; automated tools; structural factors; situational factors etc.). Individual situational awareness occurs within the construction of a mental model, where an individual is in a complex reality (where they rarely have a concept of the whole system), and they are encumbered by impediments (language, lack of knowledge etc.) [18]. The identifier of the incident will need to know the type of incident, the source (who, what) and the target. Hence, specific training on incident types is crucial to empowering SA. A strong perception of the incident is needed in order to conduct the correct decision, consequently, individual information security self-efficacy is required [27]. Nofi [18] suggests that some people who are shaped by structural factors (training, experience etc.) and situational factors (stress, complexity etc.) are better ‘noticers’ than others. The individual will need to project the implications of the incident to the future. An individual uses his/her internal heuristics to form a conceptual map of the incident and will then have to communicate it via a preliminary report [18]. This stage is incredibly important as insufficient information could lead to poor decisions but on the other hand, too much information can be overwhelming [2]. Some individuals such as end-users may not be able to identify the source and target of an incident and may use subjective judgement. However, the aim in individual SA is to enable the individual to identify, comprehend and project with the support of additional enablers such as situational, structural and automated tools. This is supported by the *Lessons Learnt* component in the SSA involving the receiving of feedback, which will be conveyed as lessons to the end-users.

##### C. Shared Situational Awareness for Multi-Actor Networks

The next phase will involve sharing the situational awareness of the incident amongst all key role players (managers, end-users etc.) including the information security incident response team (ISIRT). The enablers of SSA are shared understanding, trust, coordination, common ground and commitment. The process involves the core elements of interaction, visualization, synchronization, and sense-making. These core elements could be cyclic rather than sequential. *Interaction* involves knowledge sharing relative to the incident in order to create a shared mental model of the situation using

communication tools. According to Franke and Brynielsson [9], the human-computer interaction could also be used to understand the basis of the incident for understanding the situation in more depth. *Visualization* involves using visual tools to map the incident. Visualization is a vital component for SAA, for example, visualizing the big data attacking graphs or cognitive task analyses [10] can help form a shared mental model. *Synchronization* involves integrating the mental models of individuals, in order to create common ground towards developing a single understanding of the incident. This will involve collaboration tools. *Sense-making* involves creating a shared understanding of the collective information in order to act on the information. It is important to make sense of a situation before being able to project into the future. This will involve planning and scheduling tools and decision support tools. Sense-making can involve intelligent systems that automatically ‘fuse massive data into succinct meanings’, process meanings, ‘achieve insights’ and hypotheses, access intuition and present information in meaningful ways [28]. Sense-making helps to select a frame from multiple frames that best fits the situation [29]. A frame is a mental model that identifies gaps and makes predictions [29]. The final component is *Shared Reporting*, which communicates the full implications of the incident to the ISIRT.

**D. Information Security Incident Response Team Situational Awareness**

While it is important to involve as many stakeholders as possible during the shared situational awareness phase, the ISIRT are solely responsible for the technical aspects with respect to incident management (i.e. assessment, decision-making, and response). Their final role would be to compile a training programme into a lessons learnt package. The ISIRT work at a multilevel that is both individual and collaborative where they have to utilize information-sharing and collaborative skills [30]. The core elements of this phase involve assessment, decisions, response, and lessons learnt. *Assessment* includes assessing the trustworthiness of the report

of the previous phase; conducting an impact assessment, vulnerability analysis, situation tracking, trends and intent analysis, causality analysis and forensics. The *Assessment* element also involves projecting future incidents. This involves defining the details contained in the Shared Incident Report, confirming and classifying incidents [5]. The *Decisions* and *Response* elements involve planning and implementing the actions to constrain the incident. The *Lessons Learnt* component involves learning and building knowledge to minimize future incidents. The feedback received will be conveyed as lessons to the end-users. This incident must then become part of their situation perception. Tøndel, et al. [5] suggest assessing and evaluating end-users after every incident, followed by dissemination of incident information. This will involve usage metrics for learning effects and tuning of technical measures.

**E. Role-based Situational Awareness**

The final phase will promote the enablers of individual SA and will involve using a role-based mechanism where individuals will receive information and training based on their roles and responsibilities. This is aligned with enabling SA by reducing information overload, which increases situational awareness. This feeds back into the *Plan and Prepare* phase for the next incident.

**V. CONCLUSION**

In this paper the notion of iterating from individual SA to SSA based on end-users’ experiences is used for comprehending and projecting future outcomes. To this end, a conceptual model for multi-actor incident management to increase SSA was proposed. The model intends to increase SSA in ISIM which will improve the reporting and communication protocols that were found wanting by extant research. Future research will involve prototyping the model using design science research.

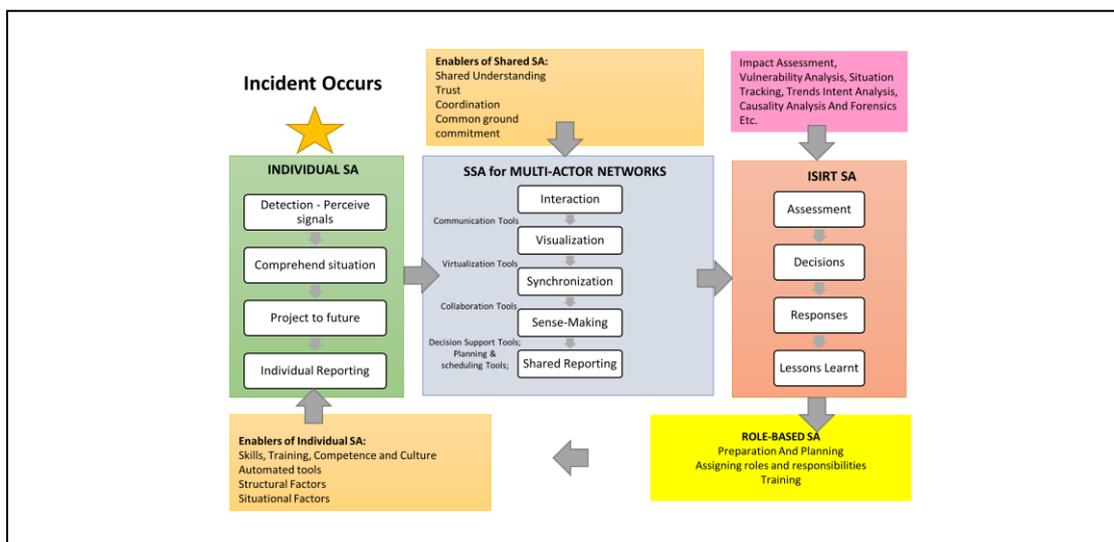


Figure 1. A Conceptual Model for Shared Situational Awareness for Information Security Incident Management

REFERENCES

- [1] H. Susanto, M. N. Almunawar, and Y. C. Tuan, "Information security management system standards: A comparative study of the big five," *International Journal of Electrical Computer Sciences IJECSIJENS*, vol. 11, pp. 23-29, 2011.
- [2] C. Hove, M. Tarnes, M. B. Line, and K. Bernsmed, "Information security incident management: identified practice in large organizations," in *IT Security Incident Management & IT Forensics (IMF), 2014 Eighth International Conference on*, 2014, pp. 27-46.
- [3] Riden. (2006). *Responding to security incidents on a large academic network*. Available: [http://www.infosecwriters.com/Papers/JRiden\\_case\\_study.pdf](http://www.infosecwriters.com/Papers/JRiden_case_study.pdf). Last accessed 13 September 2017
- [4] A. Ahmad, J. Hadgkiss, and A. B. Ruighaver, "Incident response teams—Challenges in supporting the organisational security function," *Computers & Security*, vol. 31, pp. 643-652, 2012.
- [5] I. A. Tøndel, M. B. Line, and M. G. Jaatun, "Information security incident management: Current practice as reported in the literature," *Computers & Security*, vol. 45, pp. 42-57, 9// 2014.
- [6] E. Humphreys, "Information security management standards: Compliance, governance and risk management," *information security technical report*, vol. 13, pp. 247-255, 2008.
- [7] B. Von Solms and R. Von Solms, "The 10 deadly sins of information security management," *Computers & Security*, vol. 23, pp. 371-376, 2004.
- [8] Y. He and C. Johnson, "Challenges of information security incident learning: an industrial case study in a Chinese healthcare organization," *Informatics for Health and Social Care*, pp. 1-16, 2017.
- [9] United States Coast Guard. (1989). *SITUATIONAL AWARENESS*. Available: <https://www.uscg.mil/auxiliary/training/tct/chap5.pdf>. Last accessed 13 September 2017
- [10] U. Franke and J. Brynielsson, "Cyber situational awareness—a systematic review of the literature," *Computers & Security*, vol. 46, pp. 18-31, 2014.
- [11] J. Webb, A. Ahmad, S. B. Maynard, and G. Shanks, "A situation awareness model for information security risk management," *Computers & security*, vol. 44, pp. 1-15, 2014.
- [12] W. Yu, G. Xu, Z. Chen, and P. Moulema, "A cloud computing based architecture for cyber security situation awareness," in *Communications and Network Security (CNS), 2013 IEEE Conference on*, 2013, pp. 488-492.
- [13] Cyber Security Coalition. (2015). *Cyber Security Incident Management Guide*. Available: <https://www.agoria.be/upload/agoriav3/Cyber-Security-Incident-Management-Guide-2015.pdf>. Last accessed 13 September 2017
- [14] R. Hawk. (2015). *Situational Awareness in Cyber Security*. Available: <https://www.alienvault.com/blogs/security-essentials/situational-awareness-in-cyber-security>. Last accessed 13 September 2017
- [15] S. Jajodia, P. Liu, V. Swarup, and C. Wang, *Cyber situational awareness* vol. 14: Springer, 2010.
- [16] M. Bartnes, N. B. Moe, and P. E. Heegaard, "The future of information security incident management training: A case study of electrical power companies," *Computers & Security*, vol. 61, pp. 32-45, 2016/08/01/ 2016.
- [17] M. R. Endsley, "A model of inter-and intrateam situation awareness: Implications for design, training, and measurement," *New trends in cooperative activities: Understanding system dynamics in complex environments*, pp. 46-67, 2001.
- [18] A. A. Nofi, "Defining and measuring shared situational awareness," CENTER FOR NAVAL ANALYSES ALEXANDRIA VA2000.
- [19] S. Kurapati, G. Kolfschoten, A. Verbraeck, H. Drachler, M. Specht, and F. Brazier, "A theoretical framework for shared situational awareness in sociotechnical systems," in *Proc. 2nd Workshop on Awareness and Reflection in Technology-Enhanced Learning*, 2012, pp. 47-53.
- [20] ISO/IEC 27035, "Information Technology – Security Techniques – Information Security Incident Management," ed. Geneva, 2011.
- [21] J. M. Stanton, K. R. Stam, P. Mastrangelo, and J. Jolton, "Analysis of end user security behaviors," *Computers & security*, vol. 24, pp. 124-133, 2005.
- [22] S. Metzger, W. Hommel, and H. Reiser, "Integrated Security Incident Management—Concepts and Real-World Experiences," in *2011 Sixth International Conference On IT Security Incident Management and IT Forensics (IMF)*, 2011, pp. 107-121.
- [23] G. P. Tadda and J. S. Salerno, "Overview of cyber situation awareness," *Cyber situational awareness*, pp. 15-35, 2010.
- [24] M. R. Endsley and D. Garland, "Theoretical underpinnings of situation awareness: A critical review," *Situation awareness analysis and measurement*, pp. 3-32, 2000.
- [25] P. Barford, M. Dacier, T. G. Dietterich, M. Fredrikson, J. T. Giffin, S. Jajodia, et al., "Cyber SA: Situational Awareness for Cyber Defense," *Cyber Situational Awareness*, vol. 46, pp. 3-13, 2010.
- [26] R. Klemke, S. Kurapati, H. Lukosch, and M. Specht, "Transferring an educational board game to a multi-user mobile learning game to increase shared situational awareness," in *International Conference on Learning and Collaboration Technologies*, 2015, pp. 583-594.
- [27] T. Oyewole. (2016). *Application of Situation Awareness in Incident Response*. *ISACA Journal* 3. Available: <https://www.isaca.org/Journal/archives/2016/volume-3/Pages/application-of-situation-awareness-in-incident-response.aspx>
- [28] G. Klein, B. Moon, and R. R. Hoffman, "Making sense of sensemaking 1: Alternative perspectives," *IEEE intelligent systems*, vol. 21, pp. 70-73, 2006.
- [29] M. D. Howard, R. Bhattacharyya, S. E. Chelian, M. E. Phillips, P. K. Pilly, M. D. Ziegler, et al., "The neural basis of decision-making during sensemaking: Implications for human-system interaction," in *2015 IEEE Aerospace Conference*, 2015, pp. 1-16.
- [30] T. R. Chen, D. B. Shore, S. J. Zaccaro, R. S. Dalal, L. E. Tetrack, and A. K. Gorab, "An organizational psychology perspective to examining computer security incident response teams," *IEEE Security & Privacy*, vol. 12, pp. 61-67, 2014.

