

David C. Wyld,
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

9th International Conference of Security, Privacy and Trust Management (SPTM 2021),
April 24 - 25, 2021, Copenhagen, Denmark.



AIRCC Publishing Corporation

Volume Editors

David C. Wyld,
Southeastern Louisiana University, USA
E-mail: David.Wyld@selu.edu

Dhinaharan Nagamalai (Eds),
Wireilla Net Solutions, Australia
E-mail: dhinthia@yahoo.com

ISSN: 2231 - 5403

ISBN: 978-1-925953-39-8

DOI: 10.5121/csit.2021.110501 - 10.5121/csit.2021.110507

This work is subject to copyright. All rights are reserved, whether whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the International Copyright Law and permission for use must always be obtained from Academy & Industry Research Collaboration Center. Violations are liable to prosecution under the International Copyright Law.

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

The 9th International Conference of Security, Privacy and Trust Management (SPTM 2021), April 24 - 25, 2021, Copenhagen, Denmark, 7th International Conference on Image Processing and Pattern Recognition (IPPR 2021), 8th International Conference on Computer Science and Information Technology (CSIT 2021) and 2nd International Conference on Big Data and Applications (BDAP 2021) was collocated with 9th International Conference of Security, Privacy and Trust Management (SPTM 2021). The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The SPTM 2021, IPPR 2021, CSIT 2021 and BDAP 2021 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically.

In closing, SPTM 2021, IPPR 2021, CSIT 2021 and BDAP 2021 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the SPTM 2021, IPPR 2021, CSIT 2021 and BDAP 2021.

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

David C. Wyld,
Dhinaharan Nagamalai (Eds)

General Chair

David C. Wyld,
Dhinaharan Nagamalai (Eds)

Organization

Southeastern Louisiana University, USA
Wireilla Net Solutions, Australia

Program Committee Members

Abdel-Badeeh M. Salem,
Abderrahim Siam,
Abdulhamit Subasi,
Abdulhamit Subasi,
Abhas kumar singh,
Abhishek Shukla,
Adrian Olaru,
Ahmad A. Saifan,
Ahmed Abdelgawad,
Ahmed Farouk AbdelGawad,
Alejandro Regalado Mendez,
Ali A. Al-Zuky,
Ali Abdrhman Mohammed Ukasha,
Ali El Zaart,
Ali Hussein Wheeb,
Ali Kaveh,
Amari Houda,
Amer AbuAli,
Amizah Malip,
Amleto Di Salle,
Anas M.R. AlSobeh,
Ankur Singh Bist,
Anna Bobkowska,
Anouar Abtoy,
Antony P J,
Anwar Basha H,
Athanasios Vasilakos,
Atif Bilal,
Attila Kertesz,
bdullah,
Bernard Stepien,
Beshair Alsiddiq,
Bin Zhao,
Bouchra Marzak,
Boukari nassim,
Carlo Sau,
Carlos Juiz,
Chiam Yin Kia,
Christian Mancas,
Chukwuemeka Ifegwu Eke,
Dac-Nhuong Le,
Daina Gudoniene,
Ain Shams University, Egypt
University of Khenchela, Algeria
Effat University, Saudi Arabia
University of Turku, Finland
Isroset, India
R D Engineering College, India
University Politehnica of Bucharest, Romania
Yarmouk university, Jordan
Central Michigan University, USA
Zagazig University, Egypt
Universidad Del Mar, Mexico
Mustansiriya University, Iraq
Sebha University, Libya
Beirut Arab University, Lebanon
University of Baghdad, Iraq
Iran University of Science and Technology, Iran
Networking & Telecom Engineering, Tunisia
Taibah University, Saudi Arabia
University of Malaya, Malaysia
University of L'Aquila, Italy
Yarmouk University, Jordan
Advanced Technology, India
Gdansk University of Technology, Poland
Abdelmalek Essaadi University, Morocco
A J Institute of Engineering & Technology, India
SRM Institute of Science and Technology, India
Lulea University of Technology, Sweden
Jilin University, China
University of Szeged, Hungary
Adigrat University, Ethiopia
University of Ottawa, Canada
Prince Sultan University, Saudi Arabia
Xidian University, China
Hassan II University, Morocco
skikda university, algeria
Universit degli Studi di Cagliari, Italy
University Of The Balearic Islands, Spain
University of Malaya, Malaysia
Ovidius University, Romania
University of Abuja, Nigeria
Haiphong University, Vietnam
Kaunas University of Technology, Lithuania

| | |
|----------------------------|---|
| Daniel Ekpenyong Asuquo, | University of Uyo, Nigeria |
| Dheyaa Jasim Kadhim, | University of Baghdad, Iraq |
| Dilana Hazer-Rau, | University of Ulm, Germany |
| Dimitris Kanellopoulos, | University of Patras, Greece |
| Eduardo Simas, | Federal University of Bahia, Brazil |
| Eke Chukwuemeka, | University of Abuja, Nigeria |
| Elżbieta Macioszek, | Silesian University of Technology, Poland |
| Eng Islam Atef, | Alexandria University, Egypt |
| Eugenio Zimeo, | University of Sannio, Italy |
| Fabio Gasparetti, | Roma Tre University, Italy |
| Fajiang Yu, | Wuhan University, China |
| Farouq Otoom, | Philadelphia Univeristy, Jordan |
| Farouq Saber Al-Shibli, | Philadelphia University, Jordan |
| Felix J. Garcia, | Clemente University of Murcia, Spain |
| Giuliani Donatella, | University of Bologna, Italy |
| Grigorios N. Beligiannis, | University of Patras, Greece |
| Grzegorz Karoń, | Silesian University of Technology, Poland |
| Gurjit Kaur, | Delhi Technological University, India |
| Hacer Yalim Keles, | Ankara University, Turkey |
| Hadi Amirpour, | Universidade da Beira Interior, Portugal |
| Hala Abu Khalaf, | Palestine Polytechnic University, Palestine |
| Hamid Ali Abed AL-Asadi, | Basra University, Iraq |
| Hamid Khemissa, | USTHB University Algiers, Algeria |
| Hamid Mcheick, | Université du Québec à Chicoutimi, Canada |
| Hamimah Ujir, | Universiti Malaysia Sarawak, Malaysia |
| Hariharan, | Saveetha Engineering College, India |
| Hashem H M Ramadan, | Bangalore University, India |
| Hassan Ugail, | University of Bradford, UK |
| Hayfaa A. Atee, | Middle Technical University (MTU), Iraq |
| Hiroshi Ban, | osaka university, Japan |
| Ibrahim Hamzane, | Hassan II university, Morocco |
| Ilham Huseyinov, | Istanbul Aydin University, Turkey |
| Isa Maleki, | Islamic Azad University, Iran |
| Israa Shaker Tawfic, | Ministry of Science and Technology, Iraq |
| Iyad Alazzam, | Yarmouk University, Jordan |
| Jang Eui Hong, | Chungbuk National University, South Korea |
| Jia Ying Ou, | York University, Canada |
| Jonah Lissner, | technion - israel institute of technology, Israel |
| Jong-Ha Lee, | Keimyung University, South Korea |
| Juntao Fei, | Hohai University, P. R. China |
| Kamel Benachenhou, | Blida University, Algeria |
| Ke-Lin Du, | Concordia University, Canada |
| Kenjiro T. Miura, | Shizuoka University, Japan |
| Khalid M.O Nahar, | Yarmouk University, Jordan |
| Khalid Nazim Abdul Sattar, | Majmaah University, Saudi Arabia |
| Kire Jakimoski, | FON University, Republic of Macedonia |
| Kiril Alexiev, | Bulgarian Academy of Sciences, Bulgarian |
| Klenilmar L. Dias, | Federal Institute of Amapa, Brazil |
| Koh You Beng, | University of Malaya, Malaysia |
| Lerina Aversano, | University of Sannio, Italy |
| Liansheng Tan, | Central China Normal University, China |
| Lutz Schubert, | University of Ulm, Germany |

| | |
|-----------------------------|--|
| M. Vijayalakshmi, | Thiagarajar College of Engineering, India |
| Mabroukah Amarif, | Sebha University, Libya |
| Mahdi Sabri, | Islamic Azad University Urmia Branch, Iran |
| Mario Versaci, | Reggio Calabria, Italy |
| Marzak Bouchra, | Hassan II University, Morocco |
| Mayssa Frikha, | Faculty of Sciences of Sfax, Tunisia |
| Mihoub Sofiane, | University of Tiaret, Algeria |
| Mirsaeid Hosseini Shirvani, | Islamic Azad University, Iran |
| Mohamed Ismail Roushdy, | Ain Shams University, Egypt |
| Mohammad A. Alodat, | Sur University College, Oman |
| Mohammed A. Akour, | Yarmouk University, Jordan |
| Morteza Alinia Ahandani, | University of Tabriz, Iran |
| Mourad Chabane Oussalah, | University of Nantes, France |
| Munish Saini, | Guru Nanak Dev University, India |
| Murat DENER, | Gazi University, Turkey |
| Musard Balliu, | KTH Royal Institute of Technology, Sweden |
| Nadia Abd-alsabour, | Cairo University, Egypt |
| Nahlah Shatnawi, | Yarmouk University, Jordan |
| Neha Pattan, | Carnegie Mellon University, Pennsylvania |
| Nikolai Prokopyev, | Kazan Federal University, Russia |
| Nima Jafari Navimipour, | Islamic Azad University, Iran |
| Noriko Hanakawa, | Hannan University, Japan |
| Olakanmi Oladayo O, | University of Ibadan, Nigeria |
| Otilia Manta, | Romanian American University, Romania |
| Paulo Batista, | University of Évora, Portugal |
| Pavani Konagala, | University of Mississippi, USA |
| Pavel Loskot, | Swansea University, United Kingdom |
| Pijush Barthakur, | KLS Gogte Institute of Technology, India |
| Ramadan Elaïess, | University of Benghazi, Libya |
| Richa Purohit, | Y Patil international University, India |
| Rob Fuller, | University of British Columbia, Canada |
| Said Agoujil, | Moulay Ismail University, Morocco |
| Sathyendra Bhat J, | St Joseph Engineering College, India |
| satish gajawada, | IIT Roorkee, India |
| Satyananda Reddy, | Andhra University, India |
| Sebastian Fritsch, | IT and CS enthusiast, Germany |
| Shah Khalid Khan, | RMIT University, Australia |
| Shahid Ali, | AGI Education Ltd, New Zealand |
| Shahram Babaie, | Islamic Azad University, Iran |
| Shamneesh Sharma, | Poornima University, India |
| Smain femmam, | UHA University, France |
| Soon-Geul Lee, | Kyung Hee Univ, Republic of Korea |
| Stefano Michieletto, | University of Padova, Italy |
| Subhendu Kumar Pani, | Oec, Bput, India |
| Suhad Faisal, | University of Baghdad, Iraq |
| sukhdeep kaur, | Punjab technical university, India |
| Sun-yuan Hsieh, | National Cheng Kung University, Taiwan |
| Swati Nikam, | Savitribai Phule Pune University, India |
| Taha Mohammed Hasan, | University of Diyala, Iraq |
| Tanvi Agrawal, | IIT Bombay, India |
| Tanzila Saba, | Prince Sultan University, Saudi Arabia |
| Taskeen Zaidi, | Shri Ramswaroop Memorial University, India |

Technically Sponsored by

Computer Science & Information Technology Community (CSITC)



Artificial Intelligence Community (AIC)



Soft Computing Community (SCC)



Digital Signal & Image Processing Community (DSIPC)



Organized By



Academy & Industry Research Collaboration Center (AIRCC)

**9th International Conference of Security, Privacy and Trust
Management (SPTM 2021)**

**Role-Based Embedded Domain-Specific Language for Collaborative
Multi-Agent Systems through Blockchain Technology.....01-19**
Orcun Oruc

**Federated Identity Management (FIDM) Systems Limitation and
Solutions.....21-36**
Maha Aldosary and Norah Alqahtani

**7th International Conference on Image Processing and Pattern
Recognition (IPPR 2021)**

Informative Multimodal Unsupervised Image-to-Image Translation.....37-51
Tien Tai Doan, Guillaume Ghyselinck and Blaise Hanczar

Image Classifiers for Network Intrusions.....53-59
David A. Noever and Samantha E. Miller Noever

**8th International Conference on Computer Science and Information
Technology (CSIT 2021)**

A Study of Regression Testing for Trade me Website.....61-75
Kenil Manishkumar Patel and Shahid Ali

**On-Demand Video Tagging, Annotation, and Segmentation in
Lecture Recordings to Enhance E-learning Effectiveness.....77-86**
Ken D. Nguyen and Muhammad Asadur Rahman

**2nd International Conference on Big Data and
Applications (BDAP 2021)**

**Event-Driven Time Series Analysis and the Comparison of
Public Reactions on Covid-19.....87 - 101**
Md. Khayrul Bashar

Role-Based Embedded Domain-Specific Language for Collaborative Multi-Agent Systems through Blockchain Technology

Orçun Oruç

TU Dresden, Software Technology Group, Nöthnitzer Straße 46, 01187,
Dresden

Abstract. Multi-agent systems have evolved with their complexities over the past few decades. To create multi-agent systems, developers should understand the design, analysis, and implementation together. Agent-oriented software engineering applies best practices through mainly software agents with abstraction levels in the multi-agent systems. However, abstraction levels take a considerable amount of time due to the design complexity and adversity of the analysis phase before implementing them. Moreover, trust and security of multi-agent systems have never been detailed in the design and analysis phase even though the implementation of trust and security on the tamper-proof data are necessary for developers. Nonetheless, object-oriented programming is the right way to do it, when implementing complex software agents, one of the major problems is that the object-oriented programming approach still has a complex process-interaction and a burden of event-goal combination to represent actions by multi-agents. Designated roles with their relationships, invariants, and constraints of roles can be constructed based on blockchain contracts between agents. Furthermore, in the case of new agents who participate in an agent network, decentralization and transparency are two key parameters, which agents can exchange trusted information and reach a consensus aspect of roles. This study will take the software agent development as a whole with analysis, design, and development with role-object pattern in terms of smart contract applications. In this paper, we aim to propose a role-based domain-specific language that enables smart contracts which can be used in agent-oriented frameworks. Furthermore, we would like to refer to methodology, results of the research, and case study to enlighten readers in a better way. Finally, we summarize findings and highlight the main research points by inferencing in the conclusion section.

Keywords: Software agents, Domain-specific languages, Blockchain technology, Smart contracts, Role-based programming languages.

1 Introduction

Agent-oriented programming (AOP) can be considered as a subset of object-oriented programming by showing the state of an object with human-like features such as belief, desire, intentions, and goals. Moreover, an agent should be in the interaction with other agents, in this way, agents are able to play roles as human-being does. AOP specializes the object-oriented programming methodology by fixing state and

modules (called agents) to consist of features that are coming from agent behaviors [1]. Besides, agents can handle the message passing between other agents internally.

Agent-based systems have changed their characteristics over the past few decades aspect of design, analysis, and implementation. Although one can find out the difficulty of exact definition in terms of the multi-agent system, multi-agent systems are used broadly in the application areas such as supply chain management, distributed systems, smart grids, robotic motion planning. Multi-agent systems are strongly dependent on contexts and roles. Each agent plays a role and it has a minimal set of attributes that represents the environment. Moreover, agents should have behaviors that are, in essence, related to implementing deterministic or non-deterministic behaviors of an agent that operate a role that can be in sequential order, cyclic order, or parallel order.

Agents must be in a relationship with external trusted parties to provide privacy and consistency in multi-agent systems. However, the trust was mostly provided by different logic interpretations and cumbersome ontological definitions in multi-agent systems. Blockchain technology offers a credible and private data pool that can be used with programmable contracts (smart contracts) as a shared database. As we solved the credibility problem with blockchain technology, we can enforce the data layer security, which is vital for data-driven multi-agent system communication, by implementing smart contracts on the application layer of the blockchain protocol. A smart contract is a piece of code that is stored on a blockchain by triggering coin-based transactions with saved data and which reads and writes data in a blockchain database [2]. In addition, smart contracts can ensure the testability of role features, secure transactions within the blockchain database, and separation of the business logic (model) and application logic (system architecture).

Role-based programming can be integrated into the agent concept, which is useful to reduce the complexity of the agent system design by categorizing the roles played by agents and describing the collaboration among agents [2]. We should consider equally grouping roles together in a collaborative relation or a compartment. Roles are essentially defining context-oriented software, which is an explicit data model of roles or objects that combine conditions, activated relationships of roles, and deactivate relationships of roles [3]. For instance, an agent may produce different results under different contexts, hence the given agent behaves differently in a specified environment or a collaborative agent simulation.

Compartments belong to the research of the Compartment Role Object Model (CROM) that establishes subtypes of natural types and relationship types between combined roles [4]. CROM combines the behavioral, relational, and context-dependent nature of roles in a common framework [4]. It is a research project that points out a framework for conceptual modeling that incorporates roles, graphical

¹ <https://www.coindesk.com/three-smart-contract-misconceptions>

modeling language, and a set-based formalization of roles, which has been conducted by TU Dresden Software Technology [4].

Roles can make the design of multi-agent systems easier by implementing the composition of role attributes, role invariants, role methods, and binding-interfaces. The difference between roles and objects is whether or not the roles can move hosts that exist in an environment [5]. Role-based software agents are related to context-aware multi-agent systems. Context is any information that is accessible to the program, where an entity is a person, place, or another agent that is considered relevant to the determination of behavioral variations [6]. Agents can dynamically collaborate with roles, create coalitions of trusted partners as an effective mechanism to communicate with service requestors, find services requested by them, and determine trusted services and provide services to the applicants without violating the privacy of the predefined environment [7]. A domain-specific language in an agent-oriented language can map abstraction of role-compartment to particular composition in an agent-oriented architecture. By adding design-by-contract language such as Solidity, agent-oriented language can assure role constraints regarding role-compartments.

1.1 Research Problem

Principally, multi-agent systems have been used as a software design methodology for software application problems over a few decades. Frameworks and agent communication languages that were proposed are still hard to understand and use effectively in a decentralized and centralized network. Lack of standardization in the area of analysis, design, and implementation increases software design complexity as we plan to deploy decentralized agents.

Agent communication languages such as KQML, KIF, FIPA-ACL, AgentSpeak, and major agent deployment frameworks JADE [8], JADEX [9], JASON [10], GOAL [11], JACK Framework [12], JaCaMo [13], 3APL [14], and 2APL [15] do not offer any solution for privacy, security, and trust at the level of deployment of agents. Furthermore, a variety of these languages creates a burden for language mapping. Moreover, previous solutions have no practical design-by-contract approach so as to establish the goals and actions of agents.

Researchers who deal with multi-agent systems still offer limited modeling solutions for the aforementioned problems. A domain-specific modeling language that merges general-purpose language, agent communication language, and blockchain-based design-by-contract language can make the developers' and researchers' life easier and we can map roles and goals at the analysis phase to the deployment phase through smart contract language.

Thus, current challenges of the programming aspect of the multi-agent systems lead us to create a new approach and a solution as we named role-based blockchain-enabled domain-specific language for collaborative multi-agent systems.

To conclude up regarding the research problem, we have defined research questions(RQ) as below:

- *RQ1*: Can a domain-specific language that comprises the main features of agent communication language, agent framework, and smart contract language be created?
- *RQ2*: How can roles, goals, and compartments be implemented with the domain-specific language?

1.2 Motivation and Challenges

The main motivation of this study is to create a goal-driven (so-called cognitive) agent-oriented language with blockchain technology to provide goals, desires, and intentions in multi-agent systems. The current challenges of programming in multi-agent systems lead us to create a new approach and solution in order to solve the aforementioned problems in the Introduction section.

- Multi-agent systems or swarm management should assign trust and privacy levels for new agents that consist of roles, goals, and plans to increase efficiency in the network.
- Multi-agent systems should ensure trust and privacy in data-driven domains. A human operator or an external participant should see it as a black-box process.
- System planning with the belief-desire-intention reasoning engine suffers the vulnerability of critical decisions. Such decisions may be capabilities, role assignment between agents, limitations of follower agents, and leader agents while changing positions.
- Protection against malicious agents is dependent on mostly language virtual machine environments. A developer should know the specifications of a language that relies on a virtual machine such as Java, which is a cumbersome and error-prone task. If an agent is allowed to communicate with external agents, the smart contract can alleviate the complexity of the security task.

1.3 Outline of Objectives and Contributions

Goal: The main goal of the present study is to implement a domain-specific language to demonstrate role constraints, types, invariants, and relationships using design-by-contracts, which can be done by external blockchain programming language such as smart contracts, with a secured and trusted environment for software agents.

Objectives:

- Identify existing roles from Compartment Role Object Model (CROM) in the context of compartments.

- Mapping from natural types, role types, compartment types, and relationship types to the data structures to a smart contract language such as Solidity. We would like to write a code generator from a general-purpose language to create a smart contract language so that we will have a common language with the agent-oriented programming language.
- Implementing role-definition to agent containers because entities can be bounded to devices on which agents are able to move, create, and deploy.
- Defining collaborative goal-driven roles for agents themselves that reside in agent containers.
- At the implementation phase, we defined the above-mentioned roles' and goals' specifications in a smart contract language for design-by-contract to combine with an agent-oriented programming framework.
- If we have enough time in the course of Ph.D., we will deploy a real multi-agent system such as a multi-agent unmanned air vehicle or robotic arm collaboration.

We will contribute to different aspects inducting from different research areas such as smart contract programming in the blockchain, multi-agent system development frameworks and communication languages, and decentralized agent networking. We have listed our conceivable contributions in this study:

- We introduce a new role-based agent-oriented domain-specific language that is capable to use blockchain technology at the application layer through smart contracts.
- We will evaluate the new language with existing agent-oriented languages aspect of performance, usability, fault tolerance, adaptability, and cost of communication between agents.
- We will reduce the overhead of software agent design, analysis, and implementation for the belief-desire-intention framework by proposed domain-specific language.

2 Background

Although software agents are not a new concept, there has not been found specific definition regarding what exactly should be. In essence, an agent can be either physical, software-based, or a combination of them. This kind of feature brought us to define that the software agents must be in a new category. In this study, we will implement software agents with role-oriented programming that works with role-constraints, role-invariants, role-relationship, and compartments using smart contracts in blockchain technology.

Role-based systems are autonomic systems, which means that the role-based multi-agent systems design planned capabilities and collaborative skills to delegate tasks to components [16]. Roles are an abstract concept of objects that can be

transferable between software agents; however, liveness is much longer than an object. Moreover, roles can have compartments that consist of states of roles, contexts in a software agent, and events. A software agent can connect with other software agents. While an agent is transferring a message to other agents, it should have two unique features which are:

- **Self-awareness:** States and context can be adaptable according to the environment. Software agents should adjust their contexts according to the environment.
- **Self-configuring:** When a new software agent has joined into the network, the agent should configure and reconfigure itself.

Context-dependency refers to the self-awareness and self-configuring definitions to provide agent awareness in a dynamic and high-flexible environment in multi-agent systems. Roles can be assigned to a specialized context and one can use multiple contexts in multiple compartments. Since contexts are strictly bounded by runtime evaluation, design-by-contract can be more useful than test-driven development which is the compile-time metaprogramming feature.

The idea of usability of the blockchain technology for multi-agent systems will be tested and implemented with this study. Commonly, trust and privacy should be provided by external components, application programming interfaces, or other intrusive technologies. In this study, we want to implement the application layer of blockchain so that one can easily employ a multi-agent system in a non-intrusive secure decentralized computing platform.

Design-by-Contract (DbC) is a software testing and correctness methodology. Principally, it uses preconditions, postconditions, asserts statements, and invariants. However, general-purpose language-based creates heavyweight code dependency while realizing the design-by-contract approach. A domain-specific language for this purpose is an elegant way to implement unit testing with test-driven development. In the aspect of the multi-agent systems, contracts may have states and changeable contexts so that developers can apply the design-by-contract into the blockchain-based domain-specific language. Actually, this programming approach is called defensive programming, because the application is responsible for figuring out what has occurred an error in postconditions or preconditions. For instance, when an agent planned a goal in a collaborative relationship, we should decide assumptions (precondition) and the effect of these assumptions (postcondition) that are valid. In this case, the effect should be the agent's goals. If the procedure that has been defined as a precondition is executed correctly, then it will terminate successfully to complete the given goal as achieved. Normally, there are a couple of ways to do it, but we will use Solidity stateful blockchain language to do so.

Embedded domain-specific language stands for incorporating a domain-specific language in a general-purpose language such as Java, Scala, or Kotlin. Despite that

it restricts language extensibility, in this study, we will use the advantage of a host language such as annotation-based code generation, runtime, or compile-time metaprogramming.

Last but not least, even though the solution consists of different types of languages such as an agent communication language, general-purpose language, and smart contract language, we believe that the embedded domain-specific language can ensure the design and analysis, and implementation layer compact in terms of agent deployment. We will reduce the complexity of the design, analysis, and implementation layer as much as possible and also maintain existing agent-oriented programming languages at the implementation layer.

3 Methodology and Expected Outcome

In the Ph.D. journey, we will focus on the different domain-specific language approaches. To this end, we will use the simplified methodology called Prometheus [17] methodology as shown in [1]. We will separate the research into layers to create a simple application.

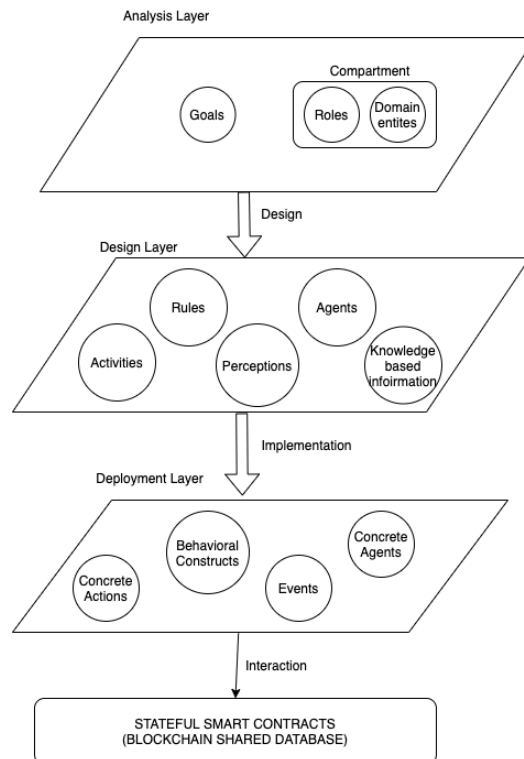


Fig. 1. Concept of Role-based Blockchain-Enabled Agent Programming

In the Analysis Layer, we should define domain entities and roles regarding the application. This will be likely a planning phase for large-scale applications when each agent involves in the network, there will be a dynamic model to create roles, goals, and domain entities. This step is a necessity to abstract computation behavior for realizing as software agents or programs. In the analysis layer, we need to implement the role-object pattern for agent applications because role-oriented languages have no direct connection with object-oriented applications.

We will design the agent features without considering the environment that is uncertain, unpredictable, time-sensitive, and highly dynamic. Having identified these abstraction characters of the planned environment, let us give a desirable definition of agents to perform in the implementation layer. The design and Analysis layer is a kind of requirement engineering phase because we should understand and specify the requirements of the given environment through a dynamic model. As shown in Figure 1, we can assign rules, activities, and perception features with an agent object so that we can implement role-oriented programming features such as constraints and relationships between agents. Agent rules can be assigned with beliefs (Information about the environment) and desires (agent's wishes), which means that the agent can define objectives by starting from an internal state to accomplish a goal.

Deployment of a software agent contains fundamental features such as communication languages, software components, users, hardware elements connecting to software agents. The deployment layer is closely related to the implementation layer of software agents, which is why practical frameworks follow up the design pattern that provides a recurring solution in a particular design problem. Design patterns such as Belief-Desire-Intention (BDI) or Reactive Agent Frameworks have no restrictive specifications and it has not been implemented with the design-by-contract approach. However, most of the agent-oriented practical frameworks have followed the BDI approach, but agent frameworks do not have to be dependent on the BDI approach. Agents should be synchronized with behaviors that show concurrent operations such as atomicity, thread prioritization, and lock-based synchronization.

As shown in Figure 1, agents may have interaction with a database in order to keep data in persistent storage. Blockchain technology provides smart contracts (programmable code snippets that work in the blockchain database) and database functionality in a deterministic way. Determinism means that a copy of a particular blockchain database should work in the exact same way in another environment. The database can accept stateful (Turing complete) or stateless (non-Turing complete) smart contracts to operate transactions from agent applications. Once an agent triggered action, the action performs a transaction into the blockchain database. Security and privacy of agent smart contracts can be provided by Merkle trees that present zero-knowledge proof and verifiable data structure.