# A Collaborative System for Corporate Performance Evaluation using Gamification and the Learning Vectors Model

Michelle G. Cacais

Instituto de Educação, Ciência e Tecnologia do Ceara (IFCE)  
Programa de Pós-Graduação em Ciência da Computação  
Fortaleza, Brazil 60040-215  
Email: michellecacais@gmail.com

Gilvandenys L. Sales

Instituto de Educação, Ciência e Tecnologia do Ceara (IFCE)  
Programa de Pós-Graduação em Ciência da Computação  
Fortaleza, Brazil 60040-215  
Email: denyssales@gmail.com

**Abstract**—Corporate strategies should contribute to the achievement of the objectives of companies, and, consequently, guarantee their sustainability. In order to keep the employees engaged and committed to their own income, we propose Process Planning and Institutional Evaluation (PIPA), a gamified system that uses the Learning Vector Model for corporate performance evaluation and follow-up of tasks. The system can be applied in entities and companies with the purpose of socializing, motivating or promoting the interaction of users. It was verified through field research, that PIPA improved the performance of the evaluated team and motivated the professionals. The evaluation using the Learning Vector Model helped in the monitoring and the progress of the evaluated ones, since besides accompanying their own income through constant verification and feedback of the supervisors. They also could interact more with the work team.

## I. INTRODUCTION

Human Capital consists of the skills and competencies that favor the performance of a work. It is one of the main instruments of resource generation, and the value of this is materialized in the contributions that each one brings to the enterprise. This value can be increased depending on the stimuli offered by the company, for example, a new training or learning a new technology that will help in the processes performed by the employees. It can also be diminished, as in situations where people are induced to a process of obsolescence or when the work environment leads to widespread demotivation of the teams [6]. Thus, it is important to keep Human Capital at a satisfactory level of performance, what is fundamental to the achievement of organizational objectives.

The evaluation of corporate performance can be a great ally to empower, motivate and reward employees. Through this practice, it is possible to focus employee activities properly, align individual goals with the organizational ones, joining performance at work with the medium-term objectives and strategies of the corporation and to maximize the potential of individuals and the team to benefit the organization. The evaluation allows the identification of potentialities and make a career progression, and also, alert those who are not doing well.

The Learning Vectors Model (LV) is an evaluation methodology which uses geometric representations of the performance, making possible the classification in a qualitative and quantitative way [7]. It was originally created to help teachers, tutors and students of distance education, allowing a semi-automatic evaluation. The evaluated ones have continuous feedback and grades are given from an association between a horizontal and a vertical component representing the positive and negative contributions of the interactions between the class.

Besides the constant evaluation, the adoption of a collaborative system could also improve the performance of the teams. This way, we had the idea to create a system for corporate evaluation and follow-up of tasks that uses the LV Model. This have been chosen for being a dynamic tool, since it allows the semi-automatic evaluation. So, there is no time loss for those who already have a running routine and want to facilitate such tasks as employee evaluations. On the other hand, employees will have access to constant feedback, an important resource for those who want to make a career progression and know which points to improve. In addition, it is very intuitive to use, as it is based in the Likert scale of appreciation, what does not require much learning to use and it is easy to remember.

The motivating factor of the program is the gamification, a technique to streamline the learning process or training, and make tedious or repetitive tasks more enjoyable [11]. Consists in the use of game elements in varied contexts, for purposes that are not exclusive of entertainment. Through this practice, you can improve employee performance, promote socialization, and generate a sense of achievement desired by people working for their goals and those of the company. When well used, gamification is an ally that gives positive returns by keeping people focused and at the same time entertained.

This way, we aim to provide a free software to help companies with the corporate evaluation and the follow-up of tasks. PIPA allows monitoring of individual or group projects, promoting more interaction between the members of the team. Here we present the system and the results obtained with our

field research in a real company.

## II. LEARNING VECTORS MODEL

The Learning Vectors Model is a qualitative-quantitative methodology of non-linear evaluation, that allows constant monitoring for the evaluated ones. LV Model focuses on the interaction of the group with the virtual learning environment, mainly in the use of distance activities. The LV Model is based in dynamic systems and uses vectors and numerical equations in a two-dimensional way, determined by projections on the Cartesian axis. These values represent the bipolarity between the qualitative and quantitative dimensions [7]. Visually, the vector is indicated by an arrow that rotates counterclockwise and allows users to check how his learning is going. Figure 1 shows the vectors of the LV Model.

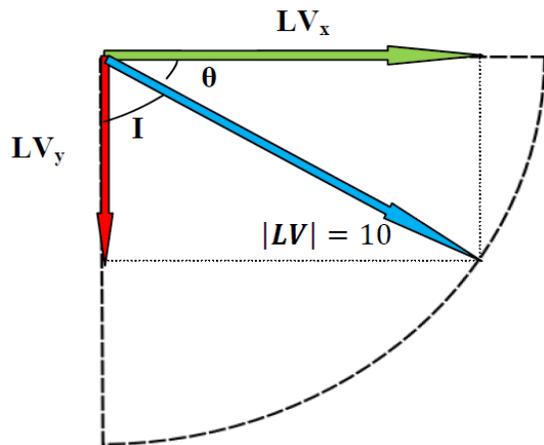


Figure 1: Graphic representation of the vectors of LV Model [7, p. 88]

Through the vectors, the scores obtained by synchronous and asynchronous activities, are presented graphically by means of a vector and numerical values. The vectors have a fixed module of 10 units, with directions starting at the angle  $\theta = -90$  (lowest score), until the value of  $\theta = 0$  (highest score). The geometric representation consists of the axes  $LV_x$ , horizontal projection, and  $LV_y$ , vertical projection. These are related to positivity and negativity factors, respectively. The module of the component  $LV_x$  is the score obtained by the user in different fulfilled activities. Equation 1 shows how the value of  $LV_x$  is obtained.

$$LV_x = 10 \times \cos[(-12\alpha + I)] \quad (1)$$

The calculation of component  $LV_x$  takes into account the variable  $\alpha$ , value of the standard learning step, stipulated in 7,5, and the variable  $I$ , for the total angular variation ( $\Delta_T$ ) of the vector and function of the  $\alpha$  angle. The movement of the vector depends on the variable  $I$  and  $\alpha$ . The value of  $\alpha$  was determined as 7,5 for dividing the angle of 90 in equal parts. The vertical component  $LV_y$ , related to the negativity of the performance of the user, is determined by the Equation 2.

$$LV_y = 10 \times \sin[(-12 + I)] \quad (2)$$

From the positivity and negativity factors, similar to bipolar dimensions, it is defined a non-linear pedagogical metric called  $\beta$  Factor, which indicates the level of the user. This factor is related to the qualitative nature of the LV Model, and involves the good and bad results of the contributions of students in tasks performed in the virtual environment. Equation 3 shows the calculation of the  $\beta$  Factor.

$$\beta = \frac{Positivity}{Negativity} \quad (3)$$

The evaluators have access to an iconic representation associated with the scale of qualitative statements (Likert scale) to assign value to the evaluated ones' performance. The LV Icons, similar to emojis, can transmit sentimental aspects, since they are graphical representations of emotions. They present categories associated to the degree of what have been presented, as well as the level of interaction with their peers. Table 3 presents the LV Icons, what they represent and the value associated with them. Table ?? shows the values of each LV Icon that will make part of the equations.

TABLE I: Description of LV Icons

Icon	Description	SC*
	<b>Very good:</b> in-depth reflections	4
	<b>Good:</b> good reflections	3
	<b>Regular:</b> medium reflections	2
	<b>Weak:</b> empty content reflections	1
	<b>Unsatisfactory:</b> person who assumes a passive position	0
	<b>Neutral:</b> messages or files that do not bring contribution	-

\* Step Coefficient

Adapted from [7, p. 113]

The values of the LV Icons are called Step Coefficient, and they are important for the calculation of the quantitative grade, influencing the the variable  $I$ . The neutral icon does not increment the note, but serves to compute the presence of the person in the activity. It is also important to point out that each type of LV activity has a different calculation that better fits this one.

The LV Model was evaluated and validated, when it was verified the effectiveness of use as a continuous process of formation. The degree of satisfaction of teachers, tutors and students was verified too. The use of the methodology proposed in this system collaborates to differentiated actions

throughout the teaching and learning process. The implementation and testing of the LV Model proves that it is possible to construct a model that relates numeric values with the subjective scale of mentions, and at the same time, presenting qualitative and quantitative values.

### III. RELATED WORK

Gamification is used in different areas to influence the behavior of people. The work seen in Steffens et al [10] proposes a framework applied to the context of software engineering, to verify the most common issues in the collaboration of the development team and how to apply game elements to overcome this issues. Other work that involves gamification, but applied to the area of enterprise environment, is Stanculescu et al [9]. These paper brings an experiment performed in a multinational company to verify how gamification could influence the performance of the employees. The results provided evidences of how a gamified experience affects the learning and the social behavior of the team in an enterprise context. The level of engagement achieved by them was high in general.

The design and development of systems for corporate evaluation is also a trend. The work seen in Ng et al [5] is one example of a system faced to these topic. The authors explain how they developed a Key Performance Indicators (KPI) system for Maintenance, Repair and Operation (MRO), in order to resolve corporate problems. The results have shown that these program enables the users to conduct analysis and develop a long-term strategy. Another work about corporate evaluation can be seen in Falcinelli et al [4], which evaluates the devices for e-learning aimed at compulsory training. The contributions of this article is the introduction of a UNI-like standard for e-learning, bringing more quality to it.