Agent-Oriented Libraries and Frameworks can be applied to role-object pattern in order to connect between agents's and roles' world. In this proposal, we implement role and natural types in a stateful contract with a contract wrapper as below:

4 Proposed Solutions

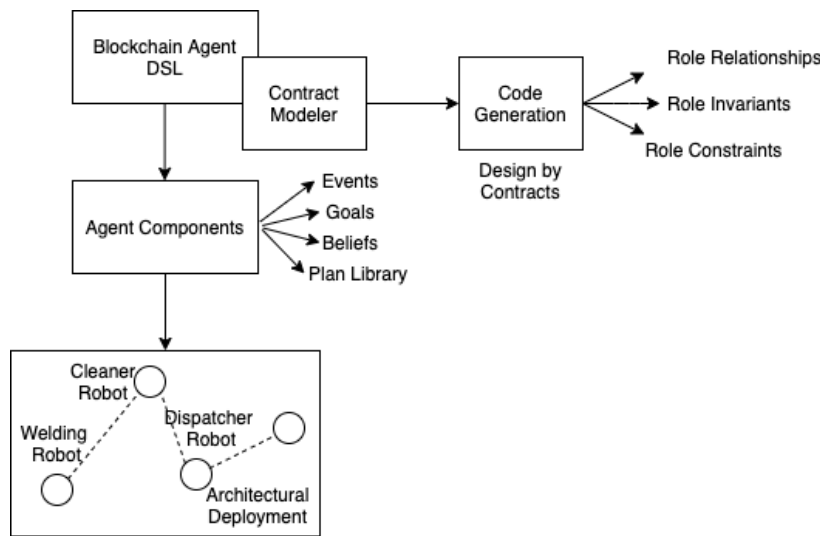


Fig. 2. System Architecture of Role-based Blockchain Enabled Language

In order to ensure all kinds of functionalities in Figure 2, we create a domain-specific language runtime interpreter with existing technologies. In the first step of implementation of the proposed language, we would like to focus on a stateful contract language, which is called Solidity, Agentspeak with Jason, and parser generator such as Another Tool for Language Recognition (ANTLR). We are planning to do language translation with various technologies such as ANTLR so that we will create the Abstract Syntax Tree to transform into the Concrete Syntax Tree. Due to the nature of agent-oriented programming, it seems to have a necessity of runtime-metaprogramming, which is allowing us to generate code from meta-objects at the runtime, we might have a solution with compile-time metaprogramming. In this case, one of the biggest challenges is to select general-purpose language because the language can identify annotations either is in a runtime-time or compile-time role type checking.

```

public static final String FUNC_ADDROLES = "addRoles";
public static final String FUNC_GETADDRESSES = "getAddresses";
public static final String FUNC_GETALIAS = "getAlias";
public static final String FUNC_REMOVEROLES = "removeRoles";

public RemoteFunctionCall<TransactionReceipt> addRoles(String addr, String alieases) {
    final org.web3j.abi.datatypes.Function function = new org.web3j.abi.datatypes.Function(
        FUNC_ADDROLES,
        Arrays.<Type>asList(new org.web3j.abi.datatypes.Address(160, addr),
            new org.web3j.abi.datatypes.Utf8String(alieases)),
        Collections.<TypeReference<?>>emptyList());
    return executeRemoteCallTransaction(function);
}

function addRoles(address addr, string memory alieases ) public{
    _addresses[msg.sender].push(addr);
    _roleAliases[msg.sender][addr] = alieases;
}

```

Fig. 3. Code Snippet from Role-Object Pattern in a Stateful Smart Contract

```

private final static String PRIVATE_KEY = "d9c5e395970994db652a448d6d192e011bd57505010178a811ceb010ea101f21";
//TODO: Ganache-cli changes the private keys and addresses each time when it operates again.

private final static BigInteger GAS_LIMIT = BigInteger.valueOf(6721975L);
private final static BigInteger GAS_PRICE = BigInteger.valueOf(20000000000L);

private final static String RECIPIENT = "0xdDd207a94a527446C1e9AC5f5009F8515Dea8f1\n" +
    "";

```

Fig. 4. Data from the Merkle Hash Tree in the Blockchain

The red rectangle In Figure 3, the address aliases that have been specified for smart contract functions need to call contract wrappers that were written in a general-purpose language. The blue rectangle in Figure 3 shows a procedure from a stateful smart contract that adds address and provides constraints checking. The green rectangle in Figure 3 demonstrates a contract wrapper in a general-purpose language in order to control contract address from Java language. In this example, we would like to simulate role attributes by adding role types into a smart contract. Role types and natural types can be represented in the object-oriented data structures and we can use smart contract addresses in order for reaching out to contracts in the blockchain consensus network. Smart contracts provide trust and security because the data will be shown as below in Figure 4. TX value represents a transaction in the blockchain that has been processed by one of the smart contracts. Private keys are assigned to accounts as shown in Figure 4 and developers can use accounts like a shared memory to realize limited concurrent applications. In the end, all values are in a Merkle tree through the hashed data structure. Moreover, one of the important security features of a distributed system is the single point of failure can be prevented by gas costs. A typical gas cost consists of an operation and a transaction cost that can prevent the consumption of general system resources to the end.

5 Limitations

In this section, we will list our limitations regarding the process of writing the thesis.

- We will present examples regarding autonomous and collaboration features. In context with collaboration, supply chain simulation between participants would be enough. As for the autonomous feature, robotic motion planning can be simulated with our proposal.
- We will focus on the existing meta-model such as CROM for specifying communicative entity types. In the context of CROM research, we will follow the guideline regarding roles and compartments that have been specified before.
- We will develop an application based on a stateful smart contract language such as Solidity the following design-by-contract approach that interacts with different agents in the context of role-oriented programming.
- This research is limited to the KQML and FIPA agent communication languages and it does not comprise stateless blockchain language. Due to the nature of the stateless blockchain language, it does not purely suitable for the object-oriented approach.
- As for ontological representation, semantic heterogeneity between agent-oriented frameworks will not be taken into consideration. So we basically will handle existing ontologies and will not advance to ontology engineering.

6 State of the Art

When we conduct a literature review, we have been investigated the following two literature research questions (LRQ):

- *LRQ 1*: How can Blockchain and Multi-agent System improve each other?
- *LRQ 2*: How does role-based programming affect collaborative multi-agent systems?
- *LRQ 3*: Can the context of agents be an affiliated aspect of role-oriented programming through blockchain technology in multi-agent systems?

ALAADIN is one of the oldest metamodels to define models of organizations for agents and this model defines a very simple description of coordination and negotiation schema [18]. The authors determine that the role is an abstract representation of an agent or service function within a group. Groups are a set of features that behave as an atomic entity so that an agent dynamically joins, creates, or leaves groups [19].

When we focus on behavioral roles for agent interaction, (Cabri et. al. 2003) proposed that an agent system defines a role as a set of capabilities and expected behaviors. BRAIN is an approach that covers a role-based interaction model, where agents' interactions and behaviors are embedded in roles [19]. Moreover, they achieved and advise to realize agent-oriented features, separation of concerns, and reuse of solutions [20]. To describe agents semantically, they defined a language called XRole that exploits built-up definitions of roles. These definitions consist of name, description, addresses, role description, and contents of the agents with relational features such as *MinOccurs* and *MaxOccurs*. RoleSystem is an interaction infrastructure that implements the model of BRAIN [21]. Roles defined by XRole can be read by humans as well as by agents and tools [21]. The RoleSystem provides two main components which are: *reqRegistration*, to register an agent in the system with a specified role; *searchForRoleAgent*, to search for agents playing a given role between agents and server agents [21].

The planning capability of multi-agent systems is one of the key features that the blockchain should take care of it. After assigning roles, plan execution of the multi-agent systems should complete distributed ordering actions. To do so, a smart contract can be used which are essentially collections of distributed code and data representing some business logic that works with the blockchain distributed consensus protocols [22]. The main idea of this paper is to coordinate the steps of multi-agents through the smart contracts aspect of distributed plan execution. In this plan execution, multiple smart contracts can be used such as oracle contract, which is allowing to exploits data in the off-chain storage, or contract of preconditions and postconditions to provide the design-by-contract pattern.

Gaia is one of the methodologies at the design and analysis phases in multi-agent systems. The main goal of this methodology is to model multi-agent systems for an

organization where different roles interact [19]. The Gaia methodology defines the features of roles as below:

- *Responsibilities*: They specify the functionalities of agents that play roles.
- *Permissions*: They are a set of rights associated with roles in which agents play.
- *Activities*: Internal computation of an agent. This does not take into consideration the relationship between agents.
- *Protocols*: This is related to interaction roles indicating agent-to-agent communication.

The role-based evolutionary programming (RoleEP) presents cooperative mobile agents to collaborate in achieving a common goal [19]. The authors of RoleEP state that an object becomes an agent by binding itself to a role that is defined in a dynamic environment [5]. The authors have defined the basic concept as below [5].

- *Environment*: An environment is composed of environment attributes, methods of environment, and roles.
- *Roles*: A role, which can move between hosts that exist in an environment, contains role attributes, role methods, and binding interfaces.
- *Objects*: An object, which cannot move between hosts, is composed of attributes and methods.
- *Agents*: An object or mental identity that binds itself with some roles and acquires traveling/collaboration functions.
- *Binding Interface*: A binding interface, which looks like an abstract method interface, is used when an object binds itself with a role.

Implementation of a domain-specific language may have metamodel design paths at the level of M1 (User Model), M2 (Unified Modeling Language), M3 (Meta Object Facility). For instance, AgentDSL is a domain-specific language for cross-cutting concerns for agents, which is supporting aspect-oriented programming, and non-crosscutting concerns [23]. The authors of AgentDSL defines a code generator that maps abstractions.

7 Research Plan

During this research, we will try to answer the research questions that we have asked in the Research Problem.

- In the first year, we will deal with the design and analysis phase from the previous studies that have been conducted by various researchers from the department of Role-based software infrastructures for continuous-context-sensitive-systems (ROSI) at TU Dresden. Roles, compartments, negotiation, and collaboration parameters will be defined and domains of case studies may expand or narrow down.

- In the second year, the design of agent architecture, topology, and sample applications of the domain-specific language will be proposed. A prototype will be shown in accordance with the supervisorship' requirements.
- In the third year, the implementation layer will be completely finished, and then we will agree on a final version of the thesis with the supervisor. If we have enough time at this stage of the research, we will develop the implementation further for the practical solution with regard to the robotic applications.

At the end of our Ph.D. journey, it is believed that developers or experts can implement role-based agent-oriented applications in the blockchain network by having an embedded domain-specific language.

The complexity of this research can easily increase because a couple of approaches should be used in the end. However, we have limited the approach with collaborative software agents either can work on hardware solutions or enterprise applications.

8 Case Study

Manufacturing scheduling is the process of assignment of timing for order, manufacturing, and delivery. So we should provide a good quality per unit and the number of units should be maximized per slot in the production line. At the same time, we need to minimize the waste of resource requirements and potential failures. Moreover, the designed system sometimes collaborates with human operators because they need to get involved in some complex problems by collaborating with robotic cells [24].

Another case study that we want to focus on the collaborative multi robots scenario. Let us assume we have two robot agents and a human agent. The human agent should work with two robot agents. The first robot agent will do actions picking material from an assembly line, finding the next slot, dropping the material, and moving towards a new position, respectively. So the second robot agent will just do action welding with the material into some raw good. Human-agent is going to check the material quality before welding it. Welding and moving can be goals for us and they need to have preconditions and postconditions. The actions of robots are pick(), find(), drop(), and goal of the robots can be the result of checking slots. We will put the goals and actions into the smart contract language with their data and then we will evaluate in terms of preconditions and postconditions. We can use these features in a domain-specific language that has been generated from an agent communication language, a general-purpose agent framework, and a stateful smart contract language.

9 Result of the Research Study

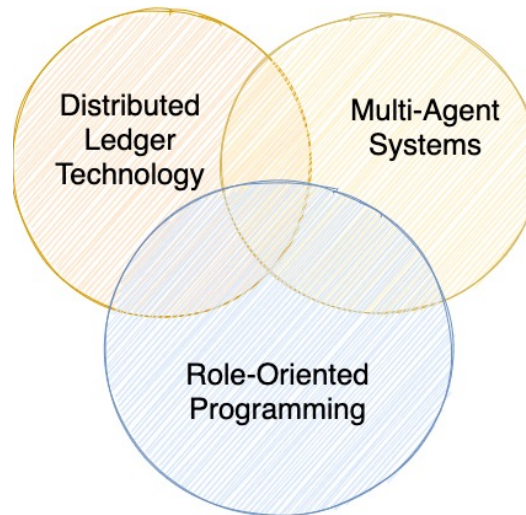


Fig. 5. Research Overview regarding different approaches

As readers can see in Figure 5, we are bringing together three fields, which are blockchain technology (BCT), multi-agent systems, and role-oriented programming, into a domain-specific language. As shown in Figure 5, we have demonstrated blockchain technology as distributed ledger technology. Aside from it being implementation-agnostic, a distributed ledger can be in a form of private, public, federated, or consortium networks.

In the distributed ledger part, we have focused on smart contract development that can give us the ability to develop an application in a decentralized environment. Multi-agent systems are suitable for the decentralized environment and they can use role-oriented attributes. Contract-oriented approach with smart contracts can benefit from role-orientation, which principally represents relations in a given model implementing an embedded domain-specific language. The language supports the following distinctive attributes:

- Agents that were created by the template language can connect to each other in a peer-to-peer manner.
- Agents can behave autonomous and partly proactive by having relationships between roles. In addition, applications that have programmed by the domain-specific language is easy to use by making certain of obtaining belief-oriented architecture and action-oriented program.

- Agents can collaborate and develop social behavior in the context of role-oriented programming by keeping the data secure with smart contracts.

10 Conclusion

The paper addressed the challenge of compelling trust and security in multi-agent systems and their role-oriented features by realizing smart contracts regarding blockchain technology (BCT). The main purpose of the research is to give a new approach to the intersection between smart contract programming, role-oriented programming, and agent-oriented programming. The course of findings among various research areas guides us to design a domain-specific language to contribute to the multi-agent system area.

- There is no common understanding in terms of multi-agent system methodology, analysis, design, or implementation. This increases the complexity of the research in the multi-agent system area.
- The limited number of domain-driven agent-oriented languages have been provided so that one can notice that multi-agent system research is most likely conceptual and it does not provide prototype and result-evaluated research.
- Synthesized metamodeling from scratch in different research areas can be ambiguous; thus, we believe that embedded domain-specific language with blockchain can solve most of the problems for multi-agent systems.
- Role-oriented programming with smart contracts is challenging because the choices of technology can affect the result of the study. For instance, stateful and stateless contracts are not advanced technologies that can employ all of the features of the object-oriented paradigm. Turing complete and non-Turing complete technologies will be scrutinized in future work.

In a nutshell, this paper presented a new significant role approach with smart contract programming implementing hash data structure and providing data security regarding roles. Presenting our approach will simplify the application development process for further researchers.

11 Future Work

In future work, a tool will be developed for a role-based multi-agent system. This tool includes an annotation processor and template-based code generator for agent behaviors. By selecting a general-purpose language, the system will be evaluated with qualitative and quantitative tools. Smart contracts will be generated through annotation processing with customized annotations and agents will be generated with a template-based code generator tool for one of the selected frameworks which have been presented in the introduction section.

Acknowledgements

The author would like to thank his supervisors, Prof. Dr. Uwe Aßmann, and Prof. Dr. Susanne Strahringer, for the patient guidance, encouragement, and comments they have provided to shape his doctoral vision. This work is funded by the German Research (DFG) within the Research Training Group Role-Based Software Infrastructures for continuous-context-sensitive Systems (GRK 1907, TU Dresden, Software Technology Group, Nöthnitzer Straße 46, 01187, Dresden).

References

1. Y. Shoham, “Agent-oriented programming,” *Artificial Intelligence*, vol. 60, no. 1, pp. 51–92, 1993. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0004370293900349>
2. H. Zhu and M. Zhou, “Role-based multi-agent systems,” *Personalized Information Retrieval and Access: Concepts, Methods and Practices*, 01 2008.
3. T. Kühn, M. Leuthäuser, S. Götz, C. Seidl, and U. Aßmann, “A metamodel family for role-based modeling and programming languages,” in *Software Language Engineering*, B. Combe-male, D. J. Pearce, O. Barais, and J. J. Vinju, Eds. Cham: Springer International Publishing, 2014, pp. 141–160.
4. T. Kühn, S. Böhme, S. Götz, and U. Aßmann, “A combined formal model for relational context-dependent roles,” in *Proceedings of the 2015 ACM SIGPLAN International Conference on Software Language Engineering*, ser. SLE 2015. New York, NY, USA: Association for Computing Machinery, 2015, p. 113–124. [Online]. Available: <https://doi.org/10.1145/2814251.2814255>
5. G. Cabri, L. Ferrari, L. Leonardi, and F. Zambonelli, “A survey about role-based interaction proposals for agents,” 01 2005.
6. B. Ferreira and A. M. Leitão, “Context-Oriented Algorithmic Design,” in *7th Symposium on Languages, Applications and Technologies (SLATE 2018)*, ser. OpenAccess Series in Informatics (OASISs), P. R. Henriques, J. P. Leal, A. M. Leitão, and X. G. Guinovart, Eds., vol. 62. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018, pp. 7:1–7:14. [Online]. Available: <http://drops.dagstuhl.de/opus/volltexte/2018/9265>
7. K. Wan and V. Alagar, “A context-aware trust model for service-oriented multi-agent systems,” vol. 5472, 04 2009, pp. 221–236.
8. F. Bellifemine, G. Caire, and D. Greenwood, “Developing multi-agent systems with jade,” *Developing Multi-Agent Systems with JADE*, pp. 1–286, 02 2007.
9. A. Pokahr, L. Braubach, and W. Lamersdorf, *Jadex: A BDI Reasoning Engine*. Boston, MA: Springer US, 2005, pp. 149–174. [Online]. Available: https://doi.org/10.1007/0-387-26350-0_6
10. R. Bordini, J. Hübner, and M. Wooldridge, *Programming Multi-Agent Systems in AgentSpeak Using Jason*, 10 2007, vol. 8.
11. K. Hindriks and J. Dix, *GOAL: A Multi-agent Programming Language Applied to an Exploration Game*, 03 2014, vol. 9783642544323, pp. 112–136.
12. M. Winikoff, *JackTM Intelligent Agents: An Industrial Strength Platform*, 01 2005, pp. 175–193.
13. O. Boissier, R. H. Bordini, J. F. Hübner, A. Ricci, and A. Santi, “Multi-agent oriented programming with jacamo,” *Science of Computer Programming*, vol. 78, no. 6, pp. 747 – 761, 2013, special section: The Programming Languages track at the 26th ACM Symposium on Applied Computing (SAC 2011) & Special section on Agent-oriented Design Methods and Programming Techniques for Distributed Computing in Dynamic and Complex Environments. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S016764231100181X>
14. M. Dastani, F. Dignum, and J.-j. Meyer, “3apl: A programming language for cognitive agents,” 01 2003.
15. M. Dastani, “2apl: a practical agent programming language,” *Autonomous Agents and Multi-Agent Systems*, vol. 16, no. 3, pp. 214–248, Jun 2008. [Online]. Available: <https://doi.org/10.1007/s10458-008-9036-y>
16. H. Zhu, “Role-based autonomic systems,” *IJSSCI*, vol. 2, pp. 32–51, 01 2010.
17. R. Bordini, M. Dastani, and M. Winikoff, “Current issues in multi-agent systems development,” vol. 4457, 09 2006, pp. 38–61.
18. *Aalaadin*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 169–179. [Online]. Available: https://doi.org/10.1007/3-540-26815-4_8
19. G. Cabri, L. Leonardi, L. Ferrari, and F. Zambonelli, “Role-based software agent interaction models: A survey,” *Knowledge Eng. Review*, vol. 25, pp. 397–419, 12 2010.

20. G. Cabri, L. Leonardi, and F. Zambonelli, "Implementing role-based interactions for internet agents," 02 2003, pp. 380–387.
21. Cabri, Giacomo and Leonardi, Letizia and Zambonelli, Franco, "Brain: A framework for flexible role-based interactions in multiagent systems," vol. 2888, 11 2003, pp. 145–161.
22. A. Shukla, S. K. Mohalik, and R. Badrinath, "Smart contracts for multiagent plan execution in untrusted cyber-physical systems," in *2018 IEEE 25th International Conference on High Performance Computing Workshops (HiPCW)*, 2018, pp. 86–94.
23. U. Kulesza, A. Garcia, C. Lucena, and P. Alencar, "A generative approach for multi-agent system development," vol. 3390, 05 2004, pp. 52–69.
24. A. Bauer, D. Wollherr, and M. Buss, "Human-robot collaboration: a survey," *Int. J. Humanoid Robotics*, vol. 5, pp. 47–66, 2008.

Author

Orçun Oruç received M.Sc. from TU Chemnitz, and he graduated from Kocaeli University with a B.Sc. degree. Currently, he is pursuing his Ph.D. in Computer Engineering-Software Technology at the Dresden Technical University. His research interests include programming languages, multi-agent systems, role-oriented programming, natural language processing, decentralized, and distributed applications.

FEDERATED IDENTITY MANAGEMENT (FIDM) SYSTEMS LIMITATION AND SOLUTIONS

Maha Aldosary and Norah Alqahtani

Department of Computer Sciences, Imam Mohammad Ibn
Saud Islamic University, Riyadh, KSA

ABSTRACT

Efficient identity management system has become one of the fundamental requirements for ensuring safe, secure, and transparent use of identifiable information and attributes. FIDM allows users to distribute their identity information across security domains which increase the portability of their digital identities. However, it also raises new architectural challenges and significant security and privacy issues that need to be mitigated. In this paper, we presented the limitations and risks in Federated Identity Management system and discuss the results and proposed solutions.

KEYWORDS

Federated Identity Management, Identity Management, Limitations, Identity Federation.

1. INTRODUCTION

Federated Identity Management (FIDM) is a concept that helps to link user's digital identities and attributes stored in several sites also allows cooperation on identity processes, policies, and technologies among various domains to simplify the user experience. FIDM typically involves Identity Providers (IdPs) and Service Providers (SPs) in a trust structure called Circle of Trust (CoT) based on a business agreement where all the identifiable information of users are federated at a central location such as the Identity Provider IdPs who is responsible to pass authentication tokens to SPs, and SPs after that provide their resource to the user. FIDM is considered a promising approach to facilitate secure resource sharing among collaborating participants in heterogeneous IT environments [1].

Many advantages demonstrated by Federated Identity Management systems such as reduce the cost provide convenience for the users and interoperability among Identity Management systems in addition to support single sign-on SSO service and other valuable services. However, it has limitations that provide several real security and privacy risks Due to the valuable information shared across domains in The FIDM using loosely coupled network protocols. The risks and limitations in FIDM require to be introduced and explained to find Appropriate solutions to mitigate these risks.

In this paper, we discussed the concept of personal identity in a real-world and digital identity as a prelude to the identity management systems. The notion of Identity Federation was discussed in this work as well some Federated Identity Management Architectures such as Liberty Alliance,

Security Assertion Mark-up Language SAML V2.0, WS-federation, and Shibboleth, etc. In this paper, we presented the limitations of Federated Identity Management based on how it affects the user. Finally, we discussed the solutions proposed to mitigate the risk of these limitations.

This paper is organised as follows: Section 2 gives background and basic information that needs to be understood before discussing the FIdM system. The concept of identity federation and the number of architectures that implement FIdM is given in section 3. Section 4 presented the limitation and risks in the FIdM. In section 5 provides a discussion of the solutions before the paper is concluded in section 6.

2. BACKGROUND

2.1. Identity

Human identity is a representation of an individual by several properties which indicates that person, reflecting its uniqueness, and distinguish that person from others. These properties could be intrinsic (e.g. DNA, retina scan, fingerprint), descriptive (e.g. name, birthplace, birthdate), demographic (e.g. gender, occupation), geographic (e.g. country, address, postcode) or psychographics (e.g. preferences, interests).[2]

The identity of an individual consists of a large number of personal properties. All subsets of the properties form partial identities of the person.[3] The person may have multiple different partial identities depending on the context. These partial identities could relate to roles the person plays. Identity involves all the primary characteristics that make each person unique but also all the characteristics that enable belonging to a particular group as well as established position within the group [4].

In today's world, living and working in the networked environment requires digital identity for each individual, it has allowed us to interact, transact, communicate, share reputations, and create trusted relationships with devices, people, and business electronically. Digital identity is the representation of identity in a digital system, Roussos et al [4] describe the digital identity as the electronic representation of personal information of an individual or organization (name, phone numbers, address, demographics, etc.).

Despite that there is a strong association between real life and digital identity, digital identity breaks from the restriction of everyday life, allowing users to exceed the boundaries of the real world[5]. clarify that digital environments granted the users the chance to get rid of the human qualities of age, race, gender, and disability.

2.2. Identity Management

Identity management (IdM) is defined as a set of procedures, policies and technologies that help authoritative sources as well as individual entities to manage and use identity information, it also provides access and privileges to end-users through authentication schemes [6]. Identity management procedures include management of the identity lifecycle, management of identity information, and management of entity authentication as an initial step for authorization.

Identity Management responsible for handling the lifecycle of identity, its creation, maintenance and eliminating a digital identity, by providing the credentials and means for identification during the preparatory process, through to authenticating and authorising access to resources, and to revoking access credentials and identities. Identity management is a crucial part of many security

services since it assures user legitimacy. Therefore, identity management is an integral part of any access management system [7].

There are numerous technologies, services and terms related to identity management such as Directory services, Service Providers, Identity Providers, Digital Cards, Digital Identities, Web Services, Access control, Password Managers, Single Sign-on, Security Token Services, Security Tokens, WS-Trust, WS-Security, OpenID, OAuth, SAML 2.0 and RBAC.

Identity management is particularly used to authenticate a user on a system and make certain whether that user is allowed or unauthorised to access a particular system. IdM also covers issues such as how users obtain an identity, the protection of that identity and the technologies supporting that protection. Digital identity management technology is an essential function in enhancing and customizing the network user experience, protecting privacy, underpinning accountability in transactions and interactions, and respecting regulatory controls [8].

3. IDENTITY FEDERATION

Federated identity management (FIdM) is when multiple enterprises allow individuals to use the same identification information or login credentials to obtain access to the services or networks of all the enterprises in the group. The partners in a FIdM system are accountable for authenticating their users and for insuring for their access to the networks.

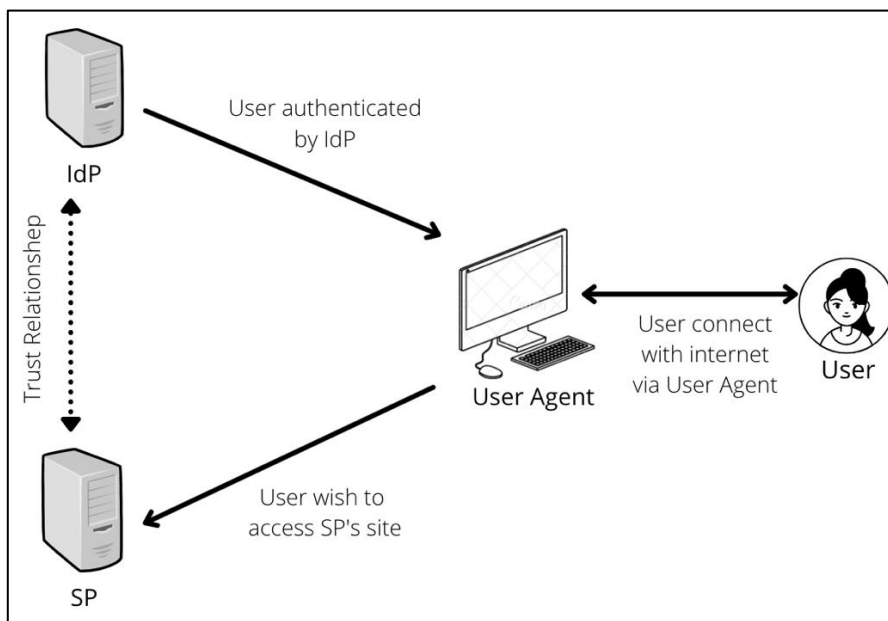


Figure 1. Component of FIdM system

The federated identity model includes four logical components: [9]

- A user is a person who acquires a specific digital identity to interact with an online network application.
- The user agent is a software application or browser that runs on any device such as PC, mobile phone and medical device. The online interactions of a user always take place via an agent, which can allow identity information flow or mediate it.