The LV Model was used in works, as Sales et al [8] and Gonçalves et al [3]. The first one addresses possible impacts that changes in the management of learning could lead to the distance education. It was compared two methodologies: spreadsheet and the LV Model. It was noticed that, with the constant monitoring of the activities through the LV Model, it is possible to infer changes in the avoidance rate. The second work presents an application of the Instruction by Peers Methodology, associated with the LV Model, to the Forum "Questions and Answers" of the Virtual Learning Environment Moodle. The authors presented a computational conception design of non-linear evaluations, as a result.

TABLE II: Summary of related work and contribution.

Reference	Gamification	Valuation	LV Model
Steffens et al [10]	✓		
Stanculescu et al [9]	✓	✓	
Ng et al [5]		✓	
Falcinelli et al [4]		✓	
Sales et al [8]			✓
Gonçalves et al [3]			✓
PIPA	✓	✓	✓

As seen in Table II, gamification has been used in the corporate environment with success, including to make per-

formance evaluation of employees. Systems to make corporate evaluation are also been developed to make it easier the measurement of the actual performance and make progressions to improve the employee income. LV Model is addressed in an educational context, and has showed itself as an alternative way for qualitative and quantitative evaluation. PIPA is the only gamified system faced to corporate evaluation that uses the LV Model.

### IV. PIPA: DESIGN AND DEVELOPMENT

The development of PIPA was planned with several studies of which technologies to use and what the final appearance would look. Firstly, technical details were considered, such as the language that would be used, the extensions and libraries that would help the system functionalities and which ones would better fit the proposal. The system was built using the PHP programming language, the script language JavaScript, HTML markup language and the CSS styling language. PHP was initially chosen because it was the same language used for programming the LV Model. In addition, MySQL database was used and some libraries, as Bootstrap<sup>1</sup>, to make the application responsive, and RGraph<sup>2</sup>, to generate dynamic charts.

The architecture of the program follows the model known as three-tiered application (3-Tier), architectural style in which the system is organized into three main layers [2]. The layers used for the architecture of PIPA are the user interface, the presentation screens that will make the interaction with people; the business layer, with the logic of the application; and the database layer, which will store score information, personal data, and employee appraisal history.

Regarding the classes of the application, a hierarchy was considered. They are three user profiles: administrator, manager and user. The users can be enrolled in activities by managers and administrators, and can access the notes assigned to them, update the status of tasks and send files and comments. Managers are responsible for registering activities, recruiting users for them and doing the evaluation at the end. They will have access to the scores of the users. The administrator is the profile that will handle the system. The manager can do what a manager does, but it is also possible to register people and classify them into any of these categories.

The design of PIPA was inspired by the minimalist design, which has the least graphic resources and texts. It does not mean that it has little content, but rather broadens the essence of what is really important, to the point of making everything else expendable before the very focus of creation [?]. The reason for choosing minimalist themes is to focus the attention of the users on the content, that is, tasks, and to divert the focus of graphic elements. The home page of the system can be seen in Figure 2. This screen contains only the dialog boxes for login, with the option of password recovery. Just like the home screen, the others also try to follow the minimalist design.

<sup>1</sup><http://getbootstrap.com/>

<sup>2</sup><http://www.rgraph.net/>

As soon as the user logs on to the system, he / she will see the last evaluation and the game elements, as ranking and badges. Figure 3 shows the user home. Other information on the home screen is the tasks still waiting to be completed or updated. In the navigation bar it is possible to see the options to view all projects, including those which have been completed and those that are still open, the general projects, and the project register, available only for manager profiles. Figures 4 and 5 show the screens of the projects of the user and general projects, respectively.

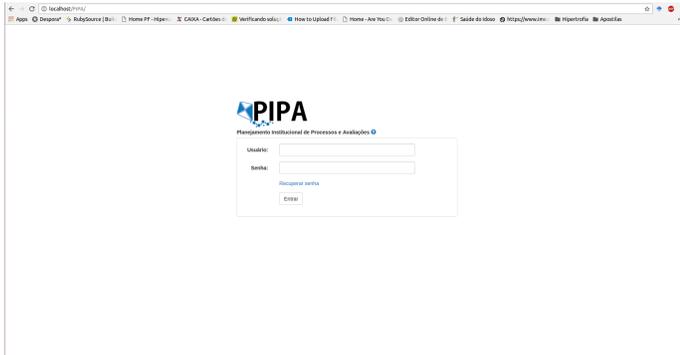


Figure 2: Home page of PIPA

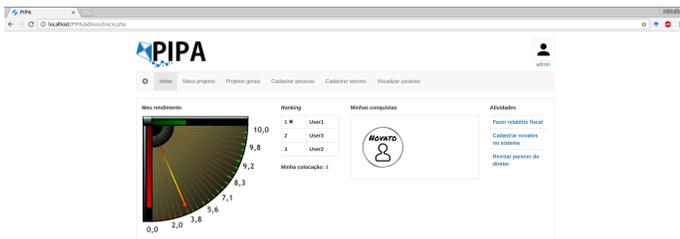


Figure 3: Page of the user of PIPA

The page of projects shows the tasks of the person is registered as a member of a team or individually. This screen shows the number, unique code of each registered task, name, type, priority and whether it is completed or not. This information is previously registered by the manager who opened the task in the system. The types are related to the sectors to which it is related, such as administrative or financial. In the projects tab, it is possible to have an overview of everything what is being produced in the company, not only the processes assigned to the user. This page is important for monitoring the projects of the colleagues. Authors are the people who registered the activity in the system, it means, users with a manager profile. 'Assigned' is related to the person who is intended the task.

The system has been made available to the general community, especially to the target audience of developers and people who want to use PIPA to manage their activities in the

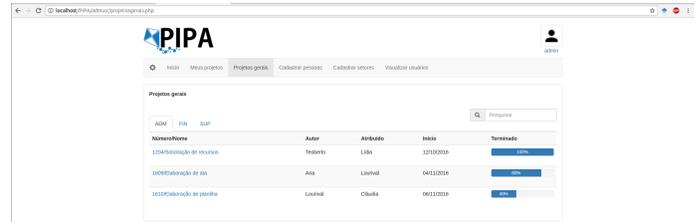


Figure 4: Page of projects of the user

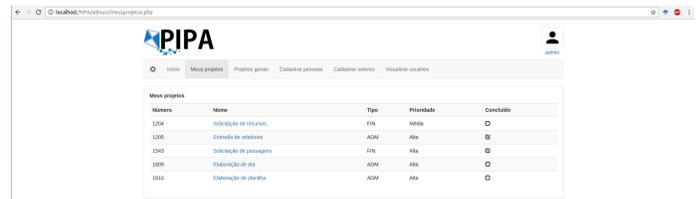


Figure 5: Page of general projects

company. The repository used to store the code, as well as do version control, was GitHub<sup>3</sup>, and can be accessed in this link: <https://github.com/Cacais/PIPA>.

### A. Learning Vectors applied to PIPA

The score obtained in the system will be made from the Learning Vectors Model adapted to PIPA. Like this evaluation proposal, PIPA point counting system aims at non-linearity. The scores will be given by the same formulas of the LV Model to calculate the positivity and negativity, and the values for  $LV_X$ , which indicates the positivity obtained by the person, are used for the point system. For better understanding, the points will be called  $P$ . It is possible to observe the relations of the equations to the values obtained in Equation 4. The value of  $P$  initially is zero and then the actual  $P$  count plus the value obtained by the calculation of the previous value of  $P$ .

$$P = P + (10 \times \cos[(-12\alpha + I)]) \quad (4)$$

Scores are obtained by calculating  $LV_X$ , while total points are the sum of all scores obtained. The ranking depends on the scores, but consists only of the sum in the week, since it is updated weekly. This can also help in accomplishing more tasks, since the more evaluations is achieved, the more

<sup>3</sup><https://github.com/>

punctuation and more chances of joining the first positions of the ranking are reached. It is possible to access the average of those evaluated at the time of reporting, but the scores that will appear on the system home page will be the last assessment.

The  $\beta$  Factor is indicative of whether or not to level up. As in LV Model, the  $\beta$  Factor is calculated by dividing positivity by negativity. At the beginning, passing the phase is easier, but with each level reached, the difficulty increases. The levels are numerically divided, for example, level 1, 2, 3, ..., n. There is no limit number. To increase the level, the score obtained must always be the double of the previous one. The calculation of the number needed to reach levels depends on other variables, as unlock badges and keep reasonable valuations.

The LV Model was chosen as a mechanism to support dynamic evaluations and to provide a constant feedback to those involved in the educational process, as well as to verify negativity and positivity. We aim to unite valuation with gamification and encourage employees to adopt positions that improve their performance, but in a fun way, trying to change the paradigm of currently available performance appraisal systems. Through these resources, it is assumed that, having in hand the mechanisms necessary to know their own performance, those evaluated will identify factors that hinder and those that improve performance and take measures to modify or maintain the progress of the task resolutions.

## V. PRELIMINARY RESULTS

For the field research, we choose a company of representation of seals and enclosures for the mineral water and beverages segment. It consists of a microenterprise with 7 employees, a receptionist, an accountant, two proprietary partners and three representatives working externally, located in the Passar neighborhood of the city of Fortaleza, Cear. The test period was a little more than a month in the first quarter of 2017. A week before the start of the tests, a meeting was held with the owners and employees for explanations of how to use the system, and then began to use our system. In this context, the system could be an ally to control the activities of everyone who works in person, but especially those who work outside.

### A. Methodology

We scheduled a day for installation of the program and explanation of how to use it for all employees. On that day, employees who work externally were also present, and it was possible to teach everyone. A local server was used inside the company and was explained individually as access, register tasks, evaluate and monitor their performance through the system. They all collaborated and seemed excited to use it. After the training, the contacts were left in case any of the employees had any questions or to solve technical problems.

The system was used for a month, more precisely thirty-six days. At the end of the process, another meeting was scheduled for delivery of software evaluation forms and feedback. Two forms were applied: one for self-evaluation and another for supervisors to evaluate the employee's income.

The questionnaires had objectives and subjective questions. It was made clear to the participants that the procedure would be anonymous, with the sole purpose of raising data for an academic research, and that colleagues would not have access to the answers. At the end of this period, the forms were collected.

The forms contained questions about the adequacy of the software to the environment, employee satisfaction and performance improvement. Thus, the evaluation models were divided into two parts: one for self-evaluation, aimed at employees in general, including the owners; and one for employee evaluation. The purpose of this would be to analyze whether employees perceived changes in their income and whether the bosses noticed it. All answered the self-assessment questionnaires, while only the owners answered the evaluation of the results of their collaborators.

### B. Results

The results of the questionnaire are organized together with the questions to which they refer. In general, the system improved employee performance. When asked about this, they said they were able to pay more attention to how the tasks were going and that the employees were more focused. As for the ease of evaluation, they felt a bit of difficulty at first, but with little time have been able to evaluate. The results can be checked as follow.

#### 1) Questionnaire of the evaluators:

- 1) Did the employees obtain improvement in income through PIPA?
- 2) The level of difficulty of evaluating employees by PIPA was:
- 3) Did the monitoring of the activity of the employees by PIPA helped the company?
- 4) Did you use to evaluate employees beforehand by any other method?