- The service provider (SP) site is a Web application that offloads authentication to a third party, which also might send the SP some user attributes. Because the SP depends on external information, it's often called a relying party (RP).
- The identity provider (IdP) is a Web site that users log in to and that occasionally stores attributes of common interest to share with several SPs.

In a Federated identity management system, the user might have one or more local identities issued by service providers (SPs), in addition to a single identity issued by the identity provider (IdP) within a specific domain called a circle of trust (CoT). A standard CoT composed of a single IdP and multiple SPs. In CoT, the IdPs must be trusted by all the SPs within it. Each SP could be a member of more than one CoT. A user can federate its IdP-issued identity with the local identities issued by SPs within the same CoT. [7]

With FIdM the user's credentials are always stored by the IdP. When a user registers into service, they do not have to provide their credentials to any of the SPs. Instead of authenticating directly with the user, the SP trusts the IdP to verify the user's credentials. The IdP then authorises the user to the application of SP, and the user is then allowed to access the service. Therefore, the user in FIdM systems never provides their credentials to anyone but the IdP.

FIdM presents numerous benefits to the various stakeholders, it offers users the single sign-on (SSO) capability that allows them to proceed between the various SPs with no need to authenticate or login again, it allows SPs to offload the cost of managing user attributes, passwords and login credentials to trusted IdPs, it provides scalability, allowing SPs to provide services to a greater number of users, it allows IdPs to maintain close relationships with end-users and sell them more services, as well as extract fees from the SPs they support [10].

Table 1. A Comparison between FIdM and SSO

| | FIdM | SSO |
|-----------------------------------|--|--|
| Single access | To multiple system across various organization | To different services within a single organization |
| User credential | Given only to Idp | Given to any system the user logging into |
| Log-in to several services | Allow | Allow |
| Use of the same credential | Allow | Allow |
| Authentication process | Only once in the same working session | Only once in the same working session |
| Identity federation | Supported | Not supported |

FIdM has aspects that are similar to single sign-on (SSO), but they are different at their core. FIdM gives you SSO, but SSO does not necessarily give you FIdM. Despite that the SSO and FIdM both allow users to log in to several services using the same login credentials, there are two things that FIdM does that SSO cannot: Firstly, SSO allows users to access multiple systems only within a single organization, whereas FIdM allows users to log into systems across various organizations. Secondly, FIdM is more secure than SSO. For SSO, the user credentials are still being provided to any system that the user is logging into. While with FIdM, user's credentials are only given to the IdP exclusively.

Certainly, FIdM depends heavily on SSO technologies to authenticate the users across diverse websites and apps, however it has advanced these technologies further. Therefore, while FIdM does provide users SSO, SSO does not offer all of the benefits that FIdM does. Table 1 present a comparison between FIdM and SSO.

Federated identity management presents economic and convenience advantages to both the users and the organizations that employ it. However, there are some serious security considerations, techniques like strong authentication must be implemented for a secure SSO because SSO system may introduce a single-point-of-failure. [9] Moreover, FIdM requires a lot of trust and open communication between partners that choose to make use of it. Organizations that are considering creating or joining an identity federation need to assure that they agree upon all factors. [10]

3.1. Federated Identity Management Architectures

3.1.1. Liberty alliance

Liberty Alliance is a project presented first in 2001. according to the official web site of the project [11], it is a consortium of more than 150 member includes governments and companies from around the world. The consortium is committed to creating an infrastructure that provides support for all existing and emerging network access devices and has defined interoperability requirements developing an open standard for federated network identity for products that meet its specifications. The specifications developed by the Liberty Alliance Project enable individuals and organizations to control their identity information securely also it is providing conveniently by supporting single sign-on (SSO) service which is the service that enables users to interact with different service providers or Web sites with trust relationships by signing in just once.

The main objectives of the Liberty Alliance Project Specifications are to Serve as open standards for SSO, management of federated identity, and web services. Also, it aims to promote permission-based sharing of personal identity attributes and Enable consumers to protect their network identity information. Additionally, aims to create an open network identity infrastructure that supports all current and emerging user agents.

The specifications in the Liberty Alliance are enclosing the following components: Liberty Identity Federation Framework, Liberty Identity Web Services Framework, Liberty Identity Service Interface Specifications, Schema Files and Service Definition Documents, and Support Documents. They are developed to enable federated network identity management. Using web redirection and open-source technologies such as SOAP and XML, they enable distributed, cross-domain interactions [12][13].

For more information about Liberty Alliance: [14] [15] [16] [17]

3.1.2. Shibboleth

Shibboleth is a project created first by US-based Internet2 in 2003. It developed an open-source, standards-based system that provides access management for individuals to a resource depending on their role instead of their identities which means that Role-based attributes are used in the Shibboleth system. Shibboleth allows the affiliated institution of the user to authenticate the user to permit access to on-campus applications and the resources licensed by the library from service providers [18]. to protect the user's privacy, Shibboleth sends anonymous identification to the service provider.

Additionally, Shibboleth provides authorisation service that helps sites to make decisions for individual's access and privileges in online resources by transport the role attributes securely between the Identity Provider site (affiliated institution) and Resource Provider site to determine whether the user has a right to access the resource or not. Single sign-on (SSO) feature is supported in the Shibboleth system which makes the system more flexible and convenient.

For more information about Liberty Alliance: [14] [19] [20] [21]

3.1.3. WS-Federation

The Web Services Security Framework is an identity management approach proposed by International Business Machines Corporation and Microsoft Corporation with other companies. As in Liberty Alliance, they provide several specifications such as WS-Security, WS-Trust, and WS-Security Policy. These specifications determine how to control the assertions (security tokens) that contain identifiable information about the user and issued by an identity provider. The security tokens help the service provider SP to decide to wither or not the user have a right to access the service resource.

According to [22] [23], WS-Federation builds upon the base WSS specifications to define mechanisms which enable resources to be shared securely between different domains. The specifications introduce many services include security token service STS which is the IdP service that issues identity tokens to users based on their authentication. Authorisation service which decides giving access right to the user.

For more information about WS-Federation: [24] [25]

3.1.4. Security Assertion Mark-up Language SAML V2.0.

The first release of the Security Assertion Mark-up Language SAML was in 2002 by the Organization for the Advancement of Structured Information Standards OASIS. SAML is standard based on Extensible Mark-up Language XML helps to manage the authentication and authorisation processes between identity providers and service providers. The SAML system includes four main concepts which are assertions, protocols, profiles and bindings. Where assertion is the declaration user information asserted by the identity provider IdP for a service provider SP. The SAML protocol is helping to determine the rules on how to embed the SAML elements inside the request/response packet and on how to process them.

Transporting protocol messages using existing widely deployed communication protocols like HTTP (Hypertext Transfer Protocol) or SOAP (Simple Object Access Protocol) was described by The SAML bindings specification (SAMLBind). besides, The SAML profile specification (SAMLProf) provides many profiles that describe how the SAML elements can be used to implement a use case and achieve interoperability. [26] [27] [28]

3.1.5. Other Architectures

In addition to the FIdM Architectures that we talked about above, there are other federated architectures that designed originally for relatively simple applications such as OpenID [29] Which is open source user-centric and decentralized Identity management system. OpenID connect [30] Which is a simple identity layer on top of the OAuth 2.0 specifications family. It is the third generation of OpenID technology. It helps the SP to authenticate the End-User based on the authentication performed by an Authorisation Server, as well as to obtain user attribute in an interoperable and REST-like manner. Besides, SCIM (System for Cross-domain Identity

Management) Which is a specification designed for cloud-based applications to manage user identities and services [31].In Table 2, we provide a comparison between Liberty alliance, Shibboleth, Security Assertion Mark-up Language SAML V2.0., WS-Federation, OpenID and OpenID connect about the target area, storage of Identity information, Single Sign-On, Single Log-Out, Identity Mapping, Security Tokens and Access to web applications. [32] [33]

Table 2. Comparison Between FIdM Architectures

| | Liberty alliance | SAML V2.0. | WS-Federation | Shibboleth | OpenID connect | OpenID |
|-------------------------------------|--|--|---|---|--|---|
| Identity mapping | By opaque identifiers | Via pseudonym service | Via pseudonym service | By short-term random IDs | Using (STS) chains and JavaScript mapping | Using (STS) chains and JavaScript mapping |
| Area targeted | Business interactions | Business interactions | Business interactions | Digital academic resource sharing | Developer and programmer | Supporting developer and programmer |
| Identity information storage | User info could be distributed and federated | User info could be distributed and federated | User info could be distributed and federated | Centrally located and only attributes sent to SP | Attributes and info are distributed IdP | Attributes and info are distributed IdP |
| Single sign-on | Supported | Supported | Supported | Supported | Supported | Supported |
| Single log-out | Supported | Supported | Not supported | Not supported | Supported | Supported |
| Security tokens | Extends SAML assertions for Communicating authentication And authorisation security Tokens between providers | Extends SAML assertions for Communicating authentication And authorisation security Tokens between providers | Builds on ws-security's Profiles and Kerberos | Extend the IdP to support info card profiles using SAML assertions as security tokens | Use json security tokens (json web token) to communicate user attributes | Use json security tokens (json web token) for user attributes |

| | | | | | | |
|--------------------|---|---|--------------------------------|--|---|--|
| Access to Services | Supports access of both Web Services and web applications | Supports access of both Web Services and web applications | Designed only for Web Services | Only supports access by web browsers to web apps | Support browser-based JavaScript , web app and native mobile apps | Support browser-based JavaScript, web app and native mobile apps |
|--------------------|---|---|--------------------------------|--|---|--|

4. LIMITATIONS AND DRAWBACKS IN FEDERATED IDENTITY MANAGEMENT SYSTEM

Federated Identity Management (FIdM) is a technique that allows the participating entities i.e., Service Providers (SPs) and Identity Providers (IdPs), to collaborate, on identity operations, technologies, and policies. FIdM also enables users of heterogeneous IT environments to share each other's resources. [34]. All the user identities in a FIdM system are federated at a central position, i.e., the Identity Provider (IdP). IdPs are responsible to proceed the authentication tokens to SPs, and after that SPs can provide their services to the requestor i.e., the user. It is also possible that the user has accounts with various IdPs, and the SP communicates with the relevant IdP for the set of attributes required. [35]. While FIdM is in general seen as a good thing, it does have some disadvantages. Based on how it effects user we determine the following limitations:

4.1. Trust

Any Federated Identity system is based fundamentally on mutual trust. The interactions in federated identity management systems occur only between pre-configured entities or closed circle of trust (CoT) due to the use of static establishment of trust which is the method where entities' trust relationship such as that between IdPs or SPs has to be pre-configured that done either during the registration phase to the system or via a trust negotiation process offline. Such limitation especially for a huge number of participating (IdPs and SPs) makes the system impractical, unscalable, and hard to establish trust relationship at runtime.[36] [37]

In any identity federation, each participating member must create and identify policies and security protocols which poses another challenge. Every member then is obligated to follow these rules, which may cause problems when various companies have different rules and requirements. Furthermore, since an organization can be a member of different federations, following these several policies and rules may become a challenge.

Current specifications of FIdM provides only the basic technical mechanisms to establish trust between participating members. However, they do not detail the requirements that need to be met before establishing these relationships.

4.2. Privacy

Privacy and data protection are a major concern in FIdM system due to personally identifiable information that shared between entities where the premier goal of FIdM models is to share identity attributes [34] there is no guarantee to prevent SPs and IdPs from misusing of identity information of users. Even though there are regulations such as [38] and [39] and privacy policies that protect the privacy of user's sensitive data but unfortunately there are no requirements to enforce these regulations and policies. Furthermore, many studies [35] [40][41] Proved that many SPs and IdPs sites are collecting, processing, and sharing data of users without user consent.

4.3. IdP discovery

The IdP discovery is the process of determining where authentication requests are going to be forwarded when a user wants to access an identity-based service. [40] One of the major significant security limitations in most of FIdM standards such as Shibboleth, Liberty and OpenID is that IdP discovery is performed on the SP server. This limitation could be exploited by a malicious SP to redirect a user to a web site masquerading as the IdP, which could then acquire the user's security credentials. [42]

Furthermore, FIdM systems rely on the constant communication between individual users and a centralized identity provider (IdP) for purpose of authenticating and grant authorisation. If the metadata used to authenticate a user to the IdP was compromised, through leaks, or any sort of attacks such as phishing attacks, an adversary would gain the same access to the federated identity provides to all other participating members.

4.4. Lack of Attribute-Aggregation Support

Another limitation of FIdM systems is that users can only choose one of their IdPs in any single working session with an SP, after that the IdP sends authentication and attribute assertion to the SP. Therefore, authorisation is restricted to a subset of the user's identity attributes. This isn't sufficient especially for Web-based services. There is a huge need for a mechanism that allows users to aggregate attributes from multiple IdPs in a single service session. This model could effectively help to protect the user's identifiers and prevents IdPs from exchanging data about users without their permission. However, each IdP still know that a federated user has several attributes at the other IdP. [43]

In Liberty, only one IdP can be queried in a single working session, and for any IdP in shibboleth, the authorisation framework only allows a single attribute authority (i.e. the Attribute Authority Service (AAS)) to be queried for user attributes. OpenID is also suffering from a lack of attribute-aggregation support. [42]

4.5. Complexity for the User

The usage of online services and transactions is growing every day, it is becoming necessary to grant the users and the service providers the tools they needed to make more transactions and expand the available services and the level of interaction and trust. [44] A drawback of FIdM based on SAML is the complexity of the protocol and resulting effort for configuration. Another limitation is the complexity for the user, especially because of the need from the user to choose their IdP at the Discovery Service (DS) and the users have to remember which federations they belong to, along with username and password. On the user side, the management of the identity is getting more complicated if the user uses multiple federations. [45]

4.6. Security

Identity theft is a serious concern in FIdM.[1] Security issues regarding a stolen identity will affect all federation partners, credentials (e.g. username and password pairs) must be protected in federated systems.

Common attacks are the impersonation attacks with stolen credentials. FIdM enabled systems to authenticate service requests by a security token attached to the request message. Therefore,

impersonation attack can also be conducted by stealing user's security token which has been authenticated, this token can be used to access resources in the federated environment. [34]

An important property of FIdM is single-sign-on (SSO). However, a crucial challenge was addressed by Madsen et al. [46] they claim that federated SSO makes the job of attackers easier. That because after the attackers conduct a successful identity theft within a federation, they could compromise resources of all federated SPs, which leads to exposure of critical data.

Another important aspect is message security, Improper message security result in concerns for identity theft. Regarding identity management, techniques to protect message confidentiality and integrity are crucial to protect sensitive identity attribute and prevent modification of identity attributes. According to Maler and Reed [40]. systems are vulnerable if it does not provide security tokens to service request messages, through digital signatures, and check the message integrity before use.

OpenID does not support any proof-of-rightful-possession methods, while in shibboleth the use of proof-of-rightful-possession methods is optional. Therefore, an IdP might not provide a user with the means to prove rightful possession of security token to an SP. Such an approach increases the risk of an attacker using a stolen token to earn access to SP resources. [42]

4.7. Revocation

In FIdM, revocation means disabling identity data, often represented as identity attributes in security tokens, therefore they can't be used for identification and authorisation purposes anymore. Current FIdM systems lack practical and efficient revocation techniques, this may lead to security violations. Revocation is an important issue in credential-based systems [44].

5. DISCUSSION AND RESULTS

This section presents the existing solutions for the challenges and limitations that been discussed in the previous section. In table 3 we provide a summary of the solutions suggested to each limitation discussed in section 4:

Table 3. Solution suggested for each limitation

| LIMITATION | SOLUTION SUGGESTED | REFERENCES |
|----------------------|---|-----------------|
| Trust | Dynamic trust establishment | [47] [48] [35] |
| | Independent trust establishment mechanisms | [34] |
| | Ensure identity trust through SAML credential | [49] |
| | Trusted Computing Technologies | [50] |
| | Identity assurance | [51] [52] |
| Privacy | Pseudonyms | [40] [53] |
| | Undetectability | [54] |
| | Unlinkability | [54] [44] |
| | Decentralized identity | [54] |
| | Privacy by design | [55] |
| IdP discovery | List of IdPs | [56] |

| | | |
|--|----------------------------------|----------------|
| Lack of attribute-aggregation support | Supporting attribute-aggregation | [43] |
| Complexity for the user | User-centric approaches | [56] [35] [52] |
| | Smart contract | [57] |
| Security | Encryption | [34] |
| | Digital Signature | [58] |
| | User identity distribution | [1] |
| | Zero-knowledge proofs | [1] |
| | Channel security | [59] |
| | Authorisation policies | [53] |
| Revocation | Limit token lifetime | [34] |

In systems like cloud computing systems or Web services trust relationship needs to be processed on-demand and at runtime which cannot be done in static trust establishment. So, the dynamic establishment of the trust relationship between entities (IdPs or SPs) in FIDM systems with the help of factors like data on the SLA and reputation of the IdP/ SP could solve such issue. In [47] and [48] a FIDM systems with dynamic trust establishment was proposed.

In this paper [35.] the researchers identified a set of factors that are fundamental for developing a holistic FIDM framework or model. These factors are Trust Management, Trust Establishment, User Privacy, Consistent User Access Rights across CoTs, Continuous Trust Monitoring, and Adaptation to Environmental or Unanticipated Changes. Based on these factors, they also presented a comparative analysis that helps identifies challenges and areas of improvements in FIDM. Choosing a Trust Management and Trust Establishment scheme depends on the user requirement, however, user privacy and alignment of user access rights across different CoTs need to be handled with both Trust Management and Trust Establishment schemes.

In this paper, [49] presented a trusted federated identity management mechanism. This mechanism helps to ensure identity trust through SAML credential, to guarantee the trustworthiness of the federated identity management procedure.

Trusted Computing Technologies can help to solve authentication, privacy and trust concerns in federated identity management systems. Khattak et al. in [50] have presented three threats in federated systems: Identity theft, Misuse of Information gathered by malicious IdPs and SPs, and trust relationship issues due to no or weak trust among users, IdPs and SPs. A Trusted platform (TP) is presented that confirms the rules of the Trusted Computing Platform Alliance (TCPA) specification to counter these threats. The presented framework can help to secure user privacy; however, it doesn't help for situations that unidentified at requirements engineering time [35.].

For preserving privacy and protect user identities, pseudonyms are an important technique, especially when multiple web services cooperate to provide an aggregated offering that requires user-attribute sharing. [40]

If SPs are trusted to link authorisation requests to identities, Pseudonymous authorisation is implemented by Project Liberty, OpenID, Passport, and Client-Side Federation. [53] However, If SPs aren't trusted with links between authorisation requests and identities, then anonymous authorisation is employed. Anonymous authorisation implemented by eliminating all unique identifiers from messages or credentials that the service provider doesn't explicitly require. For

example, Shibboleth supports anonymous authorisation, although users can choose to reveal a persistent identifier. Project Liberty lets a service provider request an anonymous, temporary, identifier for a user if the service provider elects to support anonymous authorisation. [54]

Undetectability and Unlinkability are privacy properties that help to preserve user privacy. Undetectability means users' ability to conceal actions from other parties. While Unlinkability concerns hiding correlations between combinations of actions and identities either permanently or temporarily, making it impossible to recognize two separate usages of the same credential [44]. Whether the linking between two identities was between action and identity, or between two actions, the level of trust that users grant to other parties determine the most appropriate design choice. In Project Liberty, the IdPs with established business relations create Circles of Trust (CoT). Within a CoT, a user can choose to federate two identities, in this case, the IdPs exchange information and the identities are linked [54]

In [54] the researchers identify crucial design choices essential to current identity management systems. They adopt a privacy-driven approach, which focuses on three privacy properties: Undetectability of authorisation requests which is concealing the user actions, Unlikability which is concealing correlations between combinations of actions and identities, and Confidentiality which means enabling users' control over dissemination of their attributes.

The most appropriate choice if IdPs can be trusted only with attributes that are specifically issued to them but not trusted with identity linking is a decentralized identity management system in which various, distinct IdPs each function separately using different protocols and not aware of each other. This architecture lets users select which IdPs to trust with which attributes, and spread critical attributes across distinct IdPs, thus ensuring unlikability of distinct identities. Most existing identity management systems, including Idemix, PRIME, Shibboleth, Higgins, CardSpace, OpenID, P-IMS, and U-Prove have adopted this approach.[54]

Though FIdM has mitigated the significant privacy flaws of the current situation by a number of techniques such as pseudonymous authentication and limited attribute release, however at the same time it also introduces new privacy issues, essentially by centralizing user data and making the track of user behavior easy and to link data of the same user together.

To mitigate these privacy risks the design process of FIdM systems needs to consider privacy requirements from the start. R. Hörbe and W. Hötendorfer in [55] focus on privacy by design requirements for FIdM systems. They presented a catalogue of privacy-related architectural requirements, joining up legal, business and system architecture viewpoints. Furthermore, the demonstration of concrete FIdM models showing how the requirements can be implemented in practice.

A common solution to the problem of IdP discovery is to provide a list of IdPs to the user from which the user must select the proper IdPs. however, this is could be a problem especially when the list of possible IdPs gets extensive and the user, who usually ignorant about these issues, must conduct a choice. This is called the "where are you from" problem and is a significant concern regarding usability. Rieger [56] mentions this problem and adds that because the users can be part of numerous federations this will complicate the situation more.

Liberty solves this problem by using the Liberty-Enabled Client (LEC) profile. This profile requires the participation of Liberty-Enabled User Agent (LEUA) to handle the messages sent and received during the federation and authentication processes. [42]

It would enhance the practicality of FIdM if SPs could acquire user attributes from multiple independent attribute authority to be used in association with a particular IdP. Supporting attribute-aggregation will help to solve the limitation which users is limited to choose one of their IdPs in any single working session with an SP. [42]

The Liberty Alliance was the first group to address the problem of attribute aggregation through its model of identity federation. In this model, the first IdP to authenticate the user inquires if the user prefers to be introduced to other IdPs in the federation. Afterwards, when the user authenticates to another IdP, it invites the user to federate its second identity with that from the first IdP. If the user consents, the two IdPs each create a random alias for the user and exchange secretly. Thus, neither IDP have knowledge of the user's true login identifier at the other IdP, but each can refer to the same user through the random aliases, and thereby aggregate the attributes. [43]

One solution to mitigate privacy concerns and the complexity of the user is to empower users to control their identities. Increasing users control over their information is a good solution to avoid the misuse of information and data leakage. A user-centric identity management system is developed essentially from the perspective of end-users, it aims to make the user task of managing digital identities easy by providing them with more control over their identities. [42] User-centric approaches extend the users' privacy as the user can decide which private information to send to the Consumer (e.g. SP) as in [56]. There are many advantages of user-centric identity management such as higher usability and privacy for the users, simplification of the protocol and the configuration compared to SAML-based federated identity management and helps to create trust among cloud service providers in a federated environment. [35]

In [44] the proposed system which enabling controlled access to and selective sharing of critical user attributes in FIdM solutions by integrating authenticated dictionary (ADT) into FIdM, this can help to develop a user-centric and user-friendly attribute sharing system.

In this work, [57] they presented an identity management system that provides FIdM such that a user can authenticate and transfer attributes to a relying party (RP) without the involvement of a credential service provider (CSP). They accomplish this by leveraging a smart contract running on a blockchain⁵. Their approach can increase privacy and reducing costs.

Regarding identity management, techniques to protect message confidentiality and integrity are essential to prevent compromisation of sensitive identity attributes or modification of identity attributes. This can be achieved through mechanisms such as encryption.[34]

Message security is essential in FIdM to prevent attackers and intermediaries from manipulating the messages that are in transit. Improper message security rises concern such as identity theft, false authentication, and unauthorised use of resources. Liberty Alliance specifications advised XML Digital Signature and Encryption [58] for encrypting a complete or a part of the SOAP message to preserve the integrity and confidentiality of its contents.

Bhargav-Spantzel et al. [1] recommended two kinds of techniques to protect the misuse of identity information: The distribution of user identity information among various entities and use techniques such as zero-knowledge proofs to prevent identity theft within an IdP or SP. They recommend that single central IdP is a problem in Shibboleth. Moreover, their work is also highlighting that Liberty does not consider untrusted SP or IdP within the specifications.

The availability of information in FIdM models can be ensured by having a common protocol or mechanism for communicating authentication and other information between parties and securing

communication channels and messages. Channel security can be accomplished using protocols like TLS1.0/SSL3.0 or other protocols with security characteristics that are equivalent to TLS or SSL. However, these protocols can only provide security at the transport level and not at the message level. For channel security, Liberty specifications highly recommend TLS/SSL with well-known cypher suites [58].

FIdM requires communicating parties to provide controlled access to information to authorised users. Authorisation goal is to deal with what information a user has access to or which operations a user can perform. A permission-based attribute sharing mechanism, which enables users to specify authorisation policies on the information that they want to share is recommended by Liberty specifications. [53]

A common way to mitigate revocation challenges is to limit the security token lifetime. By reducing the time-to-live to seconds or minutes the vulnerability window in cases of compromise of the token will be minimised. However, this may reduce the systems' usability as the user must reauthenticate to obtain a new valid security token. On the opposite side, when token expiration is set for a longer period user will benefit from the seamlessness, but the risk of identity theft and compromising information will increase. [34]

6. CONCLUSIONS

In our paper, we discussed the concept of identity federations well some federated identity management architectures such as liberty alliance, security assertion markup language SAML v2.0, WS-Federation, and Shibboleth with a comparison between these architectures. Furthermore, we presented the limitations of federated identity management based on how it affects the user. We determined the following limitations: trust, privacy, IdP discovery, lack of attribute-aggregation support, complexity for the user, security, and revocation. Finally, we discussed the solutions that proposed to mitigate the risk of these limitations.

In future work, an in-depth analysis of privacy, security, and trust challenges in a federated environment will be conducted. Also, we will propose a FIdM system taking into consideration the limitations and solutions we found in this paper.