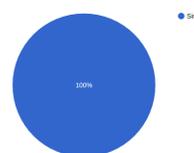


Figure 6: Question 1

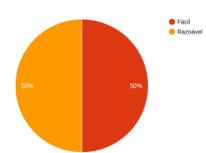


Figure 7: Question 2

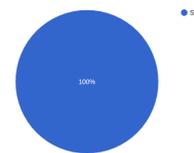


Figure 8: Question 3

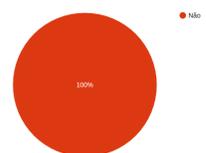


Figure 9: Question 5

#### 2) Self-assessment questionnaire:

- 1) In a general way, how do you evaluate your experience with PIPA software?

- 2) In your opinion, do your performance improved with PIPA?
- 3) Do you think your interaction with others improved using PIPA?
- 4) Was the evaluation by PIPA a positive, negative or neutral factor?

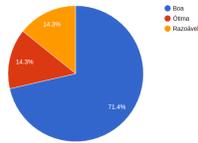


Figure 10: Answers for the first question

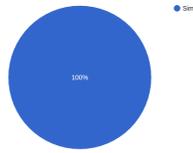


Figure 11: Answers for the second question

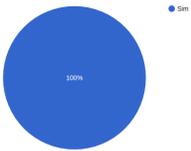


Figure 12: Answers for the third question

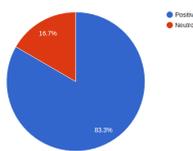


Figure 13: Answers for the fourth question

The questions were elaborated to verify the length of the work objectives. From the responses of the users, it was possible to conclude that the system was able to motivate and interact among the members of the company. Also, the evaluation enabled the evaluators to be monitored and evaluated.

## VI. CONCLUSION

PIPA come out as a proposal of a free code system that aims to stimulate the employees through monitoring of processes and evaluations. This technique proved efficient in engaging users to adopt a certain posture, so employees reach the levels of high performance teams.

Testing process was very helpful in uncovering the strengths and weaknesses of the system as well as collecting suggestions for improvements. The objectives we had at the beginning were verified and it can be concluded that, for this evaluation in the field, the system was carried out in a profitable way, fulfilling the established goals. According to the answers of the questionnaires applied, it is presumed that the one planned in the beginning was fulfilled.

It is expected that PIPA software encourages the improvement of employee performance, which will be stimulated through gaming elements to have more productivity, since mechanisms will be added for constant checking of evaluations and feedback. It is also intended to improve the monitoring of what is being produced in the company, in order to identify harmful factors, and thus be able to take preventive measures. This way, both evaluators and evaluators gain in the process. In future works, it is hoped to improve the visual aspect of PIPA, making it more attractive and intuitive to the public and

improving the user experience in using it. It is also one of the future objectives to use Fuzzy Logic for PIP Model LV.

## REFERENCES

- [1] CARRION, Wellington. "Design para web designers: principios do design para web." Rio de Janeiro (2008).
- [2] Bourque, Pierre, and Richard E. Fairley. Guide to the software engineering body of knowledge (SWEBOK (R)): Version 3.0. IEEE Computer Society Press, 2014.
- [3] Gonçalves, Alexandra Joca, Joana Laysa Lima Cunha, and Gilvandenys Leite Sales. "Concepção do Frum PR LV: avaliação formativa da aprendizagem." RENOTE 14.1.
- [4] Falcinelli, Floriana, et al. "Evaluation of an e-learning device for workers compulsory training: an example of collaboration between university and company." Research on Education and Media 8.2 (2016): 26-32.
- [5] Ng, K. K. H., M. H. M. Tang, and C. K. M. Lee. "Design and development of a performance evaluation system for the aircraft maintenance industry." Industrial Engineering and Engineering Management (IEEM), 2015 IEEE International Conference on. IEEE, 2015.
- [6] Ruzzarin, Ricardo, Augusto Prates do Amaral, and Marcelo Simion. Sistema integrado de gesto de pessoas com base em competências. Editora AGE Ltda, 2006.
- [7] Sales, G. L. Learning Vectors (LV): um modelo de avaliação da aprendizagem em EaD online aplicando mtricas no-lineares. Diss. Tese Doutorado. Departamento de Engenharia de Teleinformtica. Universidade Federal do Cear. 2010. 239f, 2010.
- [8] Sales, Gilvandenys Leite, Eliana Alves Moreira Leite, and Cassandra Ribeiro Joye. "Gerenciamento da Aprendizagem, Evasão em Ead Online e Possíveis Soluções: Um Estudo de Caso no IFCE." RENOTE 10.3 (2012).
- [9] Stanculescu, Laurentiu Catalin, et al. "Work and play: An experiment in enterprise gamification." Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. ACM, 2016.
- [10] Steffens, Flvio, et al. "Using Gamification as a Collaboration Motivator for Software Development Teams: A Preliminary Framework." (2015).
- [11] Tanaka, Samara, et al. "Gamification, Inc.: como reinventar empresas a partir de jogos." (2013).

## **Session 22: Infonomics and e-Technology**

Title: Designing an Assembly Language Using MIT App Inventor  
(Authors: Casey Wilson, Anna Sandifer, Aspen Olmsted)

Title: Undecidable Problems in Malware Analysis  
(Authors: Ali Aydın Selçuk, Fatih Orhan, Berker Batur)

Title: Ethereum Transaction Graph Analysis  
(Authors: Wren Chan, Aspen Olmsted)

Title: Detection of fake online hotel reviews  
(Authors: Anna V. Sandifer, Casey Wilson, Aspen Olmsted)

# Designing an Assembly Language Using MIT App Inventor

Casey Wilson, Anna Sandifer, Aspen Olmsted

Department of Computer Science

College of Charleston, Charleston, SC 29401

cawilson1@g.cofc.edu, sandiferav@g.cofc.edu, olmsteda@cofc.edu

**Abstract**—Many methods for learning software development neglect computer science concepts for real-world application based education. We combine one computer science concept, hardware-software interaction, and apply it in a software development context using MIT App Inventor. Through App Inventor we have created a virtual computer, implemented with its own binary machine language and assembly language. We hope that students using App Inventor to learn software development by creating apps, use our implementation to learn assembly language concepts by adjusting our languages and virtual hardware for their own purposes and interests. This serves both to help them understand a core computer science concept and to present the concept in a way that encourages practical software development skill improvement.

**Keywords**—*hardware-software interaction; Assembly language; programming; computer science*

## I. INTRODUCTION

As demand for computer and technology specialists increases, so will the number of people who turn to methods alternative to a traditional academic setting for learning the relevant skills to become effective in their target specialties. One such example is the many aspiring programmers who will likely learn through websites and courses which emphasize education in a programming goal oriented context exclusively. While this is an effective method for learning certain tasks associated with software development, this method may neglect to provide a student with a strong understanding of both fundamental or theoretical concepts.

In a traditional academic computer science setting, these fundamental and theoretical skills are generally taught by requiring students take classes that broaden their knowledge, teaching them theoretical problem-solving skills, and often teaching real-world problem solving skills as they relate to computer science. Many current and future software developers may, however, have little access to, little desire to learn, or little knowledge of where to begin learning about many of these less often applicable and tangible computer science skills.

A possible solution to this problem is to find new ways to integrate these skills with the methods that people already use to learn programming. The focus of this paper is to show an example of how this might be accomplished. One knowledge set sometimes overlooked by a programming education outside of academia is a working understanding of computer hardware-

software interaction. We propose a method to teach some basics of hardware-software interaction by providing a student with a user-designed assembly language using MIT App Inventor for Android. App Inventor is a web software which allows users to create Android apps using a drag-and-drop block programming language. It is oriented toward beginner programmers and is thus a good candidate for the integration of computer science education in a system designed to introduce and improve practical software development skills. We hope that students will take our implementation of a virtual computer and architecture and make adjustments and additions to suit their own interests. Should a student follow our method for the creation of a simple assembly language from scratch, we hope he or she will be shown some of the basics of computer hardware-software interaction, as well as receive additional practice in general software development via creating software with App Inventor.

## II. RELATED WORK

Our architecture is most closely modeled after MIPS assembly language because of its use in the popular textbook *Computer Organization and Design* [1]. Our language is, however, separate from MIPS, as this book was used as a guide.

MIPS Assembler and Runtime Simulator (MARS) is a common simulator for students to practice programming in MIPS assembly language. MARS is an IDE that implements several features for its MIPS simulator, including showing 32 registers and allowing a student to assemble and run MIPS code [3]. While these two features are implemented for our program, we focus on an Android App environment rather than a desktop application. The authors of [2] host a MIPS simulator on a mobile environment and include pipelining and dynamic scheduling in their simulator. We choose to overlook hardware specifics for a simpler program oriented toward students with less knowledge of assembly language.

Our fundamental difference from many common assembly language education tools is that our focus is on showing students how languages might be built around hardware, with the hope of making this process more enjoyable and creative for the student. Students are encouraged to use our implementation to learn about assembly and machine languages and alter our program using App Inventor and its beginner-friendly, block programming interface to change, add, or remove features to their own preference.

III. IMPLEMENTATION

The architecture that we implemented for this project was an RISC based on 16 bits. A 16-bit architecture was chosen because this project is intended to be an educational simulation more than a computer architecture designed for real world use, and while the more common 32 and 64-bit architectures could potentially support more commands and features, they would likely be beyond the scope of most students using this system to create a working architecture simulation. Conversely, an architecture based on less bits, like, for example, 8 bits, would likely make a simulation capable of performing a diverse set of computational tasks difficult due to limitations in allowing for a possible substantial set of instructions, as well as significantly limiting the size of easily accessible memory. This particularly applies to immediate-type instructions, which would need to contain at least an opcode, a register value, and an immediate value in a single 8-bit instruction. 16 bits is a good middle ground of functionality and simplicity.

Register and memory location are simulated using arrays. App Inventor provides only a “list” data type with a starting index of 1. Binary addresses are better represented using a starting index of 0 because one more location can be reached in the same number of bits when starting with an index 0 than 1. We found a user-created “array” display from [3] that had a starting index of 0, and we based our hardware simulation on this array representation. There are three array types: a “register” array of length 16 (values 0-15) to simulate a set of CPU registers, a “memory” array of length 256 (values 0-255) to hold the instructions of the program being run, and a “data” array of length 256 (values 0-255) for the user to store and load individual 16 bit words of data to and from the memory.

The app has a simple user interface for input and output, buttons to perform actions, and virtual hardware displays. There is an input text box for the user to enter either binary or assembly instructions and an “Assemble” button to convert the assembly language instructions to their equivalent binary instructions and display them in the same text box. There is also a “Load to Memory and Run” button which takes the binary input, loads it to the beginning of the memory array, runs the program to the language’s specifications, and updates the contents of the register array and data array to the values specified by the user input program. There are also 4 output text boxes should the user specify to display a formal output outside of the register array or data array. For simplification purposes and to encourage intuitive calculation for the user, the instructions moved to the memory array are written in binary, identical what was written by the user in the input text box, but the computed numbers in the register and data arrays, as well as the numbers displayed in the formal outputs, are represented in decimal format. The bottom of the screen displays the register array in one text box, and the memory array with the data array appended to it in another text box.