REFERENCES

- [1] Bhargav-Spantzel, A. Squicciarini and E. Bertino, "Establishing and protecting digital identity in federation systems", *Journal of Computer Security*, vol. 14, no. 3, pp. 269-300, 2006. Available: 10.3233/jcs-2006-14303.
- [2] Clarke, R. 2004. Identity Management, Xamax Consultancy.
- [3] S. Clauß and M. Köhntopp, "Identity management and its support of multilateral security", *Computer Networks*, vol. 37, no. 2, pp. 205-219, 2001. Available: 10.1016/s1389-1286(01)00217-1.
- [4] G. Roussos, D. Peterson and U. Patel, "Mobile Identity Management: An Enacted View", *International Journal of Electronic Commerce*, vol. 8, no. 1, pp. 81-100, 2003. Available: 10.1080/10864415.2003.11044287.
- [5] J. Bolter, "Sherry Turkle, *Life on the Screen: Identity in the Age of the Internet* (London: Weidenfeld & Nicholson, 1996), 347pp. ISBN 0 297 81514 8", *Convergence: The International Journal of Research into New Media Technologies*, vol. 3, no. 1, pp. 131-133, 1997. Available: 10.1177/135485659700300112.
- [6] "Roger Clarke's 'Authentication Model'", Rogerclarke.com, 2021. [Online]. Available: <http://www.rogerclarke.com/EC/AuthModel.html>. [Accessed: 05- Feb- 2021].
- [7] International Organization for Standardization, Genève, Switzerland. ISO/IEC Second CD 24760 – Information technology -Security techniques A framework for identity management, January 2010

- [8] Satchell, G. Shanks, S. Howard and J. Murphy, "Identity crisis: user perspectives on multiplicity and control in federated identity management", *Behaviour & Information Technology*, vol. 30, no. 1, pp. 51-62, 2011. Available: 10.1080/01449290801987292.
- [9] E. Maler and D. Reed, "The Venn of Identity: Options and Issues in Federated Identity Management", *IEEE Security & Privacy Magazine*, vol. 6, no. 2, pp. 16-23, 2008. Available: 10.1109/msp.2008.50.
- [10] Chadwick, *Federated Identity Management*. Springer, Berlin, Heidelberg, 2009.
- [11] L. Project, "Home - Liberty Alliance", Projectliberty.org, 2021. [Online]. Available: <http://www.projectliberty.org/>. [Accessed: 05- Feb- 2021].
- [12] L. Project, "Home - Liberty Alliance", Projectliberty.org, 2021. [Online]. Available: <http://www.projectliberty.org/>. [Accessed: 05- Feb- 2021].
- [13] S. Shim, Geetanjali Bhalla and Vishnu Pendyala, "Federated identity management", *Computer*, vol. 38, no. 12, pp. 120-122, 2005. Available: 10.1109/mc.2005.408.
- [14] W. Sha'lan, *Privacy and practicality of identity management systems*. Saarbrücken: VDM Verlag Dr. Müller, 2011.
- [15] Friese et al., "Bridging IMS and Internet Identity," 2010 14th International Conference on Intelligence in Next Generation Networks, Berlin, 2010, pp. 1-6, doi: 10.1109/ICIN.2010.5640948.
- [16] Mansour and A. Carlisle, *Enhancing Consumer Privacy in the Liberty Alliance Identity Federation and Web Services Frameworks*. Privacy Enhancing Technologies (9783540687900), 59., 2006.
- [17] V. Frank, 2008. Liberty Alliance Releases Identity Assurance Framework. Liberty Alliance Project.
- [18] H. Eggleston and K. Ginanni, "Simplifying Licensed Resource Access Through Shibboleth", *The Serials Librarian*, vol. 56, no. 1-4, pp. 209-214, 2009. Available: 10.1080/03615260802686981.
- [19] "Shibboleth Consortium - Shaping the future of Shibboleth Software", Shibboleth Consortium, 2021. [Online]. Available: <https://www.shibboleth.net/>. [Accessed: 05- Feb- 2021].
- [20] P. John and G. Masha, "SHIBBOLETH FOR NEW GENERATION ACCESS MANAGEMENT (UK PERSPECTIVE)", *Proceedings of the IADIS International Conference*, 2005. [Accessed 5 February 2021].
- [21] J. Paschoud, "SHIBBOLETH AND SAML: AT LAST, A VIABLE GLOBAL STANDARD FOR RESOURCE ACCESS MANAGEMENT", *New Review of Information Networking*, vol. 10, no. 2, pp. 147-160, 2004. Available: 10.1080/13614570500053874.
- [22] M. Goodner et al. *Understanding WS-Federation*. Technical report, IBM and Microsoft, May 2007.
- [23] Nadalin et al. *Web services federation language (WS-Federation) version 1.1*. Technical report, December 2006
- [24] "Understanding WS-Federation", Docs.microsoft.com, 2021. [Online]. Available: [https://docs.microsoft.com/en-us/previous-versions/dotnet/articles/bb498017\(v=msdn.10\)?redirectedfrom=MSDN](https://docs.microsoft.com/en-us/previous-versions/dotnet/articles/bb498017(v=msdn.10)?redirectedfrom=MSDN). [Accessed: 05- Feb- 2021].
- [25] J. Kallela, "Federated identity management solutions." TKK T-110.5190 Seminar on Internetworking. 2008.
- [26] S. Cantor, J. Kemp, R. Philpot and E. Maler, *Assertions and protocols for the oasis security assertion markup language*. OASIS Open, 2004.
- [27] S. Ferdous and R. Poet, "Managing Dynamic Identity Federations using Security Assertion Markup Language", *Journal of theoretical and applied electronic commerce research*, vol. 10, no. 2, pp. 53-76, 2015. Available: 10.4067/s0718-18762015000200005.
- [28] "Cover Pages: Security Assertion Markup Language (SAML)", [Xml.coverpages.org](http://xml.coverpages.org/saml.html), 2021. [Online]. Available: <http://xml.coverpages.org/saml.html>. [Accessed: 05- Feb- 2021].
- [29] N. Duan and K. Smith, "IDentiaTM - An Identity Bridge Integrating OpenID and SAML for Enhanced Identity Trust and User Access Control", ACTA Press, 2012. Available: <https://www.actapress.com/PaperInfo.aspx?PaperID=454085&reason=500>. [Accessed 5 February 2021].
- [30] "OpenID Connect | OpenID", OpenID - The Internet Identity Layer, 2021. [Online]. Available: <http://openid.net/connect/>. [Accessed: 05- Feb- 2021].
- [31] Kang J, Elmehdwi Y, Lin D. SLIM: secure and lightweight identity management in VANETs with minimum infrastructure reliance, Vol. 238. Springer; 2018. p. 823–37.
- [32] U. Frago-rodriguez, M. Laurent-Maknavicius and J. Incera-Diequez, "Federated identity architectures", *Proc. 1st Mexican Conference on Informatics Security*, vol. 2006, 2006. [Accessed 5 February 2021].

- [33] "Liberty Alliance & WS-Federation: A Comparative Overview", LIBERTY ALLIANCE PROJECT, 2003. Available: <http://www.projectliberty.org/liberty/content/download/402/2765/file/wsfed-liberty-overview-10-13-03.pdf>. [Accessed 5 February 2021].
- [34] J. Jensen, "Federated Identity Management Challenges," 2012 Seventh International Conference on Availability, Reliability and Security, Prague, 2012, pp. 230-235, doi: 10.1109/ARES.2012.68
- [35] Malik, H. Anwar and M. A. Shibli, "Federated Identity Management (FIDM): Challenges and opportunities," 2015 Conference on Information Assurance and Cyber Security (CIACS), Rawalpindi, 2015, pp. 75-82, doi: 10.1109/CIACS.2015.7395570.
- [36] G. Bendiab, S. Shiaeles, S. Boucherkha and B. Ghita, "FCMDT: A novel fuzzy cognitive maps dynamic trust model for cloud federated identity management", *Computers & Security*, vol. 86, pp. 270-290, 2019. Available: 10.1016/j.cose.2019.06.011.
- [37] Malik, H. Anwar and M. A. Shibli, "Federated Identity Management (FIDM): Challenges and opportunities," 2015 Conference on Information Assurance and Cyber Security (CIACS), Rawalpindi, 2015, pp. 75-82, doi: 10.1109/CIACS.2015.7395570.
- [38] "General Data Protection Regulation (GDPR) – Official Legal Text", General Data Protection Regulation (GDPR), 2018. [Online]. Available: <https://gdpr-info.eu/>. [Accessed: 05- Feb- 2021].
- [39] "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data - OECD", Oecd.org, 2013. [Online]. Available: <http://www.oecd.org/sti/ieconomy/oecdguidelinesontheProtectionofPrivacyandTransborderFlowsOfPersonalData.htm>. [Accessed: 05- Feb- 2021].
- [40] Maler and D. Reed, "The Venn of Identity: Options and Issues in Federated Identity Management," in *IEEE Security & Privacy*, vol. 6, no. 2, pp. 16-23, March-April 2008, doi: 10.1109/MSP.2008.50.
- [41] U. Kylau, I. Thomas, M. Menzel and C. Meinel, "Trust Requirements in Identity Federation Topologies," 2009 International Conference on Advanced Information Networking and Applications, Bradford, 2009, pp. 137-145, doi: 10.1109/AINA.2009.80
- [42] W. Sha'lan, *Privacy and practicality of identity management systems*. Saarbrücken: VDM Verlag Dr. Müller, 2011.
- [43] W. Chadwick and G. Inman, "Attribute Aggregation in Federated Identity Management," in *Computer*, vol. 42, no. 5, pp. 33-40, May 2009, doi: 10.1109/MC.2009.143.]
- [44] D. Shin, R. Lopes and W. Claycomb, "Authenticated Dictionary-Based Attribute Sharing in Federated Identity Management," 2009 Sixth International Conference on Information Technology: New Generations, Las Vegas, NV, 2009, pp. 504-509, doi: 10.1109/ITNG.2009.193.
- [45] S. Rieger, "User-Centric Identity Management in Heterogeneous Federations," 2009 Fourth International Conference on Internet and Web Applications and Services, Venice, 2009, pp. 527-532, doi: 10.1109/ICIW.2009.85
- [46] P. Madsen, Y. Koga, and K. Takahashi, "Federated identity management for protecting users from id theft," in *Proceedings of the Workshop on Digital identity management*. ACM, 2005, pp. 77-83.

AUTHORS

Maha Aldosary is currently pursuing the M.Sc. degree in information security with Imam Muhammad ibn Saud Islamic University. She graduated with a bachelor's degree in computer science from University of Tabuk. Her research interests include blockchain technology, IoT, identity management and information security.

Norah Alqahtani is currently pursuing the M.Sc. degree in information security with Imam Muhammad ibn Saud Islamic University. She graduated with a bachelor's degree in computer science from Shagra University. Her research interests include Cloud Computing, blockchain technology, identity management and information security.

INFORMATIVE MULTIMODAL UNSUPERVISED IMAGE-TO-IMAGE TRANSLATION

Tien Tai Doan^{1,2}, Guillaume Ghyselinck¹, and Blaise Hanczar²

¹Dental Monitoring, Paris, France

²IBISC Laboratory, University of Evry Val d'Essonne, Evry-Courcouronnes, France

ABSTRACT

We propose a new method of multimodal image translation, called *InfoMUNIT*, which is an extension of the state-of-the-art method *MUNIT*. Our method allows controlling the style of the generated images and improves their quality and diversity. It learns to maximize the mutual information between a subset of style code and the distribution of the output images. Experiments show that our model cannot only translate one image from the source domain to multiple images in the target domain but also explore and manipulate features of the outputs without annotation. Furthermore, it achieves a superior diversity and a competitive image quality to state-of-the-art methods in multiple image translation tasks.

KEYWORDS

Multimodal Image-to-Image Translation, Mutual Information, GANs, Manipulating Features, Disentangled Representation

1. INTRODUCTION

Image-to-image translation can be described as the general problem of mapping an image from one domain to another domain. This seemingly simple approach is the foundation of many applications in the field of computer vision such as colorization [1], style transfer [2], super-resolution [3], denoising, inpainting [4]. Moreover, image-to-image translation has been also applied for data augmentation and achieved competitive results [5] [6] [7]. Based on the availability of data, the problem can be considered as supervised learning where the dataset contains paired samples; or unsupervised learning where the dataset consists of two independent sets of images. This work focuses on the unsupervised image-to-image problem which is more applicable due to its ease of obtaining data but also more challenged in terms of training.

Unsupervised image-to-image translation leads us to the idea that an image in a domain can be translated into multiple images in the second domain, which means the translation can be multimodal. For example, in image colorization, one image can be colored in multiple ways. Some methods [8] [9] have been proposed to use a noisy vector as an additional input of the decoder. The style of the generated images can then be manipulated by changing the values of the style-vector. However, the style-vectors in existing methods are entangled and the translated images are not interpretable as a result. Lacking control over features of the output can be problematic when important information are linked to these features. In the work of Cohen et al. [10], it is shown that CycleGAN was adding/removing tumors from images when transforming MRI images from Flair to T1, especially when there is an imbalance among classes in the training data. Therefore, learning to control the features of the translated images is essential. In this paper, we

propose some improvements on MUNIT [9] - a standard in the field of multimodal image translation - by applying the mutual information maximization technique. Our method, called InfoMUNIT, generates more diverse images and especially can manipulate their textural and structural features without requiring any annotation.

2. RELATED WORKS

2.1. Multimodal unsupervised image-to-image translation

The translation of images from one domain to another has been a challenging problem in computer vision. Thanks to the evolution of convolutional neural networks, especially generative adversarial networks (GANs) [11], many deep learning models have been recently proposed to address the problem of image translation and achieved impressive outcomes.

The research of Isola et al. in [12] is one of the earliest works on image-to-image translation based on GANs. In [13], the method is upgraded using multi-scale generators and discriminators to translate high-resolution images. These methods require paired data for training which is not usually available in practice.

Learning to translate images using unpaired data is more challenging than with paired data because we do not know exactly which data-point in the source domain corresponds to which one in the target domain. Thus, it is reasonable to add some constraints to the training when it is possible. One popular assumption in most image-to-image translation research is that the structure of an image must not be changed too much by the translation. This is similar to language translation, in which, a phrase must have the same meaning after being translated to another language. Shrivastava et al. [14] propose a training strategy in which, a deep network learns to transform the style of synthesized images to make them look more real. To preserve the annotation, they add a pixel-wise loss between the input and output of the style transfer network. Similar approaches are applied in later works such as specific-task loss [15], semantic features [16], or distance between pairs of input samples [17] and so on. These constraints are useful for some specific tasks and datasets but cannot be applied robustly.

Cycle consistency is another well-known loss function being used in many bi-direction image translation models such as DualGAN [18], CycleGAN [19], and DiscoGAN [20]. In these networks, an image being translated from domain A to domain B can be also translated backward to obtain the original image. As this cycle loss is not domain related, it can be applied to most of the bi-direction translation models. In [21], Almahairi et al. extend CycleGAN for learning a many-to-many mapping by combining images with noises. Despite its ease of use, cycle loss does not assure any consistency in terms of annotation which means labels of images can be flipped by the translation. Hoffman et al. [7] proposed to use both cycle consistency and semantic consistency during the training. However, this semantic constraint is not always accessible because it requires a pretrained classifier of a similar dataset.

Another way to preserve the structural information after the transformation is to define a shared latent space where domain-independent features are stored. In UNIT [6], Liu et al. propose to break the translation into two stages: encoding the source image to a latent code and then decoding this code to an image in the target domain. To gain some control over features of the translated image, Huang et al. develop MUNIT as an extension of UNIT, by splitting the latent code into two parts: content and style. With this network, multimodal translation can be done by combining a content code of an image with randomized style codes. The output images inherit content (or structure) from the input image but differ in style (Eg. textures or colors). In DRIT++ [22], a similar idea to

MUNIT is introduced but differs slightly in style transformation techniques. Both MUNIT and DRIT++ store image style in a completely entangled manner, they offer no control over the style of output images despite their diversity. In this work, we extend MUNIT by disentangling the style code without requiring any additional annotations or pretrained networks.

2.2. Unsupervised disentangled representation learning

Learning the features of images in an unsupervised fashion has received attention from the computer vision community for years.

Most methods in the early stage were based on restricted Boltzmann machines [23] and stacked auto-encoders [24]. Models in [25] and [26] were proposed for semi-supervised learning and achieved promising results on the MNIST dataset. In [27], a GANs-based method were shown to represent the dataset in a code space where basic linear structures are supported.

Another branch of research uses labeled data to learn disentangled representation. The representation is divided into two parts: one for the given labels and one for other features. Similar fashions of training can be found in different model structures such as bilinear models [28], multi-view perceptron [29], variational autoencoders (VAEs) [30] and adversarial autoencoder [31].

For minimizing the dependency on variation labels, weakly supervised methods were developed. Reed et al. [32] propose correspondence-based training strategies for a higher-order Boltzmann machine consisting of hidden units groups and each group represent a factor of variation. A similar technique is applied to VAE in [33] to manipulate brightness and pose in images of 3D objects. These two methods share one drawback that they require grouped data points which are difficult to collect in real-life applications.

There are not many works on completely unsupervised disentangled representation learning. In [34], hossRBM is introduced as a generalized version of spike-and-slab restricted Boltzmann machine, which entangles variation factors using its higher-order interactions on latent variables. However, the method is not effective in terms of computation cost.

In InfoGAN [35], Chen et al. develop an extension of GAN which maximizes the mutual information between certain variables in the latent code and samples from an unlabeled dataset. This technique enables the model to learn the disentangled representation of images without asking for labels. In this work, we upgrade MUNIT with the mutual information learning objective from InfoGAN to enable it to manipulate features of the translated images.

3. METHOD

Our objective is to translate images from a source domain A to a target domain B , and at the same time to learn the representation of the target domain. Following the idea called *partially shared latent space* in [9], we assume that each image can be encoded as a content code which contains general structural information and a style code which defines how the image will look like. In the state-of-the-art methods, this style latent code is entangled. In this work, we disentangle this style code by maximizing the mutual information between this code and the generated image.

3.1. Network architecture

Let x_A and x_B be two images from domain A and B respectively. Our objective is to learn a function $F_{A \rightarrow B}$ that projects images from domain A to domain B , $\hat{x}_{A \rightarrow B} =$

$F_{A \rightarrow B}(x_A)$. This function can be decomposed into parts: the encoder and the generator. The encoder E_A^c extracts the content code c_A from the image. The content code is a matrix representing the content of an image independently of its style. The generator G_B generates images in domain B from an content code and a style code s : $\hat{x}_{A \rightarrow B} = G_B(c_A, s) = G_B(E_A^c(x_A), [s', i])$. Since we want a one-to-many projection, a style code s_B is inputted in the generator to introduce variability in the generated images. The style code s is a vector created by concatenating two parts s' and i where s' stores entangled style of the generated images, i contains disentangled features of the generated images. s' and i are drawn from a normal distribution $N(0, I)$. The generator learns a function that links the points from a Gaussian distribution to the different ways to apply the style of domain B to a content code. In the same way, we define the function that projects images from domain B to domain A with generator G_A and encoder E_A^c . Notice that the content space and style space are common to both domains. This is the generators that project a couple of points from these common spaces to the image sub-spaces corresponding to their domain.

For the learning of these functions, we need to complete our architecture with autoencoders and discriminators. Autoencoders are used to reconstruct the original images from their decomposition into a content code and a style code. Let E_A^s (resp. E_B^s) denote the encoder that extracts from an image of domain A (resp. B) its style code $s_A = [s'_A, i_A]$ (resp. $s_B = [s'_B, i_B]$). The autoencoder of domain A is therefore defined by $\hat{x}_A = G_A(E_A^c(x_A), E_A^s(x_A))$. Autoencoders are also used to reconstruct the content $\hat{c}_A = E_A^c(G_B(c_A, s))$ and style codes $\hat{s} = E_B^s(G_B(c_A, s))$. The discriminator D_B is used to align the distribution of images produced by the generator G_B with the distribution of original images from domain A . It is also used to disentangle the style variables contained in the vector i . In the same way, we define the autoencoders $\hat{x}_B = G_B(E_B^c(x_B), E_B^s(x_B))$, $\hat{c}_B = E_B^c(G_A(c_B, s))$, $\hat{s} = E_A^s(G_A(c_B, s))$ and discriminator D_A . Figure 1 shows the complete architecture of InfoMUNIT.

3.2. Model learning

The training of our model consists to minimize a combination of reconstruction losses and adversarial losses while maximizing the variational mutual information.

Similar to most auto-encoder based architecture, the encoders E_A^c and E_A^s compress input images to content code and style code while the generator G_A takes them to reconstruct the original image from domain A . The image reconstruction loss $\mathcal{L}_{rec}^{x_A}$ makes sure the encoder and decoder inverse each other. L_1 loss is chosen for the image reconstruction as it usually obtains well the sharpness of the reconstructed image. For the same reason, we have similar reconstruction losses for content code $\mathcal{L}_{rec}^{c_A}$ and style code $\mathcal{L}_{rec}^{s_A}$.

$$\mathcal{L}_{rec}^{x_A} = E_{x_A \sim p(x_A)}[\| G_A(E_A^c(x_A), E_A^s(x_A)) - x_A \|_1] \quad (1)$$

$$\mathcal{L}_{rec}^{x_B} = E_{x_B \sim p(x_B)}[\| G_B(E_B^c(x_B), E_B^s(x_B)) - x_B \|_1] \quad (2)$$

$$\mathcal{L}_{rec}^{c_A} = E_{c_A \sim p(c_A), s \sim p(s)}[\| E_B^c(G_B(c_A), s) - c_A \|_1] \quad (3)$$

$$\mathcal{L}_{rec}^{c_B} = E_{c_B \sim p(c_B), s \sim p(s)}[\| E_A^c(G_A(c_B), s) - c_B \|_1] \quad (4)$$

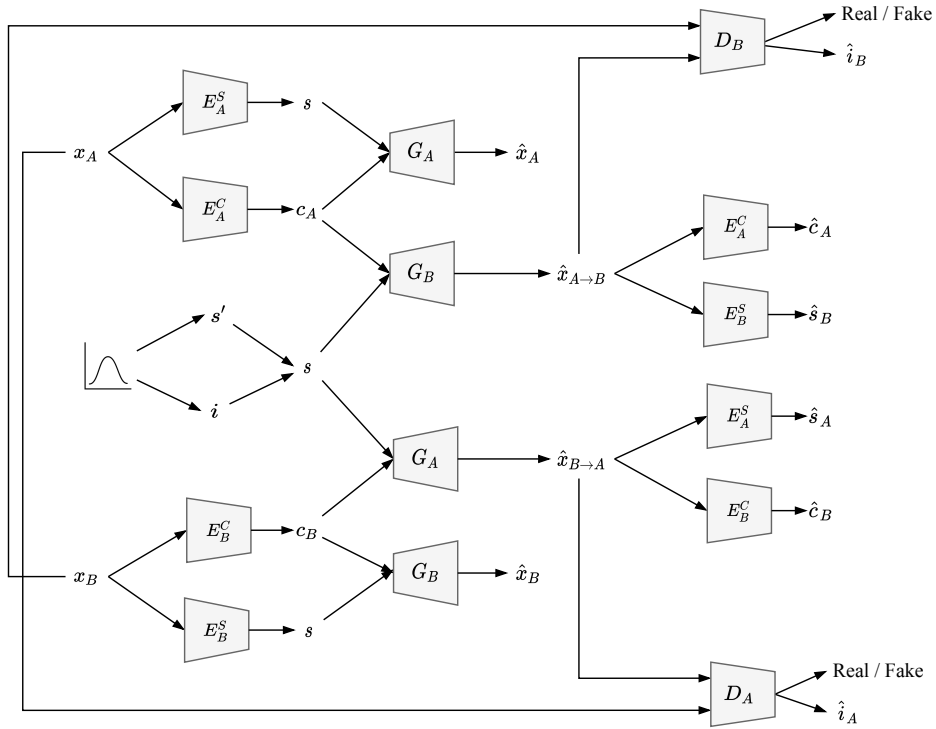


Figure 1: Overview of InfomUNIT. In this structure, each image is encoded by two encoders into a style code and a content code, and reconstructed by a decoder (also called generator). For translating an image from a domain to another domain, we firstly extract its content code, then combine it with a random style code, and send them both to the generator of the target domain. A part of the style code is used to store disentangled features of output images. We also train a pair of discriminators to distinguish between generated images and real images for each domain. The generators are also trained to maximize the mutual information between features being extracted by those discriminators and the disentangled part in the style code.

$$\begin{aligned} \mathcal{L}_{rec}^s = & E_{c_A \sim p(c_A), s \sim p(s)} [\| E_B^s(G_B(c_A), s) - s \|_1] \\ & + E_{c_B \sim p(c_B), s \sim p(s)} [\| E_A^s(G_A(c_B), s) - s \|_1] \end{aligned} \quad (5)$$

where $p(x_A)$ (resp. $p(x_B)$) is the distribution of images from domain A (resp. B), $p(c_A)$ (resp. $p(c_B)$) is the distribution of content code extracted from images from domain A (resp. B), and $p(s)$ is the distribution of style code that is the unit Gaussian distribution $N(0, I)$. Note that the distributions $p(c_A)$ and $p(c_B)$ are unknown and the learning set do not contains examples of c_A and c_B ., we need there fore to generate c_A and c_B samples from the encoders and training images $c_A = (E_A^c(x_A))$ and $c_B = (E_B^c(x_B))$.

The objective of the adversarial losses associated to the discriminators is to align the distributions of the real images with the distribution of the generated images. Like in the GAN, the discriminators try to predict if an image is a real one or an artificial image produced the generator. When the generators are frozen, the generators try to fool the discriminators in generating images close to the real ones. The adversarial losses are

defined by:

$$\begin{aligned} \mathcal{L}_{adv}^A = E_{x_B \sim p(x_B), s \sim p(s)} [\log(1 - D_A(G_A(E_B^C(x_B); s)))] \\ + E_{x_A \sim p(x_A)} [\log D_A(x_A)] \end{aligned} \quad (6)$$

$$\begin{aligned} \mathcal{L}_{adv}^B = E_{x_A \sim p(x_A), s \sim p(s)} [\log(1 - D_B(G_B(E_A^C(x_A); s)))] \\ + E_{x_B \sim p(x_B)} [\log D_B(x_B)] \end{aligned} \quad (7)$$

where the output of the discriminator $D_A(x)$ (resp. $D_B(x)$) is the probability that the image x is a real image from the domain A (resp. B).

Inspired by the idea of InfoGAN [35], we want a part of the style code to be disentangled features of the output in order to control and improve the diversity of the translated images. The style code is split into two parts $s = [s', i]$. To encourage the subvector i to represent disentangled features of the output, we maximize the mutual information between i and the generated images.

$$\mathcal{I}(i, G_B(c_A, [s, i])) \quad \text{and} \quad \mathcal{I}(i, G_A(c_B, [s, i])) \quad (8)$$

In practice, maximizing this mutual information is not achievable without access to the distribution $P(i|x)$ which is not available in our case. However, according to [36], we can define an additional distribution $Q(i|x)$ as an approximation of $P(i|x)$, and get a lower bound of the mutual information term. Thus we have:

$$\begin{aligned} \mathcal{I}(i, G_B(c_A, [s, i])) \geq L_{mi}(G_B, Q_B) = \\ E_{i \sim p(i), x_{A \rightarrow B} \sim P(G_B(c_A, [s', i]))} [\log Q_B(i|x_{A \rightarrow B})] \end{aligned} \quad (9)$$

Where $p(i)$ is a normal distribution and $P(G_B(c_A, [s', i]))$ is the distribution of the images generated by G_B with the style vector $[s', i]$. In practice, Q_B shares the same layers of the discriminator D_B as they both extract features from $G_B(c_A, [s', i])$. Q_B is implemented as a secondary output of the discriminator D_B that is notes \hat{i} . This means the closer the vector i and predicted vector \hat{i} are, the more mutual information between i and the generated image is achieved. In the same way, we define $L_{mi}(G_A, Q_A)$.

The learning of our model consists both to minimize the total loss w.r.t the encoders and generators and to maximize it w.r.t the discriminators :

$$\begin{aligned} \min_{E_A, E_B, G_A, G_B} \max_{D_A, D_B} \mathcal{L}(E_A, E_B, G_A, G_B, D_A, D_B) = \\ \mathcal{L}_{dis}^{x_A} + \mathcal{L}_{dis}^{x_B} + \lambda_x (\mathcal{L}_{rec}^{x_A} + \mathcal{L}_{rec}^{x_B}) + \lambda_c (\mathcal{L}_{rec}^{c_A} + \mathcal{L}_{rec}^{c_B}) \\ + \lambda_s (\mathcal{L}_{rec}^s) - \lambda_{mi} (L_{mi}(G_A, Q_A) + L_{mi}(G_B, Q_B)) \end{aligned} \quad (10)$$

where λ_x , λ_c , λ_s and λ_{mi} represent the importance of each loss. In our trainings, we set $\lambda_x = 10$, $\lambda_c = \lambda_s = \lambda_{mi} = 1$ as the image reconstruction is the most important loss in our structure.

4. EXPERIMENTS

4.1. Implementation Details

Our network consists of a content encoder, a style encoder, a generator, and a discriminator for each domain. We give the implementation details of each of these network.

4.1.1. Content Encoder

Input images are firstly led to the content encoder where they are down-sampled by strided convolutional layers and further processed by residual blocks. We apply Instance Normalization for all convolutional layers in the content encoder. The output of the content encoder is the content code in a form of a tensor.

4.1.2. Style Encoder

Similarly, the style encoder also down-samples input images using strided convolutional layers and a global pooling layer. A fully connected (FC) layer is applied to produce a style code as a vector consisting of 8 digits, in which, 2 final digits represent the information code (disentangled style) I_i of the image.

4.1.3. Generator

The generator takes content code and style code as inputs to reconstruct the initial input image. The content code goes through residual blocks and upsampling layers. These residual blocks are upgraded with Adaptive Instance Normalization (AdaIn) layers [37] which receive style parameters from a multilayer perception (MLP) which has the style code as its input.

4.1.4. Multi-purpose Discriminator

Our discriminator consists of two branches. The first branch is a traditional discriminator which can be found in most of the GAN-based models. The second branch consists of convolutional blocks to learn the Q distribution. These two branches share the first convolutional blocks.

4.1.5. Hyperparameters

In all our experiments in the paper, we apply Adam optimizer with β_1 and β_2 as 0.5 and 0.999 respectively. The learning rate is initially set to 0.0001, with a weight decay of 0.0001 applied every 100 thousand iterations. Our weight losses are $\lambda_x = 10$, $\lambda_c = \lambda_s = \lambda_{mi} = 1$.

4.1.6. Baselines

We compare our proposed method InfoMUNIT with following unpaired image-to-image translation techniques: CycleGAN [19], MUNIT [9] and DRIT++ [22]. The training procedures of those methods are done using official source code and configurations provided by their authors on GitHub.com.

4.2. Evaluation

We use three performance measures that estimate the quality and the diversity of the generated images, to compare InfoMUNIT with the baselines.