To manipulate the values in the array displays and output displays, which are all set to 0 by default, we first implemented a binary machine language that dictates the values represented

TABLE 2. DISTRIBUTED SERVICES BY PLATFORM

<i>Instruction</i>	<i>Opcode</i>	<i>Abbreviation</i>	<i>Format Type</i>
Add	0000	add	R
Subtract	0001	sub	R
Multiply	0010	mul	R
Divide	0011	div	R
Set if Less Than	0100	slt	R
Copy	0101	copy	R
Store Data w/Offset	0110	sdo	RWO
Load Data w/Offset	0111	ldo	RWO
Add Immediate	1000	addi	DRI
Subtract Immediate	1001	subi	DRI
Jump If Equal	1010	jeq	SRI
Jump If Not Equal	1011	jne	SRI
Store Data	1100	sd	SRI
Load Data	1101	ld	SRI
Sensor	1110	snsr	SNSR
Print	1111	print	PRINT

in the arrays and outputs. The action of each instruction is controlled by the first 4 bits of each 16-bit word. These 4 bits are called the operation code (opcode). There are 16 basic opcodes and therefore 16 basic instructions. Additionally, there are six instruction format types. These six instruction format types are Register (R), Single Register Immediate (SRI), Double Register Immediate (DRI), Register with Offset (RWO), Sensor (SNSR), and Print (PRINT). The list of instructions and the formatting procedures of each type can be seen in TABLE 1 and TABLE 2, respectively. The set of instructions allows the user to perform many important fundamental computational actions, like doing

TABLE 1. INSTRUCTION FORMATTING PROCEDURE

<i>Format Type</i>	<i>Formatting Procedure</i>	<i>Binary Example</i>	<i>Instruction</i>
	Assembly Example	Binary Example	Add
R <sup>a</sup>	Instruction rv1 rv2 tr add r8 r9 r10	OpCode rv1 rv2 tr 0000 1000 1001 1010	Subtract
RWO	Instruction rv rm imm sdo r8 r9 4	OpCode rv rm imm 0110 1000 1001 0100	Multiply
DRI <sup>b</sup>	Instruction rv tr imm addi r4 r5 20	OpCode rv tr imm 0000 00 01 00010100	Divide
SRI	Instruction rv imm jeq r8 127	OpCode rv imm 1010 1000 01111111	Set if Less Than
SNSR	Instruction sop tr sfn snsr 2 r9 3	OpCode sop tr sfn 1110 0010 1001 0011	Copy
PRINT <sup>c</sup>	Instruction rv out print r8 0	OpCode rv out 1111 1000 00 xxxxxx	Store Data w/Offset

rv = value at specified register      tr = store value to specified register  
imm = immediate      sop = sensor op code      sfn = sensor function  
rm = memory location at register value      out = output  
mi = memory location stored at immediate value  
Load Data w/Offset

a. The last four bits of Copy, an R format type, do not affect the instruction. SLT sets rv1 to a value of 1 if rv2 < tr  
b. The specified registers of the DRI type go to registers r4 – r7  
c. For PRINT instructions, the last 6 bits do not affect the instruction

TABLE 3. SENSOR BINARY INSTRUCTION INFORMATION

<i>Sensor</i>	<i>SOP</i>	<i>Return Information</i>	<i>SFN</i>
Accelerometer	0000	x Acceleration	0000
		y Acceleration	0001
		z Acceleration	0010
Gyroscope	0001	x Angular Velocity	0000
		y Angular Velocity	0001
		z Angular Velocity	0010
Orientation	0010	Azimuth	0000
		Pitch	0001
		Roll	0010
		Magnitude	0011
		Angle	0100
Proximity	0011	Distance	0000
		Maximum Range	0001

arithmetic, storing variables, if statements, looping, and formally printing to an output.

One aspect of our method of learning assembly language that sets it apart from many others is that this language allows a user to collect data from a smartphone’s sensors that can then be treated the same as any other value represented in the registers, data memory, or output. This means that the user can program the information input from the sensors to create simple apps from the app itself. This is also intended to give the user of the app an intuition for the power of a programming language, even a simple one. The sensors implemented, their opcodes and their individual functions are listed in TABLE 3.

It is important to note that while we wrote a specific architecture as we saw appropriate for this application, students are encouraged to alter the architecture and its’ implementation for their own purposes and interests as they see fit. For this reason, we focus less on an ideal implementation, and more on a student’s creation and implementation of some working architecture, to encourage creative thinking and practical development. If a student doesn’t prefer an aspect of our architecture or sees a superior implementation, he or she is encouraged to change it. This principle applies to all aspects of our implementation.

#### IV. FUTURE WORK AND CONCLUSION

One possible adjustment to be made could be the planning and implementation of a 32 or 64-bit architecture. This would allow for more reasonable use binary implementation in both memory and registers. 16 bits limits the size of a binary number to 65,536 unique integers to be represented without implementing double or larger memory device representations. The implementation of a larger bit architecture would also make more practical the implementation of binary representation of the numbers represented in the register and data arrays, as well as binary arithmetic. Floating point numbers are supported with

our implementation, but they are only either the result of division or sensor input, and floating-point arithmetic could also be a useful addition.

Another common programming feature that our architecture lacks in its current implementation is the ability to call functions, which could be an easily added feature with the addition of an additional stack-type array section appended to the memory array.

None of our virtual registers have specific reserved purposes or limitation as to what numbers can be stored, except that the DRI type instructions must be loaded into registers r4 – r7. This being the case, all registers can store all value types either directly or by performing arithmetic operations. An improvement to the current architecture could be the addition of restrictions on the values that certain registers can hold and conventions for certain individual registers or sets of registers. For example, r0 could be restricted to holding a constant value of 0, and certain registers could be reserved for holding the current memory address of the end of the program or the next stack memory location if a stack array is implemented.

While we implemented sensor inputs, we only implemented four types, and others could be added much more extensively with virtually no change to our current implementation. As sensor input is what separates our assembly language from many others that are used for educational purposes, this would be a useful and simple addition for a student learning assembly. We also designed only an opcode for sensors and did not implement abbreviated instructions to be assembled.

In the future, we would like to test the real-world efficacy of our system and see the results. Our assembly implementation on App Inventor is meant to be used as a starting point for a student interested in learning hardware-software interaction. We have implemented a substantial enough architecture to perform many computing problems, but have intentionally left certain potential aspects of implementation easily improvable. Our hope is that students learning assembly language can use our implementation as a good starting point, and they can make adjustments tailored their own interests, helping them learn about hardware-software interaction, and improving their general software development skills by making these alterations using our source code in App Inventor.

#### REFERENCES

- [1] D. A. Patterson, J. L. Hennessy, *Computer Organization and Design, the Hardware-Software Interface*, 5<sup>th</sup> ed. Burlington, MA: Morgan Kaufman, 2015
- [2] M. A. Capuz *et al.*, “Mips-64 Simulator on a Mobile Environment,” in *2016 IEEE Region 10 Symposium (TENSYP)*, Bali, 2016, pp. 219-224
- [3] K. Vollmar and P. Sanderson. (2006). “MARS: An Education Oriented MIPS Assembly Language Simulator” [Online]. Available: <http://courses.missouristate.edu/KenVollmar/MARS/tp288-vollmar.pdf>. [Accessed 04/24/17].
- [4] E. Mitchell. (2015, September 29). “Implementing and ‘Array’ in App Inventor” [Online]. Available: <http://pevest.com/appinventor2/?p=898>. [Accessed 04/24/17]

# Undecidable Problems in Malware Analysis

Ali Aydın Selçuk

Dept. of Computer Engineering  
TOBB University of Economics and Technology  
Ankara, Turkey  
aliaydinselcuk@gmail.com

Fatih Orhan, Berker Batur

Comodo Security Solutions, Inc.  
Clifton, NJ, USA  
{fatih.orhan,berker.batur@comodo.com}

**Abstract**— Malware analysis is a challenging task in the theory as well as the practice of computer science. Many important problems in malware analysis have been shown to be undecidable. These problems include virus detection, detecting unpacking execution, matching malware samples against a set of given templates, and detecting trigger-based behavior. In this paper, we will give a review of the undecidability results in malware analysis and discuss what can be done in practice.

**Keywords**- Computer viruses, malware analysis, virus detection, undecidability.

## I. INTRODUCTION

The number of malware programs encountered by security companies multiplies every year. Each of these programs needs to be analyzed by static and dynamic analysis tools. The task of running each program in a controlled environment and analyzing its behavior manually is a tedious and labor-intensive task. Therefore, there is a great need for automation of this process and for tools that will help with the analysis.

One of the most significant theoretical results in malware analysis is from the seminal works of Cohen on computer viruses [Coh87, Coh89] where he showed that a program that detects all computer viruses precisely is impossible. Later, Chess and White [4] gave an example of a polymorphic virus that cannot be precisely detected by any program. Other results followed [CJS+05, RHD+06, BHL+08] which stated the impossibility of certain critical tasks in static and dynamic malware analysis.

In this paper, we give a brief survey of the major undecidability results found in the malware analysis literature. Then we give examples from the positive side showing what can be done on these undecidable problems in practice.

## II. MALWARE ANALYSIS AND UNDECIDABILITY

Since Cohen [6] gave the first formal treatment of computer viruses, many problems in malware analysis have been shown to be undecidable. Many of these results are based on the fact that precisely deciding whether a given program/input satisfies a certain post-condition, for an arbitrary post-condition, is undecidable. The proofs are based on two general techniques: Either they build a self-contradictory program assuming the existence of a decider for the given problem, similar to [6], or they give a reduction from a well-known undecidable problem, such as the Halting Problem, similar to [7]. In this section, we

review some of the most significant undecidability results in the field.

### A. Undecidability of the General Virus Detection Problem

The first result on the undecidability of the general virus detection problem is due to Cohen [6]. Using a well-known proof technique, he argued that:

“In order to determine that a given program ‘ $P$ ’ is a virus, it must be determined that  $P$  infects other programs. This is undecidable since  $P$  could invoke any proposed decision procedure ‘ $D$ ’ and infect other programs if and only if  $D$  determines that  $P$  is not a virus. We conclude that a program that precisely discerns a virus from any other program by examining its appearance is infeasible.”

He gave the following piece of program “contradictory-virus” as an example that cannot be detected by a virus detector  $D$  in a correct way:

```
program contradictory-virus :=
{...
  main-program :=
  {if ~D(contradictory-virus) then
    {infect-executable;
     if trigger-pulled then
       do-damage;
    }
    goto next;
  }
}
```

As Cohen [6] observed, “... if the decision procedure  $D$  determines  $CV$  to be a virus,  $CV$  will not infect other programs, and thus will not act as a virus. If  $D$  determines that  $CV$  is not a virus,  $CV$  will infect other programs, and thus be a virus. Therefore, the hypothetical decision procedure  $D$  is self-contradictory, and precise determination of a virus by its appearance is undecidable.” A minor flaw in this argument was observed by Steinparz [15], who noted that this argument only shows the impossibility of a virus detector which is not a virus itself. Otherwise, if  $D$  is a virus itself, it can return “true” on contradictory-virus and be correct.

A more formal proof was again given by Cohen himself [7] by a reduction from the Halting Problem. He showed that the existence of a precise virus detector would imply a decider for the Halting Problem and hence is not possible.

### B. Existence of an Undetectable Virus

As summarized above, Cohen [Coh87, Coh89] showed the impossibility of a virus detector that detects all viruses precisely. Chess and White [4] extended this result by showing that there are viruses, in theory, with no error-free detectors. They explained, “That is, not only can we not write a program that detects all viruses known and unknown with no false positives, but in addition there are some viruses for which, even when we have a sample of the virus in hand and have analyzed it completely, we cannot write a program that detects just *that* particular virus with no false positives.”

The result of Chess and White is based on an extension of the contradiction argument in Cohen’s first paper [6]: Consider a polymorphic virus  $W$  that is able to modify its code. This virus modifies its spreading condition such that if some particular subroutine in it returns “false” on  $W$  itself, it spreads. Furthermore, this subroutine is subject to change as a part of  $W$ ’s polymorphism. Now, if some detector code  $C$  were to detect  $W$ , there is at least one instance of this polymorphic virus, where the subroutine is replaced by  $C$ , that cannot be detected by  $C$ : Just like Cohen’s argument, detection by  $C$  would result in the virus’ not spreading, and hence would imply a false positive.

The same argument shows the non-existence of a detector for  $W$  under a looser notion of detection as well: Say a program “detects” a virus  $V$  if it (i) returns “true” on every program infected with  $V$ , (ii) returns “false” on every program not infected with *any* virus, (iii) may return “true” or “false” on a program that is infected with some virus other than  $V$ . The impossibility argument above applies to this looser notion of detection verbatim. Hence, Chess and White [4] concluded that there exists, in theory, some virus that cannot be detected precisely by any virus detector even under this looser notion of detection.

### C. Semantic-Aware Malware Detection

A malware detector based on a pattern matching approach is fundamentally limited against obfuscation techniques used by hackers. The goal of malware obfuscation is to morph or modify the malware to evade detection. A piece of malware can modify itself by, for example, encrypting its payload, and then later decrypting it during execution. A polymorphic virus tries to obfuscate its decryption code using several transformations, such as code transposition, nop insertion, and register reassignment. Metamorphic viruses, on the other hand, try to evade detection through obfuscating the entire code. When they replicate, these viruses change their code by techniques such as substitution of equivalent instruction sequences, code transposition, register reassignment, and change of conditional jumps. The fundamental limitation of the pattern-matching approach for malware detection is that it is mainly syntactic and does not consider the semantics of the program flow and the instructions.

Christodorescu et al. [5] studied a method to overcome this limitation by incorporating instruction semantics to detect malicious code traits. In their framework, malicious behavior is defined by hand-constructed “templates”. The problem of deciding whether a given piece of code contains such a template behavior is modeled as the “Template Matching Problem”.

The Template Matching Problem turns out to be undecidable. Christodorescu et al. [5] gave a reduction from the Halting Problem to the Template Matching Problem, and stated that a precise solution for the general Template Matching Problem is impossible.

### D. Automatic Unpacking for Malware Detection

An obfuscation mechanism that is much used by modern malware is to hide the malicious portion of the payload as data at compile time, and then transform it into an executable at run time, a behavior known as “unpack and execute”. The unpack transformation can be something simple, such as an XOR by a block of random-looking data, or something more complex, such as decryption by a cipher like AES.

Royal et al. [12] worked on detecting such polymorphic viruses by focusing on the result of the unpack operation. The idea is to compare the executable code during the run time with that before the run time. When a change is detected, it is written out for further analysis.

The code and the data sections of a program are formally modeled as a Turing machine  $M$  and its input  $w$ . Then the unpack detection problem becomes whether  $w$  contains another program in it that will be emulated by  $M$  during computation. This problem can be formulated as the following formal language:

$$\text{UNPACKEX}_{\text{TM}} = \{ \langle M, w \rangle : M \text{ is a UTM, and } M \text{ simulates a Turing machine on its tape in its computation on } w \}$$

Royal et al. [12] gave a theorem which stated that the  $\text{UNPACKEX}_{\text{TM}}$  language is undecidable. They proved this result by a reduction from the Halting Problem. Hence, it turns out that determining precisely whether a given program contains some unpack-execute behavior in it is impossible.

### E. Automatically Identifying Trigger-Based Behavior

A common feature found in modern malware is to contain some hidden malicious behavior that is activated only when triggered; such behavior is called trigger-based behavior. Various conditions are used for triggering, such as date and time, some system event, or a command received over the network.

Brumley et al. [2] studied how to automatically detect and analyze trigger-based behavior in malware. Their approach employs mixed symbolic and concrete execution to automatically explore different code paths. When a path is explored, a formula is constructed representing the condition that would trigger execution down the path. Then a solver is employed to see whether the condition can be true, and if so, what trigger value would satisfy it.

Like many other problems in malware analysis, an exact, automatic identification of trigger-based behavior turns out to be undecidable by a reduction from the halting problem. Brumley et al. [2] observed that “Identifying trigger-based behaviors in malware is an extremely challenging task. Attackers are free to make code arbitrarily hard to analyze. This follows from the fact that, at a high level, deciding whether a piece of code contains trigger-based behavior is undecidable, e.g., the trigger condition could be anything that halts the program. Thus, a tool that uncovers all trigger-based behavior all the time reduces to the halting problem.”

#### F. NP-Complete Problems

Although the general cases of the aforementioned problems are undecidable, it turns out that it is possible to obtain their decidable versions by assuming some bound on the time or memory available to the malware. Spinellis [14] showed that a length-bounded version of Cohen’s problem is decidable and NP-complete. Borello and Mé [BM08] showed that detecting whether a given program  $P$  is a metamorphic variant of another given program  $Q$  is decidable and NP-complete. Bueno et al. [3] showed that the space- and time-bounded versions of the unpacking problem are decidable, and the time-bounded version is NP-complete.

Of course, a problem’s being NP-complete is not exactly good news. It is usually interpreted as that no efficient solution exists for the worst case of that problem. However, efficient solutions may exist for the average case, or it can be possible to obtain reasonably good solutions by heuristics or approximation algorithms.

### III. PRACTICAL SOLUTIONS

Despite the negative theoretical results on undecidability of some fundamental questions in malware analysis, practical tools have been in action since the very early days of computer viruses. By tolerating some degree of inaccuracy (i.e., tolerating some degree of false positives or negatives, or allowing inconclusive results), it is possible to build algorithms that are very effective in practice. In this section, we summarize some of the tools developed for the problems reviewed in Section 2.

#### A. Detecting Malware by Template Matching

Despite the fact that the general Template Matching Problem is undecidable, it is possible to detect malware using template matching algorithms that are mostly accurate. Christodorescu et al. [5] developed a toolkit for that purpose. The toolkit works in two phases: First, the binary program to be analyzed is disassembled, a control graph is constructed, one per program function, and an intermediate representation (IR) is generated. The IR is further processed and put into an architecture- and platform-independent form. In the second phase, the IR is compared against a given set of malware templates. Each comparison either returns “yes” or “don’t know”. Suggested malware templates for comparison include procedures such as a decryption loop or mass mail sending.

Christodorescu et al. [5] tested their tool on a real-world malware sample consisting of seven variants of Netsky (B, C,

D, O, P, T, and W), seven variants of Bagle (I, J, N, O, P, R, and Y), and seven variants of Sober (A, C, D, E, F, G, and I), all being email worms with many diverse forms found in the wild. The authors tested the malware against templates capturing the decryption loop and mass mailing functionalities. The tool detected all Netsky and Bagle variants with 100% success. The Sober worm was not detected due to a limitation in the prototype implementation, related to matching calls into the Microsoft Visual Basic runtime library. Nevertheless, their test demonstrated the success of their template matching algorithm on diverse forms of malware.

The tool was tested on a benign sample as well in order to test its false positive rates. 97.78% of the programs in the given sample were detected as benign after successful disassembly, while 2.22% could not be disassembled.

#### B. Detecting Unpack-Execute Behavior

Although the general problem of unpack-execute behavior is undecidable, Royal et al. [12] gave an algorithm for a bounded version of this problem. Let  $n$  denote the number of instructions of a given program  $P$  to execute before it halts. The program `ExtractUnpackedCode( $P, n$ )` works in two phases:

- **Phase 1: Static Analysis.** Program  $P$  is disassembled to identify code and data. Blocks of code that are separated by non-instruction data are partitioned into sequences of instructions. These sequences form the set  $I$ , which will be queried repeatedly in the next phase to detect if  $P$  is executing unpacked code.
- **Phase 2: Dynamic Analysis.** Program  $P$  is executed one instruction at a time. The current instruction sequence is captured by in-memory disassembly starting at the current value of the program counter until some non-instruction data is encountered. The current instruction sequence is compared against each instruction sequence in the set  $I$ . If the current sequence is not a subsequence of any instruction sequence in  $I$ , then it did not exist in  $P$ .

Royal et al. [12] developed this algorithm into a practical tool for MS Windows systems, called PolyUnpack. They tested the tool on the OARC malware suspect repository and compared its performance with that of the Portable Executable Identifier (PEiD), a popular reverse-engineering tool which uses a specific set of signatures to detect unpack-execute behavior [11]. PolyUnpack performed very well and was able to identify many samples with unpack-execute behavior which PEiD was unable to detect.

#### C. Detecting Trigger-Based Behavior

Detection of trigger-based behavior by manual analysis is a virtually impossible task due to the intensive labor required. On the other hand, a precise automatic analysis is not possible either; as explained in Section 2.5, the general problem of automatic identification of trigger-based behavior is undecidable. Nevertheless, a great deal of help can be obtained from automatic analysis to alleviate the burden of manual analysis. Brumley et al. [2] designed a tool for this task. Their approach consisted of several phases: First, the different types

of triggers of interest are specified. Then, different code paths are explored using mixed symbolic and concrete execution. For a path explored by this process, a formula is constructed representing the condition that would trigger execution down the path. Then a solver is employed to see whether the condition can be true, and if so, what trigger value would satisfy it.

Brumley et al. [2] developed this approach into a program called MineSweeper. They tested MineSweeper on real-world malware containing trigger-based behavior. On every case, MineSweeper was able to detect the trigger condition and the trigger-based behavior. The analysis time varied depending on the complexity of the malware, from 2 to 28 minutes. In general, MineSweeper is not guaranteed to detect every piece of malware containing trigger-based behavior, but it can definitely be used as a tool of great assistance over the impractical alternative of manual analysis.

#### D. Malware Protection by Default Deny Approach

Given that some fundamental problems in malware analysis and detection are undecidable, an alternative solution applied by practical security tools (e.g., Comodo's Endpoint Security [8]) is the *default deny* approach to protect users from malware infections: Rather than blocking only blacklisted malware applications and allowing all other safe and unknown applications, only whitelisted applications [13] are permitted to run on a host's real operating system (OS). Samples blacklisted as malware are blocked by default, and unknown programs are permitted to run in some virtualized environments such as containments, sandboxes, etc. Creating these shadow file systems, registries, and communication ports helps blocking most damages caused by malicious application, and correctly defining a program for virus detection [9].

Although the default deny approach provides a higher level of protection compared to the default allow approach, one of the current problems using these virtualized environments is encountered in application usability. Purdila and Terzis [10] developed a dynamic browser containment environment to protect users from web-based malware, where they intercept system service requests of processes and limit browser's access to critical system resources to prevent malware damages. Although they managed to provide the desired protection, their proposed technique introduced some overhead; an 11.8% increase in latency and a 13.4% decrease in throughput.

#### IV. CONCLUSION

Malware detection has been a major problem since the early days of computing. Theoretical results have been given on the inexistence of perfect detectors on various problems. Nevertheless, there is a great deal of work to be done using less-than-perfect tools. Bounded versions of the undecidable malware detection problems are in fact decidable. By assuming certain bounds on the time or memory available to the malware, it should be possible to develop detectors that work quite accurately in practice.