4.2.1. Conditional Inception Score

Based on Inception Score (IS) [38], Huang et al. [9] introduced Conditional Inception Score (CIS) specified for evaluation of multimodal image-to-image tasks. While IS measures the quality and diversity of all generated images at once, CIS focuses on the diversity of images that are translated from the same input image. Having multiple input images in the test set, we compute CIS for each group of images generated from the same input, and finally, take the mean CIS for the whole test set.

4.2.2. Frechet Inception Distance

Frechet Inception Distance (FID) [39] computes the distance between the set of generated images and the set images in the target domain. It is computed by calculated the distance between the Inception feature vectors for the two sets of images. Thus, FID can be used for evaluating networks that are trained on specific datasets without requiring a classifier pretrained on an alike dataset. The lower FID we have, the more realistic the generated images are. Normally, those feature vectors are taken from the third pooling layer of the Inception model which contains 2048 features. Due to the small size of our datasets, we compute the distance using features of the second pooling layer containing 192 features.

4.2.3. LPIPS Distance

The translation diversity is also measured by LPIPS distance which is shown in [40] to be highly correlated with human judgment. We compute LPIPS distance on generated samples of each input image, then take the average value. The larger distance among them, the more diverse they are.

4.3. Datasets

We use multiple datasets for evaluating InfoMUNIT and compare its performance with state-of-the-art techniques on the task of image-to-image translation. Each dataset contains two sets of images and our network is trained to transform images between the two domains. We crop and down-sample all images to the size of 64×64 , in RGB-color mode.

4.3.1. Edges \leftrightarrow Shoes and Edges \leftrightarrow Bags

These two datasets contain images of shoes and handbags along with their edges, introduced in the work of Isola et al. [12]. The edges \leftrightarrow shoes dataset contains 138667 pairs of samples while the edges \leftrightarrow bags dataset contains 49925 pairs. From each dataset, we keep 200 pairs of samples for testing and the rest for training. Note that we do not use the paired information of these two datasets.

4.3.2. Cats \leftrightarrow Dogs

The dataset is comprised of 1364 photos of dogs and 871 photos of cats, cropped to their heads [22]. We keep 100 images from each set for testing while the rest is used for training.

4.3.3. Portraits (Painted \leftrightarrow Real)

This dataset consists of 1814 painted portraits and real 6452 portraits captured by cameras [22]. We keep 100 images from each set for testing while using the rest for training.

5. RESULTS

5.1. Image Quality

The qualitative comparison of InfoMUNIT and other methods is shown in Figure 2. The objective of InfoMUNIT is to increase the diversity and ability to control features of generated images compared to MUNIT and the state-of-the-art, while not hurting their quality. As being shown in Figure 2, the quality of images generated by InfoMUNIT are at least as good as the images from other methods. The result is confirmed in Table 1 where we apply FID to quantitatively evaluate the realism of the generated images. Even



Figure 2: Random samples generated by our method and baselines, trained on two datasets: edges→bags (left) and cats→dogs (right). The input images (and ground-truths) are displayed in the first column. Other columns show random outputs of baseline methods and InfoMUNIT.

Table 1: Frechet Inception Distance (FID). Lower value means better performance.

| | InfoMUNIT | MUNIT | CycleGAN | DRIT++ |
|------------|------------------|--------------|-----------------|---------------|
| edge2bag | 2.81 | 2.56 | 4.23 | 1.69 |
| bag2edge | 7.68 | 8.52 | 58.53 | 5.64 |
| edge2shoe | 1.28 | 1.44 | 4.86 | 1.13 |
| shoe2edge | 4.69 | 8.83 | 88.03 | 4.24 |
| dog2cat | 9.24 | 13.48 | 2.56 | 21.59 |
| cat2dog | 6.31 | 6.31 | 2.02 | 18.14 |
| paint2real | 2.96 | 3.02 | 2.56 | 7.29 |
| real2paint | 8.60 | 8.51 | 3.97 | 18.85 |
| Average | 5.45 | 6.58 | 20.84 | 9.82 |

though InfoMUNIT does not outperform other methods in terms of image quality in any task, its performance is stable across all tasks. The performance of InfoMUNIT is close to the best method for each dataset. InfoMUNIT gives performance equivalent or better than MUNIT. This shows that the disentangled features have also an impact on the quality of the images. Notice that DRIT++ is the best for the first four tasks but totally fails in the last four tasks. This is illustrated by the strange dog images generated by DRIT++ in the Figure 2. On the opposite, CycleGAN gives the best performance for the last four tasks but is bad in the first four tasks and especially in the bag2edge and shoe2edge tasks. On average, InfoMUNIT achieves the best FID value among the four image-to-image translation methods. The good quality of images generated by InfoMUNIT is stable on multiple datasets.

5.2. Image Diversity

Table 2 and Table 3 respectively shows the CIS and LPIPS scores that evaluate the diversity of generated images. CycleGAN is not a multimodal method, it can generate only one output from one input so it does therefore not appear in this table. The LPIPS and CIS scores of InfoMUNIT are clearly superior to the scores of DRIT++ and MUNIT. The only exceptions are for the shoe2edge task where the LPIPS of DRIT++ is higher and for the real2paint task where the LPIPS of MUNIT is higher. In both cases the LPIPS of InfoMUNIT is very close to the best score and still higher than the LPIPS of the third

Table 2: LPIPS distance. Higher value means better performance.

| | InfoMUNIT | MUNIT | DRIT++ |
|------------|------------------|--------------|---------------|
| edge2bag | 3.00 | 2.07 | 2.13 |
| bag2edge | 2.01 | 1.07 | 1.60 |
| edge2shoe | 2.35 | 2.23 | 1.76 |
| shoe2edge | 1.44 | 1.00 | 1.51 |
| dog2cat | 2.24 | 1.97 | 1.11 |
| cat2dog | 2.65 | 2.24 | 1.09 |
| paint2real | 1.96 | 1.91 | 1.74 |
| real2paint | 2.14 | 2.26 | 1.14 |
| Average | 2.40 | 1.88 | 1.51 |

Table 3: Conditional Inception Score (CIS). Higher value means better performance.

| | InfoMUNIT | MUNIT | DRIT++ |
|------------|------------------|--------------|---------------|
| edge2bag | 0.42 | 0.29 | 0.30 |
| bag2edge | 0.35 | 0.04 | 0.22 |
| edge2shoe | 0.26 | 0.24 | 0.22 |
| shoe2edge | 0.24 | 0.00 | 0.12 |
| dog2cat | 0.32 | 0.32 | 0.04 |
| cat2dog | 0.30 | 0.28 | 0.03 |
| paint2real | 0.25 | 0.25 | 0.11 |
| real2paint | 0.33 | 0.30 | 0.06 |
| Average | 0.31 | 0.21 | 0.14 |

method. Over all datasets, the scores of InfoMUNIT are significantly better than the other methods. Figure 3 and Figure 4 illustrate the higher diversity of InfoMUNIT compared to MUNIT. This results show that InfoMUNIT generates significantly more diverse outputs than MUNIT and DRIT++.

5.3. Controlling Features

In this subsection, we show the advantage of InfoMUNIT over its predecessor MUNIT in manipulating features. From Figure 3, we can observe that varying values of style code in MUNIT can lead to slight changes like color of the object. With InfoMUNIT, we can significantly manipulate the features of the object. The first disentangled feature controls the size of the bag and the second one control the color from white to black. We also notice that InfoMUNIT is able to propose different textures of the bag.

The performance of InfoMUNIT on the edges→shoes task is illustrated in Figure 4 and Figure 5. While MUNIT can only change some small details of the shoes, we can significantly manipulate the color of the shoes with InfoMUNIT. Varying the first info style code makes the color changed from bright to dark, while varying the second one changes the color from cold to warm. In Figure 4, we can see that the first info style code is also responsible for the style of the shoes. From the left to the right, it turns a sneaker to a pump and makes it darker at the same time. This effect makes sense as pumps are more likely to have dark colors than sneakers.

Please note that the value of each disentangled feature in this test is plotted from -2 to 2 instead of -1 to 1 in the training phase, which means the generator is receiving style code values that it has never seen before. This explains why the images on the border looks a



Figure 3: Manipulating two last digits in the style code of MUNIT and InfoMUNIT on edges→bags task.



Figure 4: Manipulating two last digits in the style code of MUNIT and InfoMUNIT on edges→shoes task.

bit extreme.

5.4. The length of information latent code

We perform some experiments to investigate the impact of the length of information latent code i on the generated images in varying this value from 1 to 8. Table 4 shows some of these results on the *edge2shoe* datasets. We see that the FID, CIS and LPIPS weakly vary with the length of i . We conclude from these results that the quality and diversity of the generated images by InfoMUNIT are robust to the length of the information latent code.

6. CONCLUSION

We proposed an extension of MUNIT called InfoMUNIT which can manipulate features of the translated images. Our method is demonstrated in multiple image-to-image translation tasks. It achieves comparable translated image quality to state-of-the-art approaches and outperforms them in terms of outputs diversity. Moreover, our method improves the control of the user on the generated images, this kind of tool can make the image manipulation method more usable for real life applications.

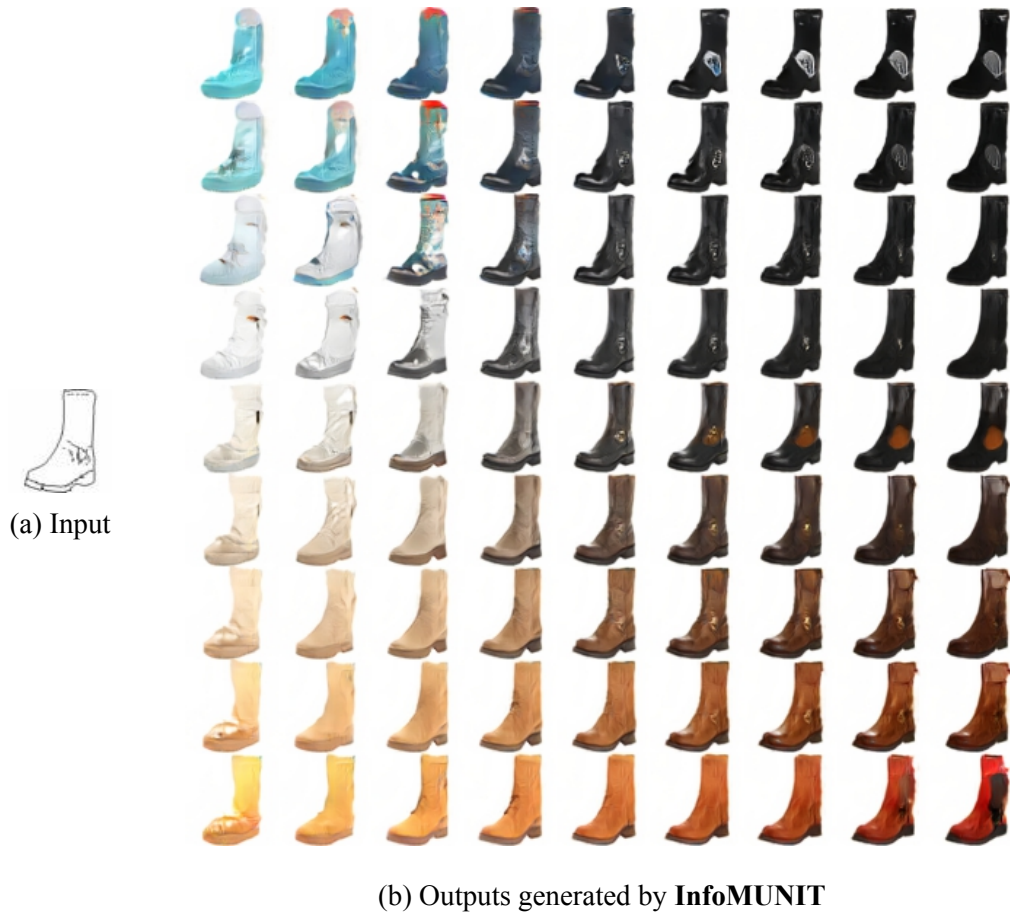


Figure 5: Combination of the two last digits in the style code of InfoMUNIT on edges→shoes task. From left to right (b), we vary the value of the first information latent code. From top to bottom, we vary the second one.

Table 4: Performance of InfoMUNIT with different lengths of information latent code.

| Length of i | 1 | 2 | 4 | 6 | 8 |
|---------------|------|-------------|-------------|-------------|-------------|
| FID | 2.99 | 2.81 | 3.19 | 3.11 | 3.37 |
| CIS | 2.59 | 3.00 | 3.64 | 3.61 | 3.65 |
| LPIPS | 0.42 | 0.42 | 0.47 | 0.47 | 0.46 |

References

- [1] S. Iizuka, E. Simo-Serra, and H. Ishikawa, “Let there be color! joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification,” *ACM Transactions on Graphics (ToG)*, vol. 35, no. 4, pp. 1–11, 2016.
- [2] L. A. Gatys, A. S. Ecker, and M. Bethge, “Image style transfer using convolutional neural networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2414–2423, 2016.
- [3] C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in *European conference on computer vision*, pp. 184–199, Springer, 2014.
- [4] J. Xie, L. Xu, and E. Chen, “Image denoising and inpainting with deep neural networks,” in *Advances in neural information processing systems*, pp. 341–349, 2012.
- [5] Z. Murez, S. Kolouri, D. Kriegman, R. Ramamoorthi, and K. Kim, “Image to image translation for domain adaptation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4500–4509, 2018.
- [6] M.-Y. Liu, T. Breuel, and J. Kautz, “Unsupervised image-to-image translation networks,” in *Advances in neural information processing systems*, pp. 700–708, 2017.
- [7] J. Hoffman, E. Tzeng, T. Park, J.-Y. Zhu, P. Isola, K. Saenko, A. A. Efros, and T. Darrell, “Cycada: Cycle-consistent adversarial domain adaptation,” *arXiv preprint arXiv:1711.03213*, 2017.
- [8] H.-Y. Lee, H.-Y. Tseng, J.-B. Huang, M. Singh, and M.-H. Yang, “Diverse image-to-image translation via disentangled representations,” in *Proceedings of the European conference on computer vision (ECCV)*, pp. 35–51, 2018.
- [9] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz, “Multimodal unsupervised image-to-image translation,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 172–189, 2018.
- [10] J. P. Cohen, M. Luck, and S. Honari, “Distribution matching losses can hallucinate features in medical image translation,” in *International conference on medical image computing and computer-assisted intervention*, pp. 529–536, Springer, 2018.
- [11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in neural information processing systems*, pp. 2672–2680, 2014.
- [12] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.
- [13] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, “High-resolution image synthesis and semantic manipulation with conditional gans,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8798–8807, 2018.
- [14] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, “Learning from simulated and unsupervised images through adversarial training,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2107–2116, 2017.

- [15] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, and D. Krishnan, “Unsupervised pixel-level domain adaptation with generative adversarial networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3722–3731, 2017.
- [16] Y. Taigman, A. Polyak, and L. Wolf, “Unsupervised cross-domain image generation,” *arXiv preprint arXiv:1611.02200*, 2016.
- [17] S. Benaim and L. Wolf, “One-sided unsupervised domain mapping,” in *Advances in neural information processing systems*, pp. 752–762, 2017.
- [18] Z. Yi, H. Zhang, P. Tan, and M. Gong, “Dualgan: Unsupervised dual learning for image-to-image translation,” in *Proceedings of the IEEE international conference on computer vision*, pp. 2849–2857, 2017.
- [19] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.
- [20] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, “Learning to discover cross-domain relations with generative adversarial networks,” in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 1857–1865, JMLR. org, 2017.
- [21] A. Almahairi, S. Rajeswar, A. Sordoni, P. Bachman, and A. Courville, “Augmented cyclegan: Learning many-to-many mappings from unpaired data,” *arXiv preprint arXiv:1802.10151*, 2018.
- [22] H.-Y. Lee, H.-Y. Tseng, Q. Mao, J.-B. Huang, Y.-D. Lu, M. Singh, and M.-H. Yang, “Drit++: Diverse image-to-image translation via disentangled representations,” *International Journal of Computer Vision*, pp. 1–16, 2020.
- [23] G. E. Hinton, S. Osindero, and Y.-W. Teh, “A fast learning algorithm for deep belief nets,” *Neural computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [24] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, “Extracting and composing robust features with denoising autoencoders,” in *Proceedings of the 25th international conference on Machine learning*, pp. 1096–1103, 2008.
- [25] A. Rasmus, M. Berglund, M. Honkala, H. Valpola, and T. Raiko, “Semi-supervised learning with ladder networks,” in *Advances in neural information processing systems*, pp. 3546–3554, 2015.
- [26] L. Maaløe, C. K. Sønderby, S. K. Sønderby, and O. Winther, “Improving semi-supervised learning with auxiliary deep generative models,” in *NIPS Workshop on Advances in Approximate Bayesian Inference*, 2015.
- [27] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *arXiv preprint arXiv:1511.06434*, 2015.
- [28] J. B. Tenenbaum and W. T. Freeman, “Separating style and content with bilinear models,” *Neural computation*, vol. 12, no. 6, pp. 1247–1283, 2000.
- [29] Z. Zhu, P. Luo, X. Wang, and X. Tang, “Multi-view perceptron: a deep model for learning face identity and view representations,” in *Advances in Neural Information Processing Systems*, pp. 217–225, 2014.
- [30] D. P. Kingma, S. Mohamed, D. J. Rezende, and M. Welling, “Semi-supervised learning with deep generative models,” in *Advances in neural information processing systems*, pp. 3581–3589, 2014.

- [31] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, “Adversarial autoencoders,” *arXiv preprint arXiv:1511.05644*, 2015.
- [32] D. Barber and F. V. Agakov, “Kernelized infomax clustering,” in *Advances in neural information processing systems*, pp. 17–24, 2006.
- [33] T. D. Kulkarni, W. F. Whitney, P. Kohli, and J. Tenenbaum, “Deep convolutional inverse graphics network,” in *Advances in neural information processing systems*, pp. 2539–2547, 2015.
- [34] G. Desjardins, A. Courville, and Y. Bengio, “Disentangling factors of variation via generative entangling,” *arXiv preprint arXiv:1210.5474*, 2012.
- [35] X. Chen, Y. Duan, R. Houthoofd, J. Schulman, I. Sutskever, and P. Abbeel, “Infogan: Interpretable representation learning by information maximizing generative adversarial nets,” in *Advances in neural information processing systems*, pp. 2172–2180, 2016.
- [36] D. Barber and F. V. Agakov, “The im algorithm: a variational approach to information maximization,” in *Advances in neural information processing systems*, p. None, 2003.
- [37] X. Huang and S. Belongie, “Arbitrary style transfer in real-time with adaptive instance normalization,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1501–1510, 2017.
- [38] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, “Improved techniques for training gans,” in *Advances in neural information processing systems*, pp. 2234–2242, 2016.
- [39] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, “Gans trained by a two time-scale update rule converge to a local nash equilibrium,” in *Advances in neural information processing systems*, pp. 6626–6637, 2017.
- [40] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, “The unreasonable effectiveness of deep features as a perceptual metric,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586–595, 2018.

IMAGE CLASSIFIERS FOR NETWORK INTRUSIONS

David A. Noever and Samantha E. Miller Noever

PeopleTec, Inc., Huntsville, Alabama, USA

ABSTRACT

This research recasts the network attack dataset from UNSW-NB15 as an intrusion detection problem in image space. Using one-hot-encodings, the resulting grayscale thumbnails provide a quarter-million examples for deep learning algorithms. Applying the MobileNetV2's convolutional neural network architecture, the work demonstrates a 97% accuracy in distinguishing normal and attack traffic. Further class refinements to 9 individual attack families (exploits, worms, shellcodes) show an overall 56% accuracy. Using feature importance rank, a random forest solution on subsets show the most important source-destination factors and the least important ones as mainly obscure protocols. The dataset is available on Kaggle.

KEYWORDS

Neural Networks, Computer Vision, Image Classification, Intrusion Detection, MNIST Benchmark.

1. INTRODUCTION

This work updates the UNSW-NB15 attack dataset [1-4] and extends the popular intrusions detection system (IDS) originally inspired by the KDD-99/DARPA challenge [5-7]. The details of the UNSW-NB15 dataset are published in a series of previous papers [1-4] which described the raw network packet captures, generated features on labeled attacks, and scored statistical methods for identifying each attack family. As illustrated in Figure 1, the current approach aims to map scaled numerical features to images, a method likened to traditional spectrogram methods. These fingerprinting techniques have proven useful when image-based neural networks have solved similar but challenging time-dependent [8] or audio [9] problems. We test the capabilities for mapping tabular features to build fast image classifiers. One advantage of this hierarchical method arises from the unique power of transfer learning to high accuracy, even when the underlying patterns prove difficult for humans to understand or classify. The dataset is available on Kaggle [10].

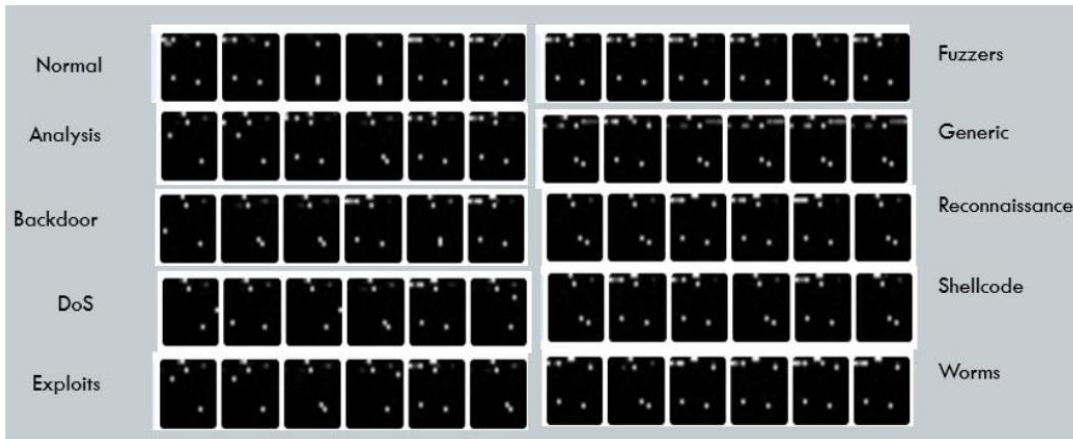


Figure 1. Nine attack types and one normal traffic dataset. We map the tabular features to grayscale thumbnails.

2. METHODS

| | | | | | | | | |
|----|----------------------|-------------------|--------------------|-------------------|-------------------|-------------------|------------------|---------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | dur | spkts | dpkts | sbytes | dbytes | rate | sttl | dttl |
| 2 | swin | stcpb | dtcpb | dwin | tcprrt | synack | ackdat | smean |
| 3 | ct_dst_src_ltm | is_ftp_login | ct_ftp_cmd | ct_flw_http_mthd | ct_src_ltm | ct_srv_dst | is_sm_ips_ports | service_ |
| 4 | service_smtp | service_snmp | service_ssh | service_ssl | protocol_3pc | protocol_a | n | protocol_aes-sp3-d |
| 5 | protocol_cbt | protocol_cftp | protocol_chaos | protocol_compaq | protocol_cphb | protocol_cpnx | protocol_crtp | protocol_crudp |
| 6 | protocol_etherip | protocol_fc | protocol_fire | protocol_ggp | protocol_gmtp | protocol_gre | protocol_hmp | protocol_i-nlsp |
| 7 | protocol_il | protocol_ip | protocol_ipcomp | protocol_ipcv | protocol_ipip | protocol_ipit | protocol_ipnip | protocol_ippc |
| 8 | protocol_iso-ip | protocol_iso-tp4 | protocol_kryptolan | protocol_l2tp | protocol_larp | protocol_leaf-1 | protocol_leaf-2 | protocol_merit-inp |
| 9 | protocol_nsfnet-igmp | protocol_nvp | protocol_ospf | protocol_pgm | protocol_pim | protocol_pipe | protocol_pnni | protocol_pri-enc |
| 10 | protocol_sat-expak | protocol_sat-mon | protocol_sccopmce | protocol_scps | protocol_sctp | protocol_sdrp | protocol_secure- | protocol_sep |
| 11 | protocol_stp | protocol_sun-nd | protocol_swipe | protocol_tcf | protocol_tcp | protocol_tlsp | protocol_tp++ | protocol_trunk-1 |
| 12 | protocol_vrrp | protocol_wb-expak | protocol_wb-mon | protocol_wsn | protocol_xnet | protocol_xns-idp | protocol_xtp | protocol_zero |
| 13 | pad | pad | pad | pad | pad | pad | pad | pad |
| 14 | pad | pad | pad | pad | pad | pad | pad | pad |
| 15 | pad | pad | pad | pad | pad | pad | pad | pad |
| 16 | pad | pad | pad | pad | pad | pad | pad | pad |
| | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 1 | sload | dload | sloss | dloss | sinpkt | dinpkt | sjit | djit |
| 2 | dmean | trans_depth | response_body_le | ct_srv_src | ct_state_ttl | ct_dst_ltm | ct_src_dport_ltm | ct_dst_sport_ltm |
| 3 | service_dhcp | service_dns | service_ftp | service_ftp-data | service_http | service_irc | service_pop3 | service_radius |
| 4 | protocol_any | protocol_argus | protocol_aris | protocol_arp | protocol_ax.25 | protocol_bbn-rcc | protocol_bna | protocol_br-sat-mon |
| 5 | protocol_dcn | protocol_ddp | protocol_ddx | protocol_dgp | protocol_egp | protocol_eigrp | protocol_emcon | protocol_encap |
| 6 | protocol_iatp | protocol_ib | protocol_idpr | protocol_idpr-cm | protocol_idrp | protocol_ifmp | protocol_igmp | protocol_igp |
| 7 | protocol_ipv6 | protocol_ipv6-fra | protocol_ipv6-no | protocol_ipv6-opt | protocol_ipv6-ro | protocol_ipx-n-ip | protocol_irtp | protocol_isis |
| 8 | protocol_mfe-nsp | protocol_mhrp | protocol_micp | protocol_mobile | protocol_mtp | protocol_mux | protocol_narp | protocol_netblt |
| 9 | protocol_prm | protocol_ptp | protocol_pup | protocol_pvp | protocol_qnx | protocol_rdp | protocol_rsvp | protocol_rvd |
| 10 | protocol_skip | protocol_sm | protocol_smp | protocol_snp | protocol_sprite-r | protocol_sps | protocol_srp | protocol_st2 |
| 11 | protocol_trunk-2 | protocol_ttp | protocol_udp | protocol_unas | protocol_uti | protocol_vines | protocol_visa | protocol_vmtp |
| 12 | state_ACC | state_CLO | state_CON | state_FIN | state_INT | state_REQ | state_RST | pad |
| 13 | pad | pad | pad | pad | pad | pad | pad | pad |
| 14 | pad | pad | pad | pad | pad | pad | pad | pad |
| 15 | pad | pad | pad | pad | pad | pad | pad | pad |
| 16 | pad | pad | pad | pad | pad | pad | pad | pad |

Figure 2. Layout template for one hot encoded images.

The use of convolutional neural networks for network attack classification depends on first converting all tabular feature sets into thumbnail images. We, therefore, recast the UNSW-NB15 tabular set of features as scaled image thumbnails [11] to solve for 9 attack types. This version of the dataset renders the corresponding UNSW attack set as 256-pixel grayscale images (16 x16). We employ one-hot-encoding [12] for the categorical inputs and rescale all numerical inputs as grayscale pixel values (0-255) between the training set's minimum and maximum values. The baseline UNSW-NB15 dataset [1-4] yields 194 values and the images are right padded with all black (255) values for any unused pixels (62) identically for all attack labels. This padding assists deep learning approaches [13] which have a stride length in powers of 2. The column labels are also included in the train and test sets as tabular formats (comma-separated value files) to compare image-based classification methods to more statistical approaches like decision trees, random forest, and support vector machines. The expectation is that all the legacy algorithms of both deep learning and statistical machine learning may assist in the new task after mapped to images of feature sets. This approach shares many characteristics with the traditional MNIST dataset [14-20] and thus can build quickly on those findings for algorithmic comparisons. Several image-based problems to solve include simply binary classifiers for attack vs. normal traffic. Like MNIST digits [14], there are 10 categories shown (0=normal; 1-9 various attacks). As shown in Figure 2, the original 42 network features expand to 194 when one-hot-encoded [12]. This process converts all categorical data (services, protocols, and states) into individual columns with their presence marked by 1 and absence by 0. For instance, protocol_http becomes pixel value 255 at the appropriate grayscale image location (row=3, column = 13) if the attack used hypertext transfer protocol. Conversely, the same pixel maps to 0 if the protocol was not used. Each row of the UNSW thus renders 256 features (of which 194 follow directly from the tabular set).

| Training | | Family | Count |
|----------|---------------|--------------------|--------------|
| Attack | 45332 | Analysis | 677 |
| Normal | 37000 | Backdoor | 583 |
| | 82332 | DoS | 4089 |
| Testing | | Exploits | 11132 |
| Attack | 119341 | Fuzzers | 6062 |
| Normal | 56000 | Generic | 18871 |
| | 175341 | Normal | 37000 |
| | | Reconnaissance | 3496 |
| | | Shellcode | 378 |
| | | Worms | 44 |
| | | Grand Total | 82332 |

Figure 3. Training and testing count per attack family count

We leave unchanged the train/test split of the original UNSW-NB15 dataset at a 1:2 ratio [1-4]. The detailed counts for each class are shown in Figure 3. It is worth noting that the UNSW-NB15 dataset updates and statistically rebalances some of the KDD99 counts based on their analysis of duplicates and potential data leakage between training and test sites. The ratio of 1:2 for training and test presents a challenging amount of previously unseen data when an algorithm gets scored or deployed. In total, we created almost a quarter-million images as 256-pixel thumbnails using ImageMagick [11]. The largest training class (normal traffic: 37,000) outnumbers the smallest attack class (worms: 44) by nearly 1000:1 as a ratio of cases.

To explore whether transfer learning from a convolutional neural network can identify network attacks, we tested the small (2 Mb) MobileNetV2 model [13] as pre-trained, then introduced both the binary and multi-class problems. The binary classifier determines whether a given image pattern represents normal or attack traffic. The multi-class problem identifies one of the 10

possible families (9 attacks in Figure 3 vs. normal). The multi-class example shares an analogous data setup to the traditional MNIST handwriting dataset [14] and thus may benefit from the various state-of-the-art approaches developed to handle those 10 classes. We solve both the binary and multi-class cases with a standard set of hyper-parameters (epochs:50, batch size:16, learning rate: 0.001). Slower learning rates disrupt the pre-trained layers of the neural network and preserve some of its beneficial weights for feature extraction in the images. We also explore the effects of smaller dataset size (<10,000 training examples vs. the full 250k) [22].

To explore the feature importance for detecting attacks, we applied a random forest algorithm (Figure 4) to the binary classifier [23]. The rank order for the top 14 contributors is shown using the Gini Index (or impurity) which effectively gauges the factors contributing to a decision split between normal and attack [23]. The highest contributors include 1) the Source to destination time to live value (sttl); 2) Number for each state (dependent protocol, e.g. ACC, CLO, CON) according to a specific range of values for source/destination time to live (ct_state_ttl); and 3) Number of connections of the same source and the destination address in 100 connections according to the last time (ct_dst_src_ltm). Not shown in Figure 4 are the least important which somewhat surprisingly include most of the one-hot-encoded protocol features that are more exotic than ordinary TCP (e.g. zero, XTP, XMN.IDP, WSN, etc.). In addition to providing a future path to reduce the intrusion detector’s dimensionality, this feature importance rank defines what cannot be safely ignored in attack datasets like UNSW-NB15 [1-4].

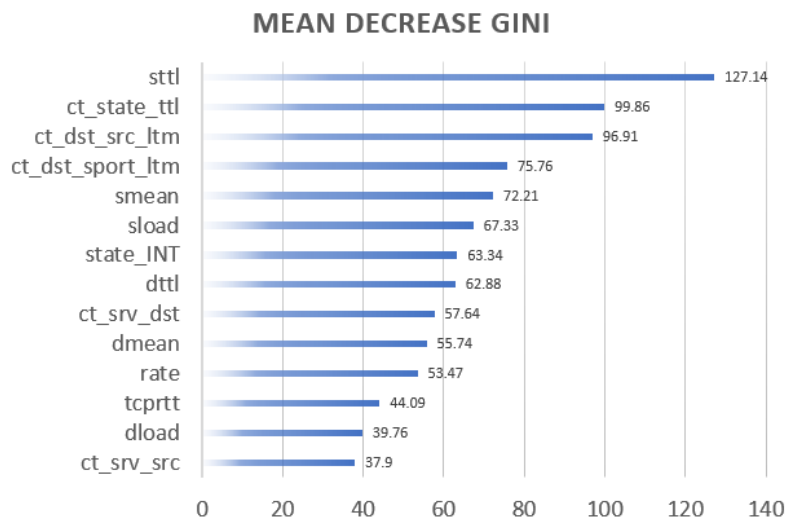


Figure 4. Feature information (GINI) Contribution to Attack Detection

3. RESULTS

As shown in Figures 4-6, the results for transfer learning with a deep convolutional neural network (MobileNetV2 [13]) demonstrate that the image-based binary classifier achieves greater than 97% accuracy in identifying whether an attack occurs (Figure 5).

| Family | Accuracy | Test |
|--------|----------|------|
| Normal | 98% | 500 |
| Attack | 97% | 542 |

Figure 5. Binary classifier results

In Figure 6, the specific identification of an attack family averages 54% accuracy between the 10 classes, with large deviations between the best (normal and generic traffic: 96+%) and the worst (worms and analysis <19%). Figure 6 shows the error matrix of which attack images confuse the neural network (worms misclassified as exploits). It is worth noting that the majority class (normal) loses no performance as more attack classes get added from the binary to the multi-classification example.

| Family | Accuracy | Test | Accuracy | Test | Family | Accuracy | Test | Accuracy | Test |
|----------|----------|------|----------|------|-----------|----------|------|----------|------|
| Analysis | 19% | 102 | 21% | 39 | Generic | 96% | 226 | | 0 |
| Backdoor | 69% | 88 | 44% | 41 | Normal | 98% | 473 | 67% | 36 |
| DoS | 34% | 161 | 54% | 41 | Reconn | 63% | 176 | 5% | 44 |
| Exploits | 66% | 264 | 44% | 50 | Shellcode | 21% | 57 | 79% | 42 |
| Fuzzers | 70% | 204 | 82% | 39 | Worms | 0% | 7 | 14% | 7 |

Figure 6. Multi class results

One contribution to this variance is the relative sparsity of UNSW-NB15 examples for the lower performing classes. To test this hypothesis, we performed the same experiment on a smaller subset (<4000) of images with a hold-out test and validation set that represents 20% of the training set (as opposed to 200% in the original UNSW-NB15 split). By better balancing the dataset, undetectable worms, and other lesser represented classes could be detected in Figure 6 (second column). Mapping attacks to images shows the dependence of accuracy on both class size and imbalance [23].

One interesting outcome of using these image-based detection maps is their portability to small hardware appliances. The small MobileNetV2 architecture [13] is tailored to run on edge devices [24], such as mobile phones. Simple network detectors thus render a complex matrix of packet features into a rapid classifier capable of running in near real-time imagery (e.g. 30 frames -or attacks- per second). The reduction of the model to use tflite (Tensorflow) as a set of stored weights represents a standard model [25] for deploying deep learning to edge devices.

| | Normal | Analysis | Backdoor | DoS | Exploits | Fuzzers | Generic | Reconn | Shellcode | Worms |
|-----------|--------|----------|----------|-----|----------|---------|---------|--------|-----------|-------|
| Normal | 98% | 0% | 0% | 0% | 1% | 1% | 0% | 0% | 0% | 0% |
| Analysis | 0% | 19% | 51% | 10% | 19% | 2% | 0% | 0% | 0% | 0% |
| Backdoor | 0% | 14% | 69% | 5% | 9% | 1% | 0% | 2% | 0% | 0% |
| DoS | 0% | 2% | 0% | 34% | 59% | 2% | 0% | 2% | 1% | 0% |
| Exploits | 1% | 0% | 1% | 19% | 66% | 3% | 0% | 9% | 0% | 0% |
| Fuzzers | 0% | 0% | 3% | 8% | 9% | 70% | 0% | 8% | 0% | 0% |
| Generic | 0% | 0% | 0% | 0% | 2% | 0% | 96% | 0% | 0% | 0% |
| Reconn | 0% | 1% | 1% | 14% | 17% | 3% | 0% | 63% | 2% | 0% |
| Shellcode | 0% | 0% | 0% | 0% | 0% | 5% | 0% | 74% | 21% | 0% |
| Worms | 0% | 0% | 0% | 0% | 86% | 14% | 0% | 0% | 0% | 0% |

Figure 7. Confusion matrix by class

4. DISCUSSION AND CONCLUSIONS

The results demonstrate a viable path of converting tabular feature data to image thumbnails, then applying convolutional neural networks to classify attack families. One potential shortcoming in our approach is any dependencies on the parametric ordering in the table format. For instance, CNNs tend to highlight close neighbors as being related in the image [26], yet there is no obvious relationship in the generated images between protocols, services, or states that justify making them into a particular attack fingerprint. One could address this flaw quantitatively by shuffling the order and determining the change of accuracy (if any). Future work should compare alternative statistical methods borrowed from the extensive machine learning literature devoted to the MNIST (and its derivative [14-21]) dataset of handwriting recognition. One can anticipate that like MNIST solutions, there exist high accuracy decision trees (like extreme gradient boosted trees – XGBoost [27]) that generate both accuracy and inference speeds comparable to the deep learning approach here. Further work could also use the image dataset [28] to design new attacks (and defenses) based on the techniques of generative adversarial networks (GANs [28]). The dataset is available on Kaggle [10].

ACKNOWLEDGMENTS

The author would like to thank the PeopleTec Technical Fellows program for encouragement and project assistance.

REFERENCES

- [1] Moustafa, Nour, and Jill Slay. "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)." Military Communications and Information Systems Conference (MilCIS), 2015. IEEE, 2015. See online <https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/>
- [2] Moustafa, Nour, and Jill Slay. "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 dataset and the comparison with the KDD99 dataset." *Information Security Journal: A Global Perspective* (2016): 1-14.
- [3] Moustafa, Nour, et al. . "Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks." *IEEE Transactions on Big Data* (2017).
- [4] Moustafa, Nour, et al. "Big data analytics for intrusion detection system: statistical decision-making using finite Dirichlet mixture models." *Data Analytics and Decision Support for Cybersecurity*. Springer, Cham, 2017. 127-156.
- [5] Özgür, Atilla, and HamitErdem. "A review of KDD99 dataset usage in intrusion detection and machine learning between 2010 and 2015." *PeerJ Preprints* 4 (2016): e1954v1.
- [6] Olusola, A. A., Oladele, A. S., &Abosedede, D. O. (2010, October). Analysis of KDD'99 intrusion detection dataset for selection of relevance features. In *Proceedings of the world congress on engineering and computer science* (Vol. 1, pp. 20-22). WCECS.
- [7] Meena, Gaurav, and Ravi Raj Choudhary. "A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA." In *2017 International Conference on Computer, Communications and Electronics (Comptelix)*, pp. 553-558. IEEE, 2017.
- [8] Hatami, Nima, Yann Gavet, and Johan Debayle. "Classification of time-series images using deep convolutional neural networks." In *Tenth international conference on machine vision (ICMV 2017)*, vol. 10696, p. 106960Y. International Society for Optics and Photonics, 2018.
- [9] Hershey, Shawn, Sourish Chaudhuri, Daniel PW Ellis, Jort F. Gemmeke, Aren Jansen, R. Channing Moore, Manoj Plakal et al. "CNN architectures for large-scale audio classification." In *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 131-135. IEEE, 2017.
- [10] Noever, David "Intrusion Detection as an Image Classifier", Kaggle.com, (2021), <https://www.kaggle.com/datamunge/intrusion-detection-as-an-image-classifier>
- [11] Salehi, Sohail. *ImageMagick Tricks*. Packt publishing ltd, 2006.

- [12] Zhang, Weinan, Tianming Du, and Jun Wang. "Deep learning over multi-field categorical data." In European conference on information retrieval, pp. 45-57. Springer, Cham, 2016.
- [13] Sandler, Mark, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. "Mobilenetv2: Inverted residuals and linear bottlenecks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510-4520. 2018.
- [14] LeCun, Yann, Corinna Cortes, and C. J. Burges. "MNIST handwritten digit database." (2010): 18. <http://yann.lecun.com/exdb/mnist/> and Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." *Proceedings of the IEEE*, 86(11):2278-2324, November 1998
- [15] Cohen, Gregory, Saeed Afshar, Jonathan Tapson, and Andre Van Schaik. "EMNIST: Extending MNIST to handwritten letters." In *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 2921-2926. IEEE, 2017.
- [16] Chen, Li, Song Wang, Wei Fan, Jun Sun, and Satoshi Naoi. "Beyond human recognition: A CNN-based framework for handwritten character recognition." In *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, pp. 695-699. IEEE, 2015.
- [17] Image Classification on MNIST, (accessed 01/2021), <https://paperswithcode.com/sota/image-classification-on-mnist>
- [18] Grim, Jiri, and Petr Somol. "A Statistical Review of the MNIST Benchmark Data Problem." <http://library.utia.cas.cz/separaty/2018/RO/grim-0497831.pdf>
- [19] Preda, Gabriel, Chinese MNIST: Chinese Numbers Handwritten Characters Images, (accessed 01/2021) <https://www.kaggle.com/gpreda/chinese-mnist>
- [20] CoMNIST: Cyrillic-oriented MNIST, A Dataset of Latin and Cyrillic Letters, (accessed 01/2021) <https://www.kaggle.com/gregvial/comnist>
- [21] Prabhu, Vinay Uday. "Kannada-MNIST: A new handwritten digits dataset for the Kannada language." *arXiv preprint arXiv:1908.01242* (2019). <https://www.kaggle.com/higgstachyon/kannada-mnist>
- [22] Warden, P. "How many images do you need to train a neural network?" (2017). <https://petewarden.com/2017/12/14/how-many-images-do-you-need-to-train-a-neural-network/>
- [23] Han, Hong, Xiaoling Guo, and Hua Yu. "Variable selection using mean decrease accuracy and mean decrease gini based on random forest." In *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSSSS)*, pp. 219-224. IEEE, 2016.
- [24] Lee, Juhyun, Nikolay Chirkov, Ekaterina Ignasheva, YuryPisarchyk, Mogan Shieh, Fabio Riccardi, Raman Sarokin, Andrei Kulik, and Matthias Grundmann. "On-Device Augmented Reality with Mobile GPUs."
- [25] Shah, Vishal, and Neha Sajnani. "Multi-Class Image Classification using CNN and Tflite." *International Journal of Research in Engineering, Science and Management* 3, no. 11 (2020): 65-68.
- [26] Liu, Li, Jie Chen, Paul Fieguth, Guoying Zhao, Rama Chellappa, and Matti Pietikäinen. "From BoW to CNN: Two decades of texture representation for texture classification." *International Journal of Computer Vision* 127, no. 1 (2019): 74-109.
- [27] Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794. 2016.
- [28] Shorten, Connor, and Taghi M. Khoshgoftaar. "A survey on image data augmentation for deep learning." *Journal of Big Data* 6, no. 1 (2019): 1-48.
- [29] Samangouei, Pouya, Maya Kabkab, and Rama Chellappa. "Defense-gan: Protecting classifiers against adversarial attacks using generative models." *arXiv preprint arXiv:1805.06605* (2018).

A STUDY OF REGRESSION TESTING FOR TRADE ME WEBSITE

Kenil Manishkumar Patel and Shahid Ali

Department of Information Technology, AGI Institute, Auckland, New Zealand

ABSTRACT

Regression testing plays a critical role to verify the functionality testing of a product. Trade Me is New Zealand based website. It is one of the major websites in New Zealand dealing in buying and selling online. The aim of this research is to find out the functionalities of Trade Me website after injecting new features. Automation regression suite is used to execute test scripts which helped the company to save time and cost compared to manual testing. Automation regression test suite also helped to prioritize test cases are designed in such a way that it can maximize the fault detection. For research analysis scrum methodology is used to meet the ultimate desires of software development companies and to increase the client satisfaction.

KEYWORDS

Regression testing, automation testing, scrum methodology, testNG, selenium

1. INTRODUCTION

Trade Me is New Zealand based website. It is one of the major internet sale websites working in New Zealand. Trade Me was founded by the New Zealand tycoon Sam Morgan in 1999. In 2015 according to Alexa Internet Trade Me was fifth most visited website in New Zealand. In a country having 5 million populations Trade Me website has 3.7 million active members and approximately eight hundred thousand visit websites on daily basis. Trade Me provides broad range of categories of new and used items. Trade Me has increased its scope over time and Trade Me now offers a wide range of listing possibilities general items, jobs, Motors, Rental property, pets and animals, Antiques & Collectables, Flatmates wanted.

The main scope and objectives of this research study is to create regression test suite and automating the test suite for testing the functionalities of Trade Me is working fine after the modification or not. This is designed for checking login, search, jobs, sort, add to watch list, remove from the watch list, logout functionalities. Trade Me have enormous number of features and is a fast-growing website approximately eight hundred thousand people are using this website on daily basis for buying and selling things. As Trade Me is a quick growing website new functionality are getting implemented frequently so to test the old functionalities are affected because of adding new functionalities we do automation regression suite.

Trade Me is rapid growing website having wide range of categories of new and used products for the vendors. As Trade Me is rapidly growing website, new functionalities are added frequently and test the old functionalities are not affected while addition of new functionalities, we do regression testing. The problem found in the Trade Me website is manual regression testing on Trade Me website is taking too long and it is monotonous job. Manual regression testing will take more time compared to automation regression suite. Because Trade Me includes much

functionality, retesting all the functionalities after a change is time consuming. While in automation, creation of test scripts takes time, but we can run it efficiently in no time. So, my goal is to automate the regression suite for Trade Me website. As full automation of Trade Me website is not possible, the main functionalities are prioritized. This will not only help to decrease the testing time for every release, but it will also help to free up the manual test resources for quality and exploratory testing.

This research paper is organized as follow: Section 2 focuses on the literature review of the automation regression testing. Section 3 is focused on the research methodology for this research. Section 4 explains execution for this research and results are discussed in section 5. Section 6 provides the discussion on the results of this research. In section 7 recommendations for future research is provided. Finally, in section 8 conclusion to the research is provided.

2. LITERATURE REVIEW

Various research papers have been referred and analysed before implementing the methodology or using any tool and technique for this research. Test case prioritization was conducted which involves scheduling test cases in order to achieve effectiveness of performance goals [1]. In this study they mentioned that Test cases should be performed in such a manner that enhances the possibility of fault detection. In this study they build up validate requirement-based system level which is used for test case prioritization method to expose more harsh faults at the beginning stage.

Another study was conducted which showed regression testing is performed to give assurance that the alterations which are made in the existing software do not harm the live behaviour of the software [2]. In this study they mentioned that if the test suites tend to grow as software progresses, so it becomes too expensive to accomplish the entire test suites. In this study regression testing is also used for prioritizing the test cases in such a way that it maximized the fault detection in the earlier stage.

In a study it was determined that regression testing is to evaluate the adjustments made to software while adding new features have not been affected to the software [3]. In this study they revealed many techniques which have been reported on how to choose the regression tests so that quantity of test cases does not expand up too big as the software grows. In this study they projected hybrid technique combines minimization, modification, and prioritization-based variety using a list of source code changes and carrying out traces from test cases run on earlier versions.

Test case Prioritization techniques plan test cases study was conducted for implementation in order that attempt to make best use of some objective function. There are various objective functions are appropriate, one such function involves pace of fault detection a measure of how speedily faults are detected within the testing process [4]. In this study they stated about how improved pace of error detection during regression testing can give quicker feedback on a system under regression test and let debuggers start their work in advance. In this study they forecast several techniques for prioritizing test cases and report the efficiency of this techniques for improving pace of fault detection.

A study was conducted on test case prioritization for continuous regression testing is limited by fixed time constraints. To ensure time constraints and accomplish test cases, testing goals must be well planned in execution [5]. Prioritization techniques are usually used to command test cases to mirror their significance according to one or more criteria. High fault detection or reduced time test are vital part. In this study they proposed test prioritization approach Rocket to get better effectiveness of continuous regression testing of industrial video conferencing software. The outcome of the study shows that using Rocket it provides quicker fault detection, it also improves

regression fault detection rate. It also reveals that “30% more faults for 20% of the test suite executed” comparing to manually prioritized test cases.

A further study was conducted on test case prioritization for web service regression testing is vital to confirm the worth of service-oriented business application in their evolutions [6]. In this study they mentioned how the test case prioritization technique plays vital role to boost the effectiveness of web service application regression testing. In this study they analyse the reliance relationship using control and data flow information in an orchestration language: WS-BPEL. They also prioritize the test cases according to wrap more alteration affected elements with the maximum weight.

A study was conducted to comprehend regression testing techniques which are most vital phase in the software development life cycle [7]. In the maintenance phase of software development life cycle the development team maintains the software and delivered to clients. In this study they mentioned about the software maintenance findings for the reasons like enhancement of capabilities, deletion of capabilities, error corrections and optimization. In this study they have presented a variety of regression testing techniques and their classifications offered by different researchers and elaborating selective and prioritizing test cases for regression testing in brief.

A research was conducted to test prioritization using system model's regression testing is important to check customized system is retested using existing test suite. The size of the test suite may be very huge, and testers are fascinated in defecting the faults in the earlier stage during the retesting process [8]. In this study they mentioned about the existing prioritization methods which are totally based on the code of the system. In this study they have introduced variety of methods of test case prioritization which are based on state-based models after changes to the model and the system. The model is implemented for the test suite and information about model implementation is used to prioritize tests.

A further research was conducted on test case prioritization approaches in regression testing, a systematic literature review states that quality of the software can be assured by going through software testing process [9]. In this study they talked about how regression testing phase is a costly process as it consumes a longer time. In this study they come up with the solution by scheduling test cases implementation with the help of prioritization approach.

Another research was conducted on prioritization for regression testing using ant colony optimization based on test factors, regression testing is believed to be one of the most expensive, time-taking, and vital activity which is executed in an environment with number of limitations for confirming the strength of a modified software [10]. In this study they talked about why it is necessary to select the sequence of test cases which is accurate, and it can traverse all the faults and take minimum execution time. In this study they come up with the solution that prioritization supports to accomplish performance requirements in which only the vital test cases are implemented. Previous studies have shown that prioritization approaches based on test factors like complexity, fault rate, importance, volatility, time, coverage have good outcomes and it also approaches based on intelligent practices like ant colony, genetic algorithm have been encouraging.

After going through various research on various aspects of regression testing and test prioritisation our aim for this research is to find whether the functionalities of the Trade Me website is working fine or not after inserting new features and for this we need to use certain tools which are discussed in next section.

3. RESEARCH METHODOLOGY

Selenium Web driver is used to build this automation framework. In this research automation framework for a web application has been designed and implemented. Using these framework testers can easily write and execute the test cases effectively and in less span of time. This framework is supportive to developer as well as tester to analyse their code and it produces the customized report to tester [11]. For browser testing we can use selenium because it is an open source tool [12]. Selenium web driver has been selected for this research and why it is selected it is shown below in table 1.

Table 1: Comparisons of Tools

| Parameter | Selenium Webdriver | QTP | Test Complete |
|-------------------------------|--|---|--|
| Programming Language | Scripts are normally designed using ruby, java, python, Perl, and PHP. | It only supports VBScript and JavaScript. | It Allows scripts to be designed in C#, C++, VBScript, JS Script. |
| Platform Supported | MAC, UNIX, Windows operating system. | It only supports windows. | It runs on windows Vista and 7. |
| Open source/ Paid Tool | Open Source | It is a paid tool, and it is based on the project. | It cost lesser than QTP and it is licensed tool. |
| Test Execution Report | It does not generate the report by itself we need plugin like TestNG to execute and generate reports in html format. | Using QTP we can generate test execution report and determine the execution status. | Execution results are generated as a separate file and displayed to the user [13]. |

For this research scrum methodology is selected because of its nonstop iteration for the development and testing throughout the life cycle of the software development process. The Sprint last for 2 weeks. The team member has to co-ordinate with the scrum master for the daily report via email or face to face meeting when needed. The scrum master reviews the work and provides suggestion. At the end of the Sprint there is retrospective meeting to see development and testing progress.

Scrum methodology has been opted for this research. While using traditional SDLC are not able to meet the market requirement and leading to customer dissatisfaction, as customer requirements are changing frequently and making it more complex for the testers. Scrum methodology has been introduced to meet the desires of the software development companies like Trade Me [14].

4. RESEARCH EXECUTION

4.1. Automation Test Plan

It is time consuming to write the automation scripts. To automate everything is not possible so spending money and time on automation is to develop a strategy that boost up the velocity for the short and long term.

Repetitive Test

- As we know login is repetitive test because we are calling login several times on the website.

High Risk

- Auto-updating of the web-browser is high risk.
- Using X-path sometimes creates a lot of error because it does not call the expected value.

Test Approach

- Selecting the appropriate automated testing tool for testing Trade Me website.
- Test cases should be given test priority, which test case has high priority should be automated first, test cases having low priority should be tested at the end.
- Test early and Test often.

Test automation phases

- Requirement analysis: - Functional requirements of Trade Me.
- Test Planning: - Functional Automation planning.
- Test case design and development: - Planning test cases and prioritizing the test cases according to the priority (High, Medium, Low).
- Test environment: - Installing Selenium, Xampp, TestNG for generating report and Eclipse.
- Test execution: - Scripts Implementation.

Roles and Responsibilities

Table 2: Roles and Responsibilities

| Members | Task | Time |
|---------------------|---------------------------------|---------|
| 1.Automation Tester | For writing automation scripts. | 6 hours |
| 2.Manual Tester | For writing the test cases | 5 Hours |

Test Environment

- **Software**
 1. Installing Selenium.
 2. Installing Xampp.
 3. TestNG for generating report.
 4. Eclipse.
 5. Database Server.
 6. Test Data.
- **Hardware**
 1. CPU 1.60 GHZ.
 2. Operating System: 64 Bit Windows 10
 3. Ram: 12 GB

Assumptions

- Testing environment stability.
- Test resources available.
- Stable application.

Risks

- Auto updating of the web-browser is high risk.
- Using X-path sometimes creates a lot of errors because it does not call the expected value.

4.2. Proposed Architecture for Automation Test Plan

The figure above justifies the architecture of automation test plan. There are 8-Page Factory Classes including login and logout which are connected to the test cases individually and that test cases are connected to the TestNG class. TestNG class is further divided in to two different parts. One TestNG class is connected with Selenium Web-Driver, Chrome Browser and test under execution. The second part is connected with the TestNG framework to generate reports and generate logs.

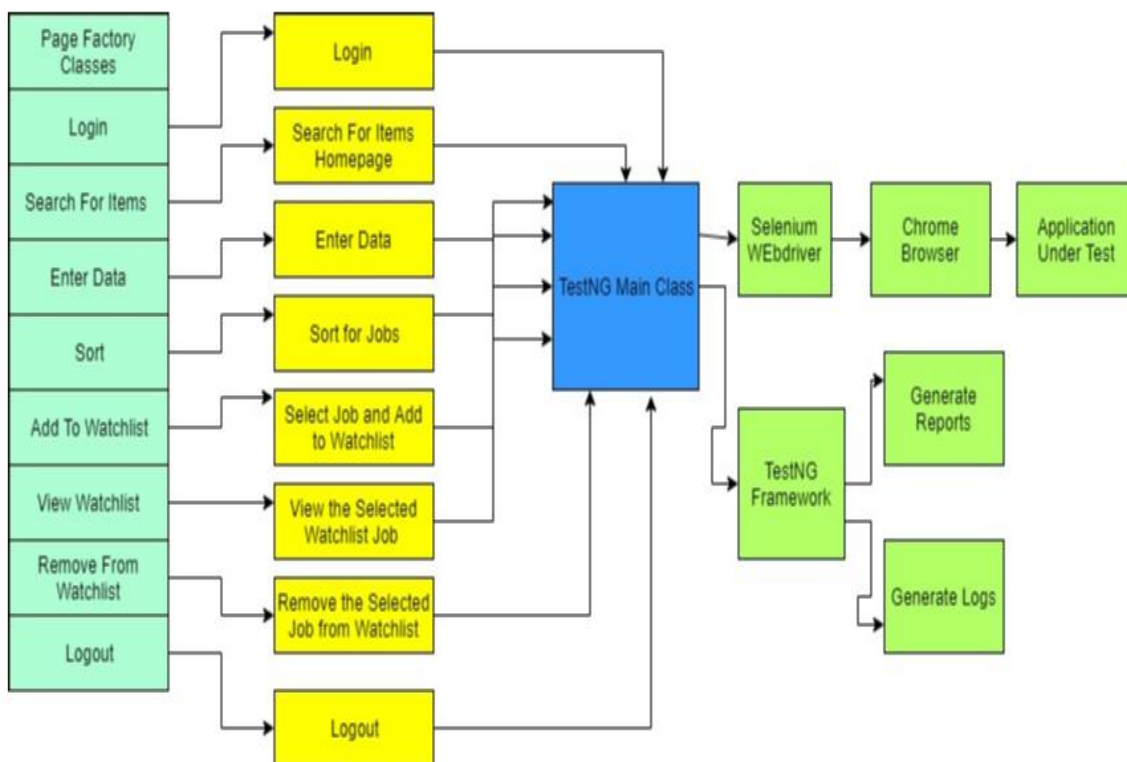


Figure 1: Architecture of Automation Test Plan

4.3. Functional Test Cases

The functional test cases are prioritized as per the requirement of research. Regression automation testing for the whole Trade Me website is not possible. So, the aim of the research is to prioritize the test cases which are listed below in table 3.