#### ACKNOWLEDGMENT

We would like to thank Hatice Sakarya for her help during the preparation of this paper.

#### REFERENCES

- [1] Jean-Marie Borello, Ludovic Mé, "Code obfuscation techniques for metamorphic viruses", *Journal in Computer Virology*, 4(3):211–220, 2008.
- [2] David Brumley, Cody Hartwig, Zhenkai Liang, James Newsome, Dawn Song, Heng Yin, "Automatically identifying trigger-based behavior in malware" In *Botnet Detection*, pp. 65–88, Springer, 2008.
- [3] Denis Bueno, Kevin J. Compton, Kareem A. Sakallah, Michael Bailey, "Detecting Traditional Packers, Decisively", *Proceedings of the 16th International Symposium on Research in Attacks, Intrusions, and Defenses (RAID 2013)*, 2013.
- [4] David M. Chess, Steve R. White, "An undetectable computer virus", *Proceedings of Virus Bulletin Conference*, vol. 5, 2000.
- [5] Mihai Christodorescu, Somesh Jha, Sanjit A. Seshia, Dawn Song, Randal E. Bryant, "Semantics-aware malware detection", *Proceedings of the 2005 IEEE Symposium on Security and Privacy (S&P'05)*, 2015.
- [6] Fred Cohen, "Computer viruses: theory and experiments", *Computers and Security*, 6(1):22-35, 1987.
- [7] Fred Cohen, "Computational aspects of computer viruses", *Computers and Security*, 8(4):325-344, 1989.
- [8] Comodo, "End Point Security and the Case For Auto Sandboxing", White Paper. <https://containment.comodo.com/resources/white-papers/White-Paper-End-Point-Security-And-The-Case-For-Auto-Sandboxing.pdf>
- [9] David Evans, "On the Impossibility of Virus Detection", 2017.
- [10] Octavian Purdila, Andreas Terzis, "A Dynamic Browser Containment Environment for Countering Web-based Malware", *Proceedings of the 8th RoEdunet International Conference*, 2009.
- [11] Jibz, Qwerton, snaker, xineohP. *PEiD*. peid.has.it, 2005.
- [12] Paul Royal, Mitch Halpin, David Dagon, Robert Edmonds, Wenke Lee, "PolyUnpack: Automating the hidden-code extraction of unpack-executing malware", *Proceedings of the 22nd Annual Computer Security Applications Conference (ACSAC'06)*, 2006.
- [13] Adam Sedgewick, Murugiah Souppaya, Karen Scarfone, "Guide to Application Whitelisting", *NIST Special Publication (800-167)*, 2015.
- [14] Diomidis Spinellis, "Reliable identification of bounded-length viruses is NP-complete", *IEEE Transactions on Information Theory*, 49(1):280–284, 2003.
- [15] Franz X. Steinparz, "A comment on Cohen's theorem about undecidability of viral detection", *Alive*, vol. 1, 1991.

# Ethereum Transaction Graph Analysis

Wren Chan

Department of Computer Science and Engineering  
New York University  
New York, NY  
wc1453@nyu.edu

Aspen Olmsted

Department of Computer Science  
College of Charleston  
Charleston, SC 29401  
olmsteda@cofc.edu

**Abstract**—Cryptocurrency platforms such as Bitcoin and Ethereum have become more popular due to decentralized control and the promise of anonymity. Ethereum is particularly powerful due to its support for smart contracts which are implemented through Turing complete scripting languages and digital tokens that represent fungible tradable goods. It is necessary to understand whether de-anonymization is feasible to quantify the promise of anonymity. Cryptocurrencies are increasingly being used in online black markets like Silk Road and ransomware like CryptoLocker and WannaCry. In this paper, we propose a model for persisting transactions from Ethereum into a graph database, Neo4j. We propose leveraging graph compute or analytics against the transactions persisted into a graph database.

**Keywords**—Ethereum; transaction graph; graph compute; graph analytics; Neo4j;

## I. INTRODUCTION

Ethereum is the second largest cryptocurrency platform after Bitcoin by market capitalization, \$24 billion to \$53 billion (as of August 1, 2017). The currency used in Ethereum is called ether (ETH) whereas the currency used in Bitcoin is also called bitcoin (BTC). Ethereum relies on public blockchain where consensus is maintained by proof-of-work like Bitcoin. A proof-of-work system is one where miners maintain the blockchain by competing to solve computationally intensive problems (which are easy to validate). Ethereum provides the following features on top of what is offered by Bitcoin:

- Ethereum Virtual Machine (EVM) which provides a runtime environment for smart contracts.
- Smart contracts that act as stateful decentralized applications that run on EVM implementations to enforce contract instructions. A taxonomy of smart contracts was proposed by Bartoletti and Pompianu [1] to be divided into five categories: financial, notary, game, wallet, and library.
- Digital tokens that represent digital assets where the issuance is governed by smart contracts. The digital tokens can be treated as currencies like ETH and BTC and can be issued through initial coin offerings (ICOs) akin to IPOs.

Since anonymity is a key promise of cryptocurrency platform and due to the popularity of Bitcoin, much of the existing research deal with weaknesses in Bitcoin and recommendations to fix them. In this paper, we will work on applying the approaches from Bitcoin onto Ethereum. Section II

will describe the related works involving the use of a graph database to persist and analyze transactions. Section III explains the motivation for applying graph analytics on Ethereum. Section IV will cover what was implemented, Section V will cover the results, and Section VI will cover the conclusions and future work.

## II. RELATED WORKS

Spagnuolo et al. [6] created a modular framework, BitIodine that took transaction data and information scrapped from the web to “test several real-world use cases” including discovering connections from a known address and investigating payment for ransomware.

- The transaction data is put through a component called Clusterizer which tries “to find groups of addresses that belong to the same user” based on two heuristics based on how Bitcoin transactions take as input, the output from prior transactions (formally called unspent transaction output). Nick [8] provides a succinct write-up of several heuristics that can be applied for clustering in his thesis.
- The Grapher component creates two graphs, transaction graph and user graph from the transaction data and output from Clusterizer. The transaction graph has addresses as vertices and transactions as edges. The user graph has “users” (effectively cluster of addresses) as vertices and aggregated transactions as edges.
- The Classifier component reads the two graphs created from the Grapher, enriches it with labels that are scrapped from the web and persists the result of the resulting classifications into a database. It provides a reverse index against the transactions, addresses, and users effectively.

Fleder et al. [2] created a similar pipeline to BitIodine which appears to be due to sharing the model for transactions and users from Reid and Harrigan [7]. Unlike Spagnuolo et al [6], Fleder et al. weren't as clear about what underpins the figures provided or any source code to replicate what was done. The PageRank algorithm was mentioned as a guide to finding interesting users in the user graph due to similarities with graph constructed by search-engines to rank websites. When PageRank algorithm was applied for transactions over a single day (October 25, 2013), Fleder et al. found the transactions corresponding to the FBI's seizure of BTC from Silk Road addresses.

Haslhofer et al. [4] created a solution, GraphSense which implements what it calls an address graph. Based on the description of the address graph, it corresponds to the transaction graph that Spagnuolo et al. [6] used for BitIodine. GraphSense creates what it calls an address cluster (“user” in BitIodine) based on heuristics that aren’t mentioned in the paper. Like BitIodine, GraphSense also applies tags based on contextual information scrapped from the web. GraphSense leverages GraphX and GraphFrame API of Apache Spark for graph construction and analytics.

### III. MOTIVATION

As the leading platform for execution of smart contracts, it is important to determine whether the promise of anonymity holds for Ethereum. Like many of the papers written concerning anonymity of Bitcoin, we’re aiming to find potential weaknesses and recommendations for fixing them. The paper is focused on whether an attacker can de-anonymize addresses from graph analytics against transactions on the blockchain.

### IV. RESEARCH AND IMPLEMENTATION

Due to the time and disk space required for downloading the Ethereum blockchain, we’ve opted for loading transactions from REST API provided by Etherscan block explorer for our initial implementation (<https://etherscan.io>). The REST API allows for us to query for all transactions to and from an address of an account. ICOs and digital token transfer were not considered as part of the implementation.

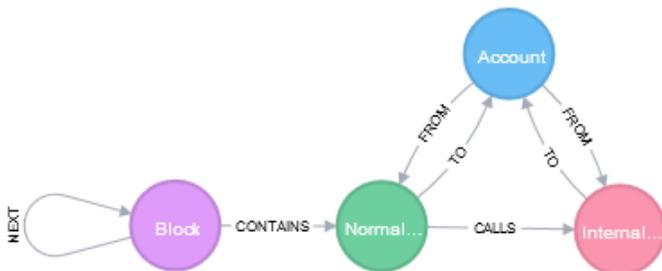


Figure 1. Graph Model

The graph model seen in Figure 1 is used in Neo4j where we have blocks that contain many “normal” transactions which in turn have calls to multiple “internal” transactions. “Normal” transactions in this context refer to transactions that are included directly in the blockchain and therefore signed by the corresponding private key. “Internal” transactions aren’t stored in the blockchain and are the results of contracts being executed in the EVM. Both “normal” and “internal” transactions, in turn, have value in ETH being transferred to and from accounts. Accounts in this model cover normal accounts that are analogous to Bitcoin addresses and contract accounts which are tied to an instance of smart contract.

The graph created for this paper started with initial addresses that are known to be associated with hacks from Etherscan (Gatecoin in May 2016 and Coindash in July 2017):

- 0x04786aada9deea2150deab7b3b8911c309f5ed90

- 0x6a164122d5cf7c840d26e829b46dcc4ed6c0ae48

From the initial addresses, we proceeded to load additional neighboring addresses that are part of the same transactions (for up to three hops away). The load logic was set to avoid loading transactions from accounts that have high in and out-degree transactions. Such accounts could be tumbler service, exchanges or gambling smart contracts which would make our results more difficult to interpret.

### V. RESULTS

Through some simple Cypher queries (Neo4j’s graph query language), we could see some transactions from the Gatecoin hack ending up in addresses that are associated with cryptocurrency exchanges (Changelly, ShapeShift, Bitfinex, Bittrex, and Poloniex) based on tagging done by Etherscan as seen in Figure 2.

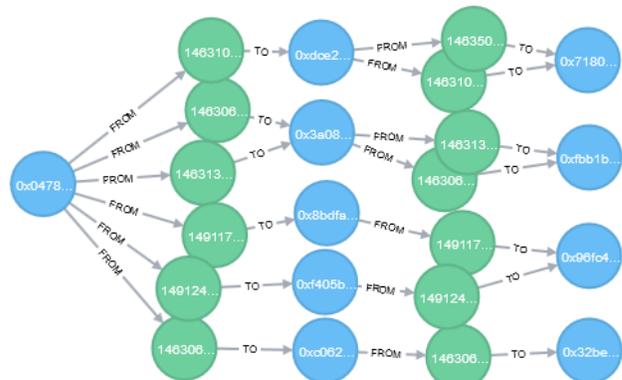


Figure 2. Transactions from Gatecoin Hack – Accounts in blue and transactions in green

### VI. CONCLUSIONS / FUTURE WORK

Due to limitations of what was loaded into the graph database, our insight was limited. Ethereum doesn’t rely on unspent transaction outputs to be inputs for a new transaction, unlike Bitcoin. This means that the heuristics for clustering addresses into users mentioned in Section II aren’t applicable. An extension of the work performed for this paper would be to take the entire Ethereum blockchain and load it into Neo4j. We can subsequently leverage Neo4j extensions or Apache Spark to perform graph analytics such as PageRank algorithm used by Fleder et al. [2]. Aside from graph analytics, we can apply web scrapping to associate addresses with tags such as those provided by Etherscan or user handles from forum posts that mention an address.