Table 3: Functional Automation Testing

| Test Scenarios ID | Test Scenarios | Test Steps | Expected Results | Actual Results | Pass/Fail |
|-------------------|---|---|--|--|-----------|
| TC-01 | Verify successful login functionality. | <ol style="list-style-type: none"> 1. Open browser enter "https://www.trademe.co.nz/Members/Login.aspx" 2. Enter valid "Username". 3. Enter valid "password". 4. Click on "Login Button". | User should be able to see the homepage. | User has been logged in successfully. | Pass |
| TC-02 | Verify that user can search for the items. | <ol style="list-style-type: none"> 1. Open browser enter "https://www.trademe.co.nz/Members/Login.aspx". 2. Enter valid credentials for login. 3. Search for "Jobs". 4. Click on "Search Button". | User should be able to see the "Trade Me Jobs HomePage". | User can see the Trade Me jobs homepage. | Pass |
| TC-03 | Verify that user can enter Keywords successfully. | <ol style="list-style-type: none"> 1. Open browser enter "https://www.trademe.co.nz/Members/Login.aspx". 2. Enter valid credentials for login. 3. Search for "Jobs". 4. Click on "Search Button". 5. Enter "Keywords". 6. Select "Job Type". 7. Select "Region". 8. Select "categories". 9. Select "Annual Income". 10. Select "Any Experience". 11. Search for the jobs. | User should be able to enter the valid details. | User can enter the details. | Pass |
| TC-04 | Verify that user can select "Sort". | <ol style="list-style-type: none"> 1. Open browser enter "https://www.trademe.co.nz/Members/Login.aspx". 2. Enter valid credentials for login. 3. Search for "Jobs". 4. Click on "Search Button". 5. Enter "Keywords". 6. Select "Job Type". 7. Select "Region". 8. Select "categories". 9. Select "Annual Income". 10. Select "Any Experience". 11. Search for the jobs. 12. Select "Featured Jobs". | User should be able to see the Featured jobs. | User can see the featured jobs. | Pass |

| | | | | | |
|--------------|---|--|---|--|------|
| TC-05 | Verify that user can "Add to Watchlist". | <ol style="list-style-type: none"> 1. Open browser enter "https://www.trademe.co.nz/Members/Login.aspx". 2. Enter valid credentials for login. 3. Search for "Jobs". 4. Click on "Search Button". 5. Enter "Keywords". 6. Select "Job Type". 7. Select "Region". 8. Select "categories". 9. Select "Annual Income". 10. Select "Any Experience". 11. Search for the jobs. 12. Select "Featured Jobs". 13. Click on "Add to Watchlist". | User should be able to add jobs to the watchlist. | User can add the jobs to the watchlist. | Pass |
| TC-06 | Verify that user can "View the Watchlist". | <ol style="list-style-type: none"> 1. Open browser enter "https://www.trademe.co.nz/Members/Login.aspx". 2. Enter valid credentials for login. 3. Search for "Jobs". 4. Click on "Search Button". 5. Enter "Keywords". 6. Select "Job Type". 7. Select "Region". 8. Select "categories". 9. Select "Annual Income". 10. Select "Any Experience". 11. Search for the jobs. 12. Select "Featured Jobs". 13. Click on "Add to Watchlist". 14. Click on "View Watchlist". | User should be able to view the watchlist. | User can view the watchlist. | Pass |
| TC-07 | Verify that user can "Remove from the Watchlist". | <ol style="list-style-type: none"> 1. Open browser enter "https://www.trademe.co.nz/Members/Login.aspx". 2. Enter valid credentials for login. 3. Search for "Jobs". 4. Click on "Search Button". 5. Enter "Keywords". 6. Select "Job Type". 7. Select "Region". 8. Select "categories". 9. Select "Annual Income". 10. Select "Any Experience". 11. Search for the jobs. 12. Select "Featured Jobs". 13. Click on "Add to Watchlist". 14. Click on "View Watchlist". 15. Click on "Remove from Watchlist". | User should be able to Remove from the watchlist. | User can remove jobs from the watchlist. | Pass |

| | | | | | |
|--------------|---|--|---|--|------|
| TC-08 | Verify that user can "Logout Sucessfully" | <ol style="list-style-type: none"> 1. Open browser enter "https://www.trademe.co.nz/Members/Login.aspx". 2. Enter valid credentials for login. 3. Search for "Jobs". 4. Click on "Search Button". 5. Enter "Keywords". 6. Select "Job Type". 7. Select "Region". 8. Select "categories". 9. Select "Annual Income". 10. Select "Any Experience". 11. Search for the jobs. 12. Select "Featured Jobs". 13. Click on "Add to Watchlist". 14. Click on "View Watchlist". 15. Click on "Remove from Watchlist". 16. Click on "Logout". | Verify that user has logged out successfully. | User has been logged out successfully. | Pass |
|--------------|---|--|---|--|------|

4.4. Gantt Chart for research

The Gantt chart used for this research is given in below Table 4. The Gantt chart in Table 4 explains brief about the Sprint 0 and Sprint 1. It explains activities from day 1 to day 25. The Gantt chart also have the start and the end date so we can get the clear idea about the Sprints and have the duration that which Sprint Activity took more hours to complete.

Table 4: Gantt chart for research

| Sprint | Actions/Deliverables | Day | Start date | End Date | Durations |
|-----------------|--|-----|------------|----------|-----------|
| Sprint 0 | Selecting and understanding the research. | 1 | 19/08/19 | 19/08/19 | 7 |
| | Understanding the plan of the proposal. | 2 | 20/08/19 | 20/08/19 | 6 |
| | Collecting the information related to proposal. | 3 | 21/08/19 | 21/08/19 | 5 |
| | Create plan for the proposal and check the literature for proposal. | 4 | 22/08/19 | 22/08/19 | 6 |
| | Meeting supervisor for the proposal feedback and changes made according to the guidance of the supervisor. | 5 | 23/08/19 | 23/08/19 | 6 |
| | Installing and setting up the environment. | 6 | 26/08/19 | 26/08/19 | 6 |
| Sprint 0 | Analysing the application. | 7 | 27/08/19 | 27/08/19 | 6 |
| | Analyse the technical architecture and design. | 8 | 28/08/19 | 28/08/19 | 6 |
| | According to business requirements create a test plan. | 9 | 29/08/19 | 29/08/19 | 5 |
| | Creating scenarios for regression testing. | 10 | 30/08/19 | 30/08/19 | 7 |
| Sprint 1 | Creating scripts for regression testing. | 11 | 02/09/19 | 02/09/19 | 6 |
| | Scripts execution | 12 | 03/09/19 | 03/09/19 | 8 |
| | Generate reports | 13 | 04/09/19 | 04/09/19 | 8 |
| | Meeting supervisor for the research report. | 14 | 05/09/19 | 05/09/19 | 4 |
| | Changes can be made according to the | 15 | 06/09/19 | 06/09/19 | 4 |

| | | | | | |
|-------------|--|----|----------|----------|----|
| | guidance of supervisor. | | | | |
| Sprint 2 | Gathering information for final report. | 16 | 09/09/18 | 10/09/18 | 10 |
| | Preparing the final report for the research. | 18 | 11/09/18 | 16/08/19 | 26 |
| | Preparing the final presentation. | 22 | 17/09/19 | 17/09/19 | 8 |
| | Research monitoring. | 23 | 18/09/19 | 18/09/19 | 6 |
| | Meeting supervisor for the feedback about the report and presentation. | 24 | 19/09/19 | 19/09/19 | 4 |
| | Final changes made as per the requirement of research supervisor. | 25 | 20/09/19 | 20/09/19 | 4 |

4.5. Snippets of code

The code snippets for the main class is displayed below in figure 2. @Test (annotations) in the figure states that all the test scenarios are implemented in a chronological way. (Priority=0) states the test case priority that login test case should be executed first. In the Main TestNG class sub-methods are created.

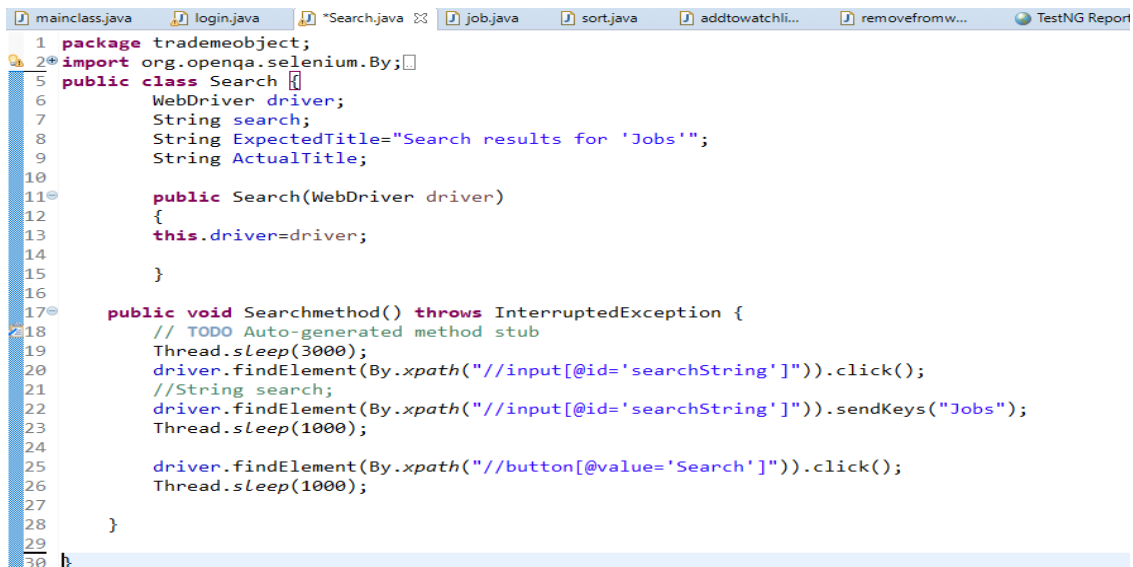
```

17 public class mainclass {
18     WebDriver driver;
19     @BeforeTest
20     public void launch() throws InterruptedException {
21         System.setProperty("webdriver.chrome.driver", "C:\\ChromeDriver\\chromedriver.exe");
22         driver = new ChromeDriver();
23         driver.get("https://www.trademe.co.nz/Members/Login.aspx");
24         Thread.sleep(2000);
25         driver.manage().window().maximize();
26         driver.manage().timeouts().implicitlyWait(10, TimeUnit.SECONDS);
27     }
28     // login
29     @Test(priority=0, enabled=true)
30     public void login() throws InterruptedException {
31         login loginmain=new login(driver);
32         loginmain.Username();
33         loginmain.Password();
34         loginmain.login();
35     }
36
37     // search
38     @Test(priority=1, enabled=true)
39     public void search() throws InterruptedException {
40         Search Searchmain=new Search(driver);
41         Searchmain.Searchmethod();
42         System.out.println(driver.getTitle());
43         //assertEquals(driver.getTitle(), "Jobs");
44
45         //Searchmain.search();
46     }

```

Figure 2: Snippets of Main class

The code snippets for search test suite are show in Figure 3. The Figure 3 below defines methods in each method and defined the actions for each method to be performed.



```

1 package trademeobject;
2 import org.openqa.selenium.By;
3
4 public class Search
5     WebDriver driver;
6     String search;
7     String ExpectedTitle="Search results for 'Jobs'";
8     String ActualTitle;
9
10
11 public Search(WebDriver driver)
12 {
13     this.driver=driver;
14 }
15
16
17 public void Searchmethod() throws InterruptedException {
18     // TODO Auto-generated method stub
19     Thread.sleep(3000);
20     driver.findElement(By.xpath("//input[id='searchString']")).click();
21     //String search;
22     driver.findElement(By.xpath("//input[id='searchString']")).sendKeys("Jobs");
23     Thread.sleep(1000);
24
25     driver.findElement(By.xpath("//button[@value='Search']")).click();
26     Thread.sleep(1000);
27 }
28
29
30

```

Figure 3: Snippets of Search Class

5. RESULTS

Research results of this study are given below.

5.1. TestNG Reports

TestNG is used to create testing reports and tester can easily come to know how many test are passed, failed and skipped. TestNG is an automation Framework where NG stands for next generation. TestNG is motivated from Junit which uses the annotations (@). Using TestNG tester can also execute the failed test cases separately. Also, TestNG provides a separate option “testing-failed.xml file” in the test output folder. So that if tester wants to run the failed test cases it can run with the help of XML file.

TestNG is used to capture the metrics which is used to improve the effectiveness and efficiency of a software testing process. The below points show how to capture TestNG report.

- TestNG will generate default report.
- After executing the java project should be refreshed to get test-output folder.
- Right clicking on the emailable-report.html.
- Same output can be seen in the web browser.

The Figure 4 shows that how many test cases are passed, skipped, failed. It also shows time taken for one test to complete. The Figure 4 defines that there are 7 number of test cases have been executed successfully. The number of test cases failed and skipped are 0.

| Test | # Passed | # Skipped | # Failed | Time (ms) | Included Groups | Excluded Groups |
|-----------------------|---------------------|---------------|-----------|-----------|-----------------|-----------------|
| Default suite | | | | | | |
| Default test | 7 | 0 | 0 | 100,540 | | |
| Class | Method | Start | Time (ms) | | | |
| Default suite | | | | | | |
| Default test — passed | | | | | | |
| tradememain.mainclass | addtowatchlist | 1568607253876 | 25351 | | | |
| | job | 1568607225414 | 23656 | | | |
| | login | 1568607210436 | 8623 | | | |
| | logoutm | 1568607290880 | 2639 | | | |
| | removefromwatchlist | 1568607279228 | 11650 | | | |
| | search | 1568607219061 | 6352 | | | |
| | sortby | 1568607249071 | 4804 | | | |

Figure 4: TestNG Reports

5.2. Chronological view

The chronological view of Trade Me main class is shown in the below Figure 5. The Figure 5 shows chronological view order of the test cases and what is the flow of test case. The flow of Trade Me main class is shown in the Figure 5.

| Class | Method | Time (ms) |
|-----------------------|---------------------|-----------|
| tradememain.mainclass | launch | 0 ms |
| | login | 16703 ms |
| | search | 25901 ms |
| | job | 32731 ms |
| | sortby | 55485 ms |
| | addtowatchlist | 59421 ms |
| | removefromwatchlist | 85148 ms |
| | logoutm | 97201 ms |
| | tearDown | 100065 ms |

Figure 5: Chronological View

5.3. Times

The Figure 6 shows time taken for each test to passed. It shows how much time taken by one test to get executed. Add to watchlist has taken 25,726(MS) approximately which is highest to execute one test. And logout took approximately 2863(MS) which least time to execute the test case.

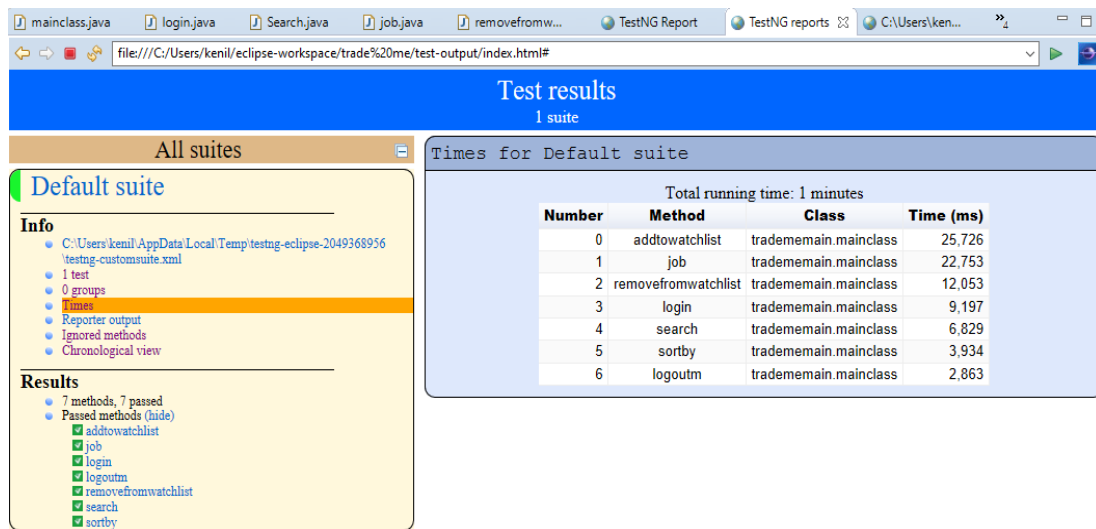


Figure 6: Times

6. DISCUSSION

The result shown in Figure 4 defines that, all the test scenarios has been passed. But what happens when one of the test cases failed? Let us take test scenarios search for the job has been failed. So, it will not have allowed me to do further regression testing on different scenarios. To overcome this challenge, we need to fix the error and run the test successfully. Using TestNG tester can run the failed test cases individually. TestNG also provides separate option like “testing-failed.xml file” in the test output folder. So that tester can run the failed test cases rather than running whole scripts which will help to reduce time. Let’s suppose what heppens if the test cases got skipped? Sometimes QA team needs to execute the last minute issue verification at that moment QA team does not have enough time to execute all the test cases so in this case QA team can skipped the test case by @Test(enabled=false).

The Chronological view of Trade Me website is shown in Figure 5. Chronological means in a specify order. The Figure shows the test suites in step by step order. What if it is not in chronological order? If the test suites are not in order the test cases will be failed. Test suites should be in a specify order to execute the test cases.

The time taken by each test to execute successfully is shown in the Figure 6. Add to watch list has taken 25,726 (MS) to execute the test which is highest and logout has taken least amount of time to execute the test which is 2,863 (MS). Now, what happens if the test case is taking too long or test cases get success? We can just simply add wait time between the element to get the accurate result and after the execution of all test cases we can simply decrease the wait time between the element so that we don’t have to wait everytime.

7. RECOMMENDATIONS

In this research regression automation test suite has been used to overcome the problems. First problem faced was automating the whole Trade Me website is not possible because it has lot of scenarios. To overcome this problem came up with the solution of prioritizing test cases. Prioritizing test cases helped to reduce the time and it also helped in reducing the cost. Second challenge faced during the execution phase. As, Trade Me is dynamic website we cannot automate with the help of X-path because it changes after every few minutes. To overcome this

problem ID and Name is used to identify the correct path. Regression automation test suite helped this research to reduce the time and it also helped to reduce the cost. It also helped to free up the manual test resources for quality and exploratory testing.

8. CONCLUSION

In this research we have covered the plan and the problem of the Trade Me website. For this research selenium Webdriver was selected and it helped to test the new features and processes of Trade Me website. I successfully implemented Regression testing which helped me to prioritize the test cases in such a way that it maximized the fault detection. Scrum methodology is used to meet the desires of the software development companies and increase the client satisfaction. The automation test plan, functional automation test cases, sprints, architecture of automation test plan explains the workflow in detail about the research. Regression testing is the best technique, but we cannot automate the “Trade Me” website fully. So, we prioritized the test cases and automate. This research will also help in future for similar kind of fast-growing websites.

REFERENCES

- [1] Raju, S., & Uma, G. V. (2012). Factors oriented test case prioritization technique in regression testing using genetic algorithm. *European Journal of Scientific Research*, 74(3), 389-402.
- [2] Yoo, S., & Harman, M. (2012). Regression testing minimization, selection and prioritization: a survey. *Software Testing, Verification and Reliability*, 22(2), 67-120.
- [3] Wong, W. E., Horgan, J. R., London, S., & Agrawal, H. (1997, November). A study of effective regression testing in practice. In *PROCEEDINGS The Eighth International Symposium On Software Reliability Engineering* (pp. 264-274). IEEE.
- [4] Rothermel, G., Untch, R. H., Chu, C., & Harrold, M. J. (1999, August). Test case prioritization: An empirical study. In *Proceedings IEEE International Conference on Software Maintenance-1999 (ICSM'99). Software Maintenance for Business Change* (Cat. No. 99CB36360) (pp. 179-188). IEEE.
- [5] Marijan, D., Gotlieb, A., & Sen, S. (2013, September). Test case prioritization for continuous regression testing: An industrial case study. In *2013 IEEE International Conference on Software Maintenance* (pp. 540-543). IEEE.
- [6] Chen, L., Wang, Z., Xu, L., Lu, H., & Xu, B. (2010, June). Test case prioritization for web service regression testing. In *2010 Fifth IEEE International Symposium on Service Oriented System Engineering* (pp. 173-178). IEEE.
- [7] Duggal, G., & Suri, B. (2008, March). Understanding regression testing techniques. In *Proceedings of 2nd National Conference on Challenges and Opportunities in Information Technology*.
- [8] Korel, B., Tahat, L. H., & Harman, M. (2005, September). Test prioritization using system models. In *21st IEEE International Conference on Software Maintenance (ICSM'05)* (pp. 559-568). IEEE.
- [9] Khatibsyarbini, M., Isa, M. A., Jawawi, D. N., & Tumeng, R. (2018). Test case prioritization approaches in regression testing: A systematic literature review. *Information and Software Technology*, 93, 74-93.
- [10] Ahmad, S. F., Singh, D. K., & Suman, P. (2018). Prioritization for regression testing using ant colony optimization based on test factors. In *Intelligent Communication, Control and Devices* (pp. 1353-1360). Springer, Singapore.
- [11] Gojare, S., Joshi, R., & Gaigaware, D. (2015). Analysis and design of selenium webdriver automation testing framework. *Procedia Computer Science*, 50, 341-346.
- [12] Holmes, A., & Kellogg, M. (2006, July). Automating functional tests using selenium. In *AGILE 2006 (AGILE'06)* (pp. 6-pp). IEEE.
- [13] Jain, C. R., & Kaluri, R. (2015). Design of automation scripts execution application for selenium webdriver and test NG framework. *ARNP J EngApplSci*, 10, 2440-2445.
- [14] Mahalakshmi, M., & Sundararajan, M. (2013). Traditional SDLC Vs Scrum Methodology—A Comparative Study. *International Journal of Emerging Technology and Advanced Engineering*, 3(6), 192-196.

AUTHORS

Kenil is a graduate of BE computer science engineering in India. He moved to New Zealand in 2018 to pursue his career in Graduate Diploma in Software Testing at AGI Education limited. He started exploring about testing concepts and practices. He has an interest in Functional automation testing as well as performance Testing.



Dr. Shahid Ali is IT program leader and senior lecturer at AGI Education Limited, Auckland, New Zealand. He has published number of research papers on ensemble learning. His expertise and research interests include ensemble learning, machine learning, data mining and knowledge discovery.

© 2021 By AIRCC Publishing Corporation. This article is published under the Creative Commons Attribution (CC BY) license.

ON-DEMAND VIDEO TAGGING, ANNOTATION, AND SEGMENTATION IN LECTURE RECORDINGS TO ENHANCE E-LEARNING EFFECTIVENESS

Ken D. Nguyen and Muhammad Asadur Rahman

Department Dept of Computer Science & Information Technology,
Clayton State University, Georgia, USA

ABSTRACT

The COVID-19 pandemic has forced much of the academic world to transition into online operations and online learning. Interactions between the teachers and students are carried out via online video conferencing software where possible. All video conferencing software available today is designed for general usage and not for classroom teaching and learning. In this study, we analyzed the features and effectiveness of more than a dozen major video conferencing software that are being used to replace the physical face-to-face learning experiences. While some of the video conferencing software has pause feature but none allow annotation and segmentation of the recording. We propose tagging and annotation during the live streaming to improve direct access to any portion of the recorded video. We also propose automatic segmentation of the video based on the tagging so that the video is short, targeted, and can easily be identified.

KEYWORDS

Live and On-Demand Video, Video Tagging, Video Annotation, Video Segmentation, E-Learning, Remote Learning, Synchronous Lectures, User Interface Design, Human-Computer Interaction.

1. INTRODUCTION

Since the outbreak of the COVID-19 pandemic, the ways in which we teach and learn have shifted dramatically from largely in-person to completely virtual. Teaching in the fields of computer science and technology often requires close interaction between students and instructors, and online teaching and learning require tremendous organization and discipline. Furthermore, effective online education greatly depends on the vehicle in which a course is delivered. Software tools that are used in synchronous or asynchronous delivery of courses must have user-friendly interfaces to promote effective and timely interaction between instructors and students.

The COVID-19 pandemic has shifted much of the US education system onto remote learning, changing from instructor-led face-to-face interactive mode to distance learning via online web learning management systems (LMS) such as Desire2Learn, KMI LMS, GoToTraining, Grovo, Edwiser remUI, Moodle, Edmego, SmarterU, Paradiso Solutions, DHx Software, MindQuest Learning, Brainier, Talent LMS [1,2,3,4,5,6,7,8,9,12,13,14] ScholarLMS, Eurekos LMS, InnTier, Knolyx, Wisetail, Learnupon, GO[15], etc.

Video lectures and synchronous video conferencing are two primary mechanisms for instructors to emulate the in person native interactive environment of teaching and learning. A recent study [16] conducted on a random control monitoring of 157 students in a for-credit online course at a 4-year university found that on average, students watch all assigned and required videos with embedded assessment 1.29 times, about 9% of videos are skipped by students, and about 2% of the videos are watched after the assigned deadline. These figures indicate that students mostly watch the videos once to complete the assignment to get the grade and then almost completely ignore the video. This intriguing fact seems to contradict the common assumption that a recorded lecture would allow the student to review the material multiple times which could lead to better understanding of course materials and enhance their performance. In addition, in the studies [17,18], recorded lectures seem to support significant improvement in student performance if they are used as supplements.

Studies have shown that segmenting videos of 6 minutes or less in length is more effective for leaning [17,18,19,20], It allows learners to engage with small pieces of new information, gives them control over the flow of information, and helps encode the information into long-term memory. Long videos are ineffective as they overload the intrinsic cognitive capability of learning working memory [19,20,21,22,23].

While remote learning is not new, effective remote learning has always been a challenge as most younger and inexperienced learners need constant guidance; moreover, courses with higher levels of abstraction and logical thinking such as in the STEM (Science, Technology, Engineering, Mathematics) fields require repetitive guidance and often hands-on experience, thus requiring instructor-led class time as seen in most education institutions in the US and around the world. Because on-site operations at educational institutions in the US are now closed to slow down the spread of the coronavirus, most instructors depend on video conferencing tools such as Zoom, Microsoft Teams, WebEx, GoogleMeet, GoToMeeting, TeamViewer, YouTube, Facebook, Tiktok, Telegram, Slack, Discord and the likes to replicate direct hands-on face-to-face guidance during remote learning. This switch is a big challenge in teaching as instructors are often not well-equipped or trained, nor do they have the time to become effective video content makers and video editors. Moreover, none of the listed video tools are capable of or designed for handling video conferencing in academic environment.

Table 1. Supported features by various online synchronous video meeting software.

| Software | Recording | Pause/Resume | Annotated | Video Segmentation | Transcription | Chat | Multi-User |
|----------------------|-----------|--------------|-----------|--------------------|---------------|------|------------|
| Microsoft Teams | Yes | No | No | No | Yes | Yes | Yes |
| Zoom | Yes | Yes | No | No | Yes | Yes | Yes |
| Google Meet | Yes | No | No | No | Yes | Yes | Yes |
| WebEx | Yes | Yes | No | No | Yes | Yes | Yes |
| Team Viewer | Yes | Yes | No | No | No | Yes | Limited |
| Telegram | Limited | No | No | No | No | Yes | Limited |
| Facebook Live Stream | Yes | Yes | No | No | No | Yes | Limited |
| Goto Meeting | Yes | No | No | No | Yes | Yes | Yes |
| Discord | No | No | No | No | No | Yes | Limited |
| Slack | No | N/A | No | No | No | Yes | Limited |
| Youtube Live | Yes | No | Yes | No | Yes | Yes | No |
| Tiktok | No | No | No | No | No | No | No |

Table 1 depicts a summary of our review of popular software that support online video recording and posting. We found that none of the popular video conferencing tools being used in education have proper interface and functionality to allow the instructors effectively create segmented instructional videos for use in their courses. From our study, we propose new feature and interface design to provide the much needed and effective tool for instructors to create targeted and topic specific videos to help promote effective remote teaching and learning.