There are several studies that rely on network layer traffic to de-anonymize transactions in Bitcoin. One study completed by Biryukov et al. [3] relied on a custom Bitcoin peer to capture the propagation of transactions to fingerprint the entry node where a transaction likely originated from. In effect, these studies function as a side-channel attack that exploited how the Bitcoin platform operated underneath the covers to be able to group transactions together and perform proactive tagging. An attempt to replicate the methodology for these studies would require a

deep understanding of the Ethereum platform and would likely yield better results compared to graph analysis.

The observation that the stolen ETH went into addresses associated with cryptocurrency exchanges suggests that it may be beneficial to link the Ethereum transaction graph with Bitcoin transaction graph. This would provide additional insight to see if ETH was converted into BTC by the hackers and whether there is an associated BTC address.

#### REFERENCES

- [1] M. Bartoletti and L. Pompianu, "An empirical analysis of smart contracts:platforms, applications, and design patterns," *arXiv preprint arXiv:1703.06322*, 2017.
- [2] M. Fleder, M. S. Kester and S. Pillai, "Bitcoin Transaction Graph Analysis," *arXiv preprint arXiv:1502.01657*, 2015.
- [3] A. Biryukov, D. Dhovratovich and I. Pustogarov, "Deanonymisation of Clients in Bitcoin P2P Network," in *2014 ACM SIGSAC Conference on Computer and Communications Security*, Scottsdale, Arizona, USA, 2014.
- [4] B. K. R. F. E. Haslhofer, "O Bitcoin Where Art Thou? Insight into Large-Scale Transaction Graphs," in *SEMANTICS 2016*, Leipzig, 2016.
- [5] R. Boyd, "Graph Compute with Neo4j: Built-in Algorithms, Spark & Extensions," 9 March 2016. [Online]. [Accessed 10 July 2017].
- [6] M. Spagnuolo, F. Maggi and S. Zanero, "BitIodine: Extracting Intelligence from the Bitcoin Network," in *FC 2014: Financial Cryptography and Data Security*, Christ Church, Barbados, 2014.
- [7] F. Reid and M. Harrigan, "An Analysis of Anonymity in the Bitcoin System," in *Security and Privacy in Social Networks*, New York, Springer, 2013, pp. 197-223.
- [8] J. D. Nick, "Data-Driven De-Anonymization in Bitcoin," ETH-Zürich, Zürich, 2015.

# Detection of fake online hotel reviews

Anna V. Sandifer

Department of Mathematics &  
Computer Science  
The Citadel  
Charleston, SC  
asandife@citadel.edu

Casey Wilson

Department of Computer Science  
College of Charleston  
Charleston, SC  
cawilson1@g.cofc.edu

Aspen Olmsted

Department of Computer Science  
College of Charleston  
Charleston, SC 29401  
olmsteda@cofc.edu

**Abstract**— Individuals use online reviews to make decisions about available products and services. In recent years, businesses and the research community have shown a great amount of interest in the identification of fake online reviews. Applying accurate algorithms to detect fake online reviews can protect individuals from spam and misinformation. We gathered filtered and unfiltered online reviews for several hotels in the Charleston area from yelp.com. We extracted part-of-speech features from the data set, applied three classification models, and compared accuracy results to related works.

**Keywords:** fake review detection, machine learning, part-of-speech tags, Multinomial Naïve Bayes classifier, Bernoulli Naïve Bayes classifier, logistic regression classifier.

## I. INTRODUCTION

There are thousands of reviews online, which makes it convenient for people to make decisions, but the amount of data makes it difficult to sort through [1]. The real value of online reviews is in its content and the certainty that reviewer indeed received products or services prior to writing the review. Promotion or demotion of the products and services is one of the main reasons for deceptive reviews. At times, to create better ratings for the venue, hotel owners pay employees to fabricate false reviews [2]. Alternatively, some reviewers write negative reviews for malicious reasons, like to distort the reputation of the business reviewed [3].

Yelp.com is one of the biggest online review sites. It uses a filtering algorithm to detect fake reviews. However, the algorithm is a trade secret. In this work, we collected reviews from yelp.com for 100 random hotels in the Charleston area. We labeled filtered reviews as real and unfiltered reviews as fake. We extracted part-of-speech features, trained and tested the data set, built a model and compared results to related work.

The rest of the paper is organized as follows: Section II reviews related work. Section III describes the motivation for our work. Section IV describes the data collection mechanism and analysis. Finally, Section V will provide a conclusion and future work.

## II. RELATED WORK

Ott et al. [4] define deceptive opinion spam as “fictitious opinions that have been deliberately written to sound authentic in order to deceive the reader.” To analyze deceptive opinion

spam, authors performed a hotel review analysis, using a data set of 800 reviews. Truthful reviews were gathered from tripadvisor.com, and fake reviews were written by Amazon Mechanical Turk online workers (known as Turkers). Each Turker was tasked to write a deceptive review on one of the 20 Chicago hotels. Reviews were written from the perspective of the hotel’s marketing department employees. They satisfied certain length and complexity criteria and portrayed the hotel in a positive light. The Turker got paid \$1 per accepted review. There were 6977 truthful reviews for the same Chicago hotels that were mined from tripadvisor.com, but only 400 reviews were selected for research.

Two approaches were chosen to detect deceptive opinion spam: human performance, and automated approach. Three graduate students were selected to analyze the reviews. For an automated approach, authors used n-gram features as well as a combination of the n-gram features and psycholinguistically-motivated features. Accuracy results using human judges was reported at 61.9%, using bigram features 89.6%, and using psycholinguistic deception detection approach 76.8%.

Mucherjee et al. [5] used filtered and unfiltered data from yelp.com to analyze real and pseudo reviews. Yelp’s filter claimed to be very accurate with a few false positives. Authors used reviews from 85 hotels and 130 restaurants in the Chicago area. While analyzing reviews, authors noticed that the quantity of fake reviews on yelp.com is much smaller than real reviews. They used balanced data (50% fake and 50% real reviews) to train the data set and used natural distribution to test the set. Mucherjee et al. used POS (part-of-speech) based features to build classifiers. All experiments were based on 5-fold Cross Validation and yielded 67.8% accuracy.

McCallum and Nigam [6] performed an evaluation of two event models for the Naïve Bayes Text Classification. The first model was a multi-variate Bernoulli event model. This model calculates the probability of the document, by multiplying the probability of all attribute values, including the probability of non-occurrence for words that were not present in the document. The second model was the multinomial event model. This model calculates the probability of the document, by multiplying the probability of words that occurred in that document. Authors completed the empirical study with different data sets, such as the Newsgroups and WebKB. Results showed that the second model outperformed the first one in terms of large vocabulary sizes. The multinomial model reduced accuracy error by 27%.

### III. MOTIVATING EXAMPLE

Product online reviews provide an information base for consumer experience. Online reviews appear on websites that sell products and offer services. Consumers are likely to look for information online before making purchase decisions. According to Fang et al. [1], 65% of leisure service consumers will read online reviews, before deciding on the vacation destination. The quality of individual hotels is measured through different means, such as stars or ratings on characteristics of cleanliness, view, and staff friendliness. However, online reviews provide support for those means. It is critical to be able to identify fake online reviews with great accuracy.

### IV. IMPLEMENTATION

Our first step was to collect the data. We used Chrome Web Scraper to gather reviews from yelp.com for 100 random hotels in the Charleston area. Non-filtered reviews were tagged as real and filtered reviews were tagged as fake. Table I provides dataset statistics.

TABLE I. DATASET STATISTICS

	Hot els	Fake	Non-fake	% of fake	Total reviews
Our work	100	640	3310	16.2%	3950
Mucherjee et al.	85	4876	4876	14.1%	5678
Ott et al.	20	400	400	50%	800

Next, we tokenized both datasets, extracted adjectives, adverbs, and verbs and tagged words as real or fake. Most informative features and their probabilities are shown in Table II, where  $w$  is the word,  $P(w|R)$  is the probability of the word occurring in the real review, and  $P(w|F)$  is the probability of the word occurring in the fake review.

TABLE II. MOST INFORMATIVE FEATURES

w	P(w R)	P(w F)
assume	1.0	14.4
suggested	1.0	11.1
packing	1.0	10.3
tells	1.0	10.3
easier	1.0	8.7

We calculated a frequency distribution of extracted features. To train the data, we used half of the features, and the other half was used for testing. For model building, we used NLTK [7] and Scikit-Learn [8] modules. To identify fake online reviews, we used Multinomial Naïve Bayes, Bernoulli Naïve Bayes, and logistic regression models. The best accuracy result was achieved by applying the Multinomial Naïve Bayes model and yielded 85.9% accuracy. Figure 1 compares accuracy results for our work and related work.

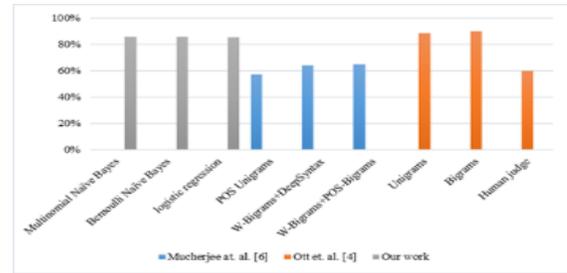


Figure 1. Accuracy Result

### V. CONCLUSION AND FUTURE WORK

In this paper, we collected data, extracted part-of-speech features, and applied three different classification models to identify fake online reviews. We found that highest accuracy was achieved by applying the Multinomial Naïve Bayes classification model to our dataset.

In our future work, we plan to improve on our methods for extracting features from datasets. We also intend to use more classification models, develop voting algorithms, and introduce a reliability score.

### REFERENCES

- [1] B. Fang, Q. Ye, D. Kucukusta and R. Law, "Analysis of the perceived Value of online tourism reviews: Influence of readability and reviewer characteristics," *Tourism Management*, pp. 498-506, 2016.
- [2] K. L. Short, "Buy My Vote: Online Reviews for Sale," *Vanderbilt Journal of Entertainment and Technology Law*, 2013.
- [3] S. Banerjee and A. Y. Chua, "Understanding the process of writing fake online reviews," in *Ninth International Conference On Digital Information Management (ICDIM 2014)*, 2014.
- [4] M. Ott, Y. Choi, C. Cardie and J. Hancock, "Finding Deceptive Opinion Spam by Any Stretch of the Imagination.," in *49th Annual Meeting of the Association for Computational Linguistics*, Portland, 2011.
- [5] A. Mukherjee, V. Venkataraman, B. Liu and N. Glance, "Fake Review Detection: Classification and Analysis of Real and Pseudo Reviews," Department of Computer Science (UIC-CS-2013-03), Chicago, 2013.
- [6] A. McCallum and K. Nigam, "A comparison of event models for Naive Bayes text classification.," 2010.
- [7] "Natural Language Toolkit," [Online]. Available: <http://www.nltk.org/>. [Accessed 26 March 2017].
- [8] "Scikit-learn Machine Learning in Python," [Online]. Available: <http://scikit-learn.org/stable/>. [Accessed 26 March 2017].



**ICITST**