2. RECORDING TAGGING AND ANNOTATING PROPOSAL

Synchronous video class meetings between faculty and students tend to cover many topics that span over an hour or longer. The recordings of these meetings could be more effectively used if they could be fragmented into individual segments such that each segment is focused on a specific item, topic, question, or idea. There should be an easy option for quickly marking the starting and ending of each fragment on the fly during class discussion. The author should be able to go back to the recorded video to adjust the markers and then extract the fragments as smaller videos for posting.

2.1. New Features to Consider

In our study, we look at Microsoft Teams, Zoom, WebEx, and Google Meets as the most popular conferencing software being used in academia because they can handle many users simultaneously and can integrate with institutional digital infrastructure. These software allow the user to trim the recordings and have the recordings transcribed. They also have a good auto-generated transcript service for recording videos. Zoom and WebEx allow the recording to be paused and resumed during the meeting, thus allowing some of the discussion to be included in the recordings. While these software have many great features, the final product – the recorded meeting videos – are often long and not effective for learning in academia. While the instructor can prepare a pre-recorded video using a camcorder or a recording software such as Kaltura Capture[24] or similar software, these videos do not represent the class dynamic or the specific topics and concepts in a way that the students would approach in real time. In addition, the majority of course-specific materials are specific to individuals, groups of students, and each semester. This issue is intensified much more in courses that have frequent updates to materials such as IT and CS. Therefore, synchronous meetings and recorded meeting videos are still very popular. Instructors are often faced with the decision to edit and fragment videos or leave them as-is. The second option is often the choice as time, resources, and training are limited.

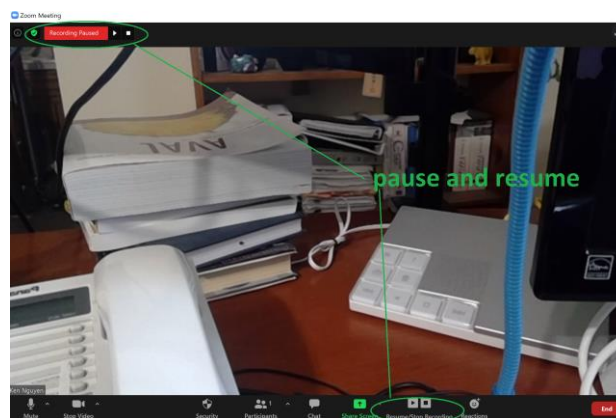


Figure 1. Zoom meeting with Pause/Resume option

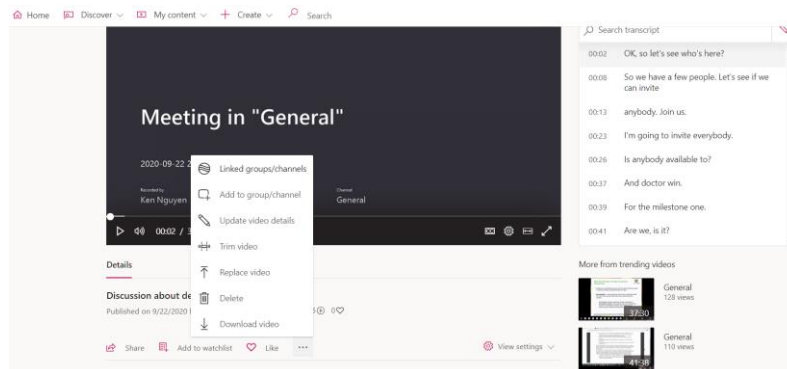


Figure 2. Microsoft Teams Streaming Transcription

Moreover, video editing software such as Adobe Premiere and others require tremendous amounts of time and effort, and video editing and video production are skills in which most instructors are not trained or exposed to. Figures 1 shows a screen capture of Zoom synchronous video meeting and Figure 2 shows the transcription of a Microsoft Teams recording.

Moreover, the option to pause the recording has a big impact as to which institutions can and cannot use the software. Without the pause option, the recording will be contiguous, thus making the recording more transparent and easier to verify any possible allegation that arises from attendees. Not allowing pausing on the recording would allow tracing and tracking of any potential legal claim against the meeting participants. Any editing of the original recording would show the alteration, doctoring, or tampering of the evidence. Therefore, software such as Microsoft Teams are the primary choice for most public institutions. Not allowing pausing on the recording would allow tracing and tracking of any potential legal claim against the meeting participants. On the other hand, having the pause option allows certain parts of the meeting to be excluded from the recording; thus, this option works well for private and nonessential meetings. However, in almost all instructional synchronous meetings, there are many times a topic is reiterated, so the same questions may be asked by different individuals in the same meetings, and some repetitive and insignificant discussions are recorded during lectures. Thus, having all of these in the recording videos makes them very long, and students lose focus and interest. We want to have a feature that can automate the fragmentation process and retain the best of these two models.

2.2. Proposal to Enhance Video Recording with Targeted Tagging and Annotations

In terms of interrupting a recording, there could be two models of video conference recording: (a) those with a pause option such as Zoom and WebEx, and (b) those without a pause option. The platforms with the pause option allow the user to skip recording of certain periods of time when the discussion is not directly related to or of particular significance for those who may review the video later. These tools are good for personal and private environments, where there is less scrutiny and little chance that legal action may arise or be taken for or against what is being recorded. The second model, without the pause option, is often appropriate for public use and institutions involving many stakeholders. Thus, the recording should be continuous and untampered so that it can withstand all scrutiny, showing transparency, and can be used for legal purposes.

2.2.1. Designs

While these video conferencing applications are excellent at facilitating synchronous meetings, they lack certain features to support an effective online learning environment. For example, students generally lose focus when the video duration exceeds 6 minutes; that amount of time or less in length is more effective for learning[19]. Thus, having long recording sessions usually discourages students from reviewing the videos. Secondly, long recordings generally cover a lot of material and contain many interweaving concepts. It is best to make short video segments for each concept paired with illustrative examples. However, video making requires extensive amounts of preparation time and video editing skills that most faculty would not have, especially for those who have been teaching face-to-face and now are suddenly thrown into fully online instruction.

Our proposed design enhances these video conferencing applications so that the user can make annotations during the recording of conference meetings and have the recordings segmented into shorter videos automatically. The user then can make these video into shorter segments with specific context to their students. For video conference recording software that do not have the “pause” feature, we propose the recording enhancement as seen in Figure 3.

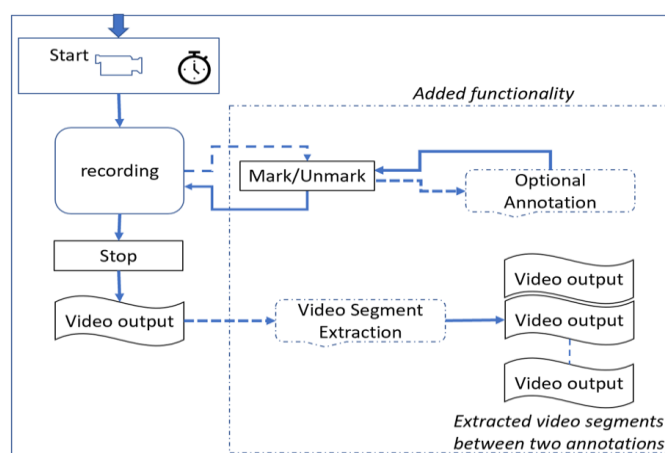


Figure 3. Enhanced recording feature for video conference software without a “pause” feature

The recording can run continuously from the moment it is started until the stop option is triggered. However, in our model, a permanent alternating “mark” and “unmark” button will be displayed on the interface, allowing the user to trigger segmentation markings on the recording. The button will alternate between mark and unmark. This option allows the user to place an annotation mark on the video at the moment of choice. A semi-transparent fading balloon for annotation will appear for a few seconds following the “mark,” allowing the user to type in a short title or description for the marked segment. If the user ignores the balloon, the first transcribed sentence immediately following the marking will be used to annotate the title of the recorded segment. This feature allows the user to keep full attention on video conferencing without distraction. The user can edit the automated text after the video recording session is finished.

For video conference recording software that do not have a “pause” feature, we propose the recording enhancement as seen in Figure 4.

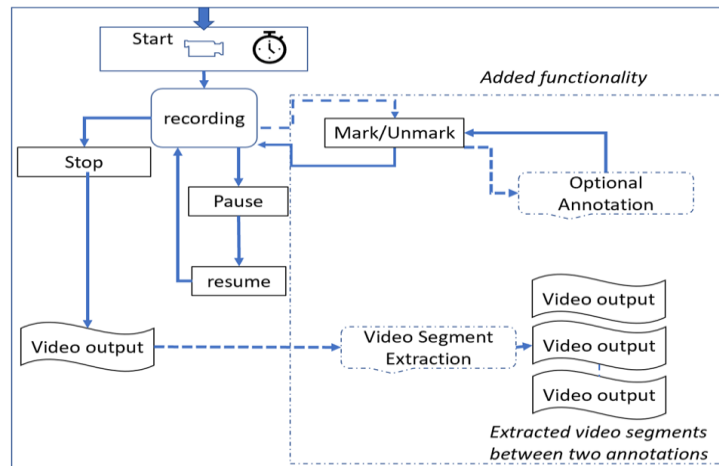


Figure 4: Enhanced recording feature for video conference with 'pause' feature

2.2.2. Recording with Enhanced Video Tagging for Marking and Annotating Segments of Recording Video

Enhanced tagging for the marking/unmarking and annotating feature would run in conjunction with the traditional “pause” option in any video recording service, where the user can put the recording on hold and resume the recording when ready. The pause and resume functions could share the same button with alternating options displayed. However, the user will be presented with a semi-transparent fading pop-up balloon/screen which can be annotated at the beginning of a recording segment. The user can either choose to click on the balloon and enter some text to annotate or describe what would be immediately discussed in the video right after the resume option is triggered or to ignore it completely. The balloon will fade away in a few seconds if the user is not clicking on it; however, an annotation button will appear next to the resume button so that the user can pull up the annotation text box.

Since the recording can be paused, the user would have full control of the recording. Thus, it is not necessary to add a new button that takes up precious interface space, especially when the user is using small mobile devices such as a smart phone. For this model, the pause button would function as a traditional pausing option. The enhancement appears when the user clicks on the annotated balloon and a message can be added. We also propose to have an “annotated” button to be displayed on the control bar which the balloon fades into. This button allows the user to bring back the annotation balloon to annotate the beginning of a recording segment.

2.2.3. Automatic Video Segmentation and Extraction with the New Tagging and Annotation Feature

This option is added into both models, giving the user an option to extract video recording segments between annotations as individual videos. In online education, this feature is a great time saving tool for preparing teaching materials and learning. An example use case would be as follows: a computer science professor describes the concept of a data structure, its usage, and how to implement it to students. Thus, he would have examples to illustrate the concept. The examples would be very appropriate and precise due to his experience and knowledge in the field. However, these kinds of examples may not relate well with students as they do not have the same background as the professor, and the precision of the examples may require a certain level of influence, ability to recall, utilize and readily apply previously learned concepts. Thus, when the user makes his own video lecture, the students may still not be able to comprehend it very

clearly. A discussion is usually followed by the posting of the lecture so students can ask for clarification or alternate explanations. These kinds of questions are individual and specific to help students to grasp the concept. Thus, the professor then responds to the students' questions with specific examples and illustrations. These answers are often more related to the students. Thus, having a way to extract these specific illustrations and place them along with the lecturing concept will help the students to learn and allow the professor to utilize their time effectively. This use case assumes that the professor is given a time slot to interact with students. With current video conference tools, the professor does not have that option. He may have to make several videos for each of these examples. Certainly, these video segments will not be professionally done; however, most professors are not professional video editors anyway.

2.2.4. Legal Implication of Contiguous Recordings

The option to record a meeting without a pause option allows institutions and corporations to have a continuous and unaltered record that can withstand potential legal tests when there are complaints about what has been said or shared during the meeting. With our proposal, the original recordings are the same as the original system would produce, and our proposal allows the system to generate a by-product that is beneficial and supportive to an online academic environment.

2.3. Implementing the Enhanced Tagging and Annotation Features

To implement these proposed features with minimal modification to the system and without changing its structure, we propose adding only a button onto the interface to provide the user with a way to manually trigger the new feature. Figure 5 depicts the implementation of the design where an auxiliary file is being used to record all annotations during the recording session in sync with the recording clock.

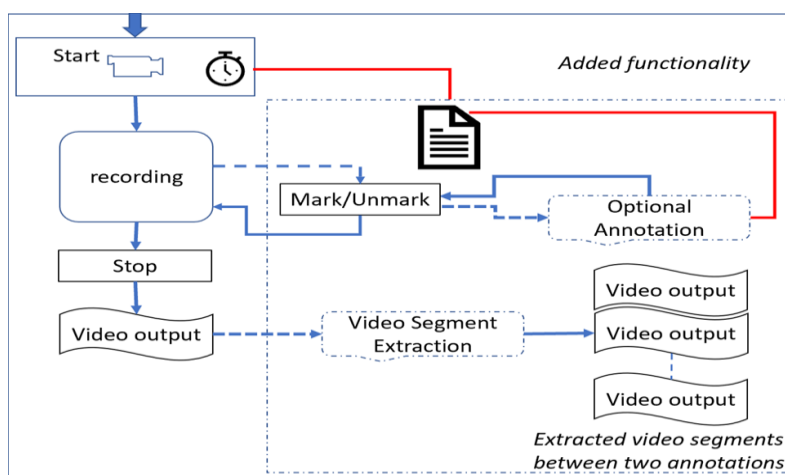


Figure 5: Enhanced recording feature for video conference with 'pause' feature

All annotations can be done as a floating semi-transparent text box interface. To support the creation of the video segments, we will rely on the timer and a description file. In general, most systems employ a clock to show elapsed time in the recording, which is also the length of the recording, and we will use the same clock to support the new feature. When the new button is triggered, the system recording duration time is saved. Similarly, every interaction with the new feature will be recorded with the content given, and the time relative to the video recording. For example, two minutes into the recording, the user presses the "mark/unmark" button, annotates a

message 3 seconds later, and then clicks the “pause/unmark” button 5 minutes after the start of the annotation. This allows the system to generate a 5 minute and 3 second long recording segment starting 2 minutes from the beginning of the recording. An example entry for the description file is:

```
2:00    2:03    the message    7:03
```

The message can either be superimposed on the beginning of the segment or kept as the beginning of the segment transcription.

Multiple annotations: To handle multiple annotations on the same segment, the user can always annotate the video by typing into the semi-transparent text box and hit “enter” to save the message. The system records the time when the first character is entered into the text box and the message for the future display in the video segment. For example, at 2 minutes into the recording, the user clicks the mark/unmark button, then types in a message 5 seconds later; then the user types another message 1 minute later and clicks the unmark button at 9 minutes into the recording. The system will record the following:

```
2:00    2:05    the first message    3:05    the second message    9:00
```

Thus, with this technique, the user can annotate any number of messages, each value is separated by a tab, and the first and last value of the entry indicates the time interval from which the original recording will be copied to create a new recording segment.

When the recording is stopped, the system will use these entries in the description file to generate the recording segments from the original recorded video, and the messages can be superimposed on each segment individually at the corresponding recorded time period.

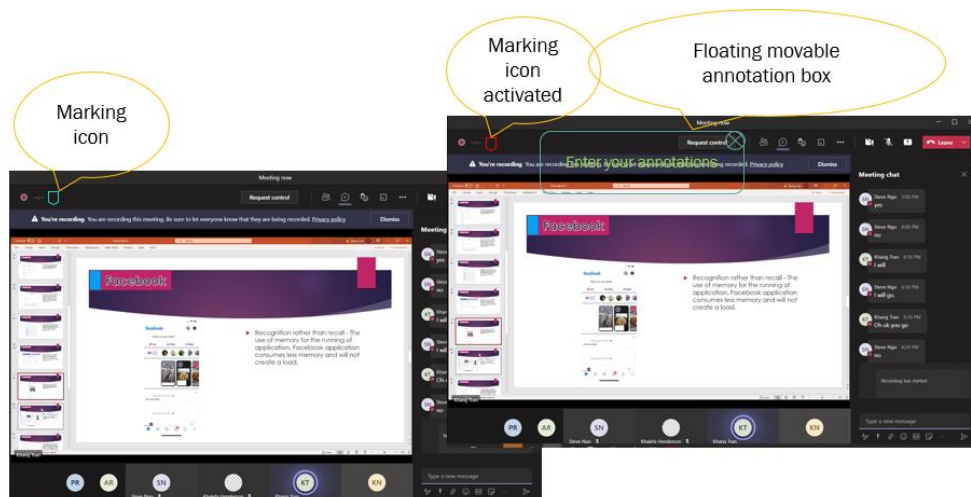


Figure 6: Interface for segment marking and annotation features

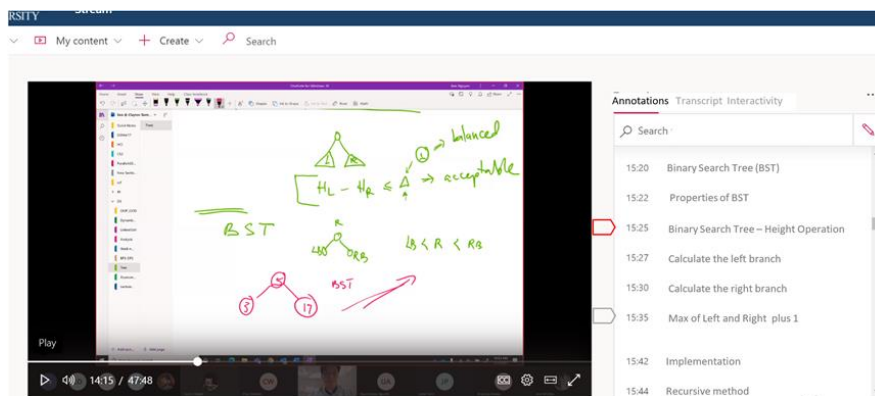


Figure 7. Segment Marking and annotations on Recording Videos

The user interface design for the segmentation marking and annotation features of the prototype system is given in Figure 6. The marking icon when clicked is activated which would become red to indicate the start of the segment. Clicking it again would mark the end of the segment. The annotation box is a floating box where annotation can be written. Figure 7 shows active marker and annotations on a recording.

3. CONCLUSIONS

In this study we analyzed over a dozen synchronous video meetings and conferencing software. We looked at the video delivery user interfaces that would be most appropriate and effective for student learning and engagement. From our study and recent research, we concluded that none of these software has the video features necessary for effective remote learning. We proposed features that would seamlessly integrate with any video conferencing software allowing the user to annotate the recording of the meeting on the fly, on-demand and obtain automated segmentation of the video. The segmentation feature would allow topic centric videos to be generated automatically from the recording without requiring further editing and minimal intervention. These new features can be added on to the existing video conferencing software without changing the underlying infrastructure. As future research, we would like to analyze and compare the effectiveness of these auto-generated targeted video segments against specialized video content created for similar topics.

REFERENCES

- [1] Tovutilms. "Learning management System". <https://www.tovutilms.com/> Date accessed January 28, 2021.
- [2] ELucid. "Build a Solid Corporate Training Culture with ELucid"., <https://elucidlearning.co/>. Accessed 26 Jan. 2021.
- [3] "KMI Learning". Custom ELearning Solutions., <https://www.kmlearning.com/>. Accessed 26 Mar. 2021.
- [4] GoToMeeting. "Online Meeting Software, Video Conferencing & Web Conferencing". <https://www.gotomeeting.com/>. Accessed 10 Dec. 2020.
- [5] Grovo. "The Microlearning Company.", 20 Dec. 2020, <https://www.grovo.com>
- [6] Edwiser. "LMS Solutions Under One Roof!". <https://edwiser.org/>. Accessed 5 Nov. 2020.
- [7] Moodle. "Open-Source Learning Platform". <https://moodle.org/>. Accessed 01 Oct. 2020.
- [8] Edmego Learning. "Edmego Learning Management System., 11 Mar. 2016, <http://www.edmegolearning.com/>. Accessed 18 Oct. 2020.
- [9] SmarterU. "Learning Management System | LMS". <https://www.smarteru.com/>. Accessed 12 Feb 2021.
- [10] Paradiso LMS. "Best Learning Management System (LMS) Software". <https://www.paradisosolutions.com/>. Accessed 30 Jan. 2021.

- [11] DHx Software. “Digital Learning Management System (LMS)”. <https://www.dhxsoftware.com/elearning>. Accessed 7 Feb. 2021.
- [12] LMS.Org. “Learning Management System Reviews”. <https://www.lms.org/>. Accessed 2 Mar. 2021.
- [13] TalentLMS. “Cloud LMS Software”. <https://www.talentlms.com/>. Accessed 6 Feb. 2021.
- [14] Thought Industries. “Powering the Business of Learning”. <https://www.thoughtindustries.com>. Accessed 15 Jan. 2021.
- [15] Complete List of Learning Management System Reviews, <https://www.lms.org/learning-management-systems-list/>. Accessed 23 Sep. 2021
- [16] Baker, R., Evans, B., Li, Q. et al. (2019) “Does Inducing Students to Schedule Lecture Watching in Online Classes Improve Their Academic Performance? An Experimental Analysis of a Time Management Intervention” *Res High Educ* 60, 521–552 <https://doi.org/10.1007/s11162-018-9521-3>
- [17] E. Péter and H. Ferenc (2017) "Analysis of video views in online courses" *40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia*, pp. 778-782, doi: 10.23919/MIPRO.2017.7973527.
- [18] Guo, P. J., Kim, J., & Rubin, R.(2014) “How video production affects student engagement: an empirical study of MOOC videos” *AMC Conference on Learning @ Scale*. Atlanta, GA.
- [19] Brame CJ. (2016) “Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content” *CBE Life Sci Educ.*;15(4):es6. doi:10.1187/cbe.16-03-0125
- [20] Birgili B, Seggie FN, Oğuz E, (2021) “The trends and outcomes of flipped learning research between 2012 and 2018: A descriptive content analysis” *Journal of Computers in Education*.1-30. doi:10.1007/s40692-021-00183-y
- [21] Suppan M, Stuby L, Carrera E, et al. (2021) “Asynchronous Distance Learning of the National Institutes of Health Stroke Scale During the COVID-19 Pandemic (E-Learning vs Video): Randomized Controlled Trial”. *J Med Internet Res*. 23(1):e23594. Published 2021 Jan 15. doi:10.2196/23594
- [22] Choe RC, Scurig Z, Eshkol E, et al.(2019) “Student Satisfaction and Learning Outcomes in Asynchronous Online Lecture Videos” *CBE Life Sci Educ*. 18(4):ar55. doi:10.1187/cbe.18-08-0171
- [23] Cooper KM, Ding L, Stephens MD, Chi MTH, Brownell SE.(2018) “A Course-Embedded Comparison of Instructor-Generated Videos of Either an Instructor Alone or an Instructor and a Student”. *CBE Life Sci Educ*.17(2):ar31. doi:10.1187/cbe.17-12-0288
- [24] Kaltura, Powering Any Video Experience, <https://corp.kaltura.com/>. Accessed 2 Mar. 2021.

AUTHORS

Dr. Nguyen is an associate professor in the department of Computer Science and Information Technology at Clayton State University, Georgia USA. Dr. Nguyen has published numerous research articles in the field of Design and Human-Computer interaction. Dr. Nguyen's areas of expertise are in Algorithms, Hi-performance, and Large-scale Computing, Artificial Intelligent, and Full-Stack Software Development. Dr. Nguyen was a software developer prior to joining the academia. His current interests are in the fields of Bioinformatics, Cloud and Mobile Computing, Robotics, and Machine Learning. He published a book on Large-scale Computing in Bioinformatics titled "Biological Multiple Sequence Alignment: Scoring Functions, Algorithms, and Evaluations", John Wiley & Sons, Inc (2016) ISBN-13: 978-1118229040.



Dr. Rahman is a professor in the department of Computer Science and Information Technology at Clayton State University, Georgia USA. Dr. Rahman earned his Ph.D. in computer science from Illinois Tech in the year 2000. Prior to academic appointments he worked as a software developer for the securities industry. Dr. Rahman has been teaching for over twenty years and many of these years he has been trained and an expert in online and remote course development and delivery. His research interests include natural language processing and computer science education. Dr. Rahman has been active in research and publication. He published numerous articles in computational linguistics, user interface design, online learning, and computer science education.



EVENT-DRIVEN TIME SERIES ANALYSIS AND THE COMPARISON OF PUBLIC REACTIONS ON COVID-19

Md. Khayrul Bashar

Tokyo Foundation for Policy Research, Tokyo 106-6234, Japan

ABSTRACT

The rapid spread of COVID-19 has already affected human lives throughout the globe. Governments of different countries have taken various measures, but how they affected people lives is not clear. In this study, a rule-based and a machine-learning based models are applied to answer the above question using public tweets from Japan, USA, UK, and Australia. Two polarity timeseries (meanPol and pnRatio) and two events, namely “lockdown or emergency (LED)” and “the economic support package (ESP)”, are considered in this study. Statistical testing on the sub-series around LED and ESP events showed their positive impacts to the people of (UK and Australia) and (USA and UK), respectively unlike Japanese people that showed opposite effects. Manual validation with the relevant tweets shows an agreement with the statistical results. A case study with Japanese tweets using supervised logistic regression classifies tweets into health-worry, economy-worry and other classes with 83.11% accuracy. Predicted tweets around events re-confirm the statistical outcomes.

KEYWORDS

COVID-19, lockdown, economic support, public reactions, polarity timeseries, statistical analysis, machine learning, sentiment comparison.

1. INTRODUCTION

Now a days, the microblogging platforms especially the Twitter has become essential tools for communication, especially for political or professional leaders including health professionals and public to interact with each-other as well as their intra-domain communication [1]. It has become a popular platform starting from the third world countries to the developed countries. This platform is playing active roles to disseminate public health information and to obtain real-time health data using crowdsourcing methods. It has already been used to disseminate information during many public health disasters such as influenza in 2009, the outbreak of Ebola virus (EV) in 2014, the spread of Middle Eastern respiratory syndrome in 2015 and the outbreak of Zika virus in the late 2015 [2]. World Health Organization (WHO) and the Centers for Disease Control and Prevention (CDC) have adopted the use of twitter and other social media by realizing the twitter’s potential to inform and educate the public and governmental agencies. Several systematic review papers identified six main uses of Twitter for public health: (i) analysis of shared content, (ii) surveillance of public health topics or diseases, (iii) public engagement, (iv) recruitment of research participants, (v) public health interventions, and (vi) the network analysis of Twitter users [3]. On the other hand, the Twitter platform has been facilitating the analysis of many political issues such the prediction of vote percentage, the political campaign and its effects, the analysis of political homophily, the detection of election fraught etc. [4]. Some researchers are using corona related tweets for political framing [5].

In December 2019, the first diagnosis of a novel coronavirus, formally named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was made in the Wuhan City, Hubei Province, China. Later, its rapid spread has drawn increasing media and public attention. Press coverages have further elevated in January 21, 2020 when the CDC activated its emergency operations center and the WHO began publishing daily situation reports. Subsequent travel bans, large-scale quarantine of Chinese residents generated significant interests by the public. Local and global media have gradually become more active to update the ongoing corona situations and to publish in-depth analysis on corona pandemic. Governments of different countries have gradually taken various measures including the declaration of emergency or lockdown and the financial support packages for flattening the peak situations of corona virus. At the same time, the public health researchers have taken many emergency projects to find the source of corona virus and to discover vaccines for controlling its spread [6]. However, there is limited insight on how the public sentiments are affected by the corona severity and/or the various government actions such as the declaration of lockdown or emergency conditions and the economic support packages.

2. RELATED WORKS

Over the past months after the onset of coronavirus, several works were published. Three main streams of works are progressing: (i) the development of models for estimating the spread of corona virus and the associated infection or death cases (ii) the development of vaccine as a remedy against this deadly disease, and (iii) the analysis of epidemic's impacts on the public health, economy and the global supply systems. Several authors have proposed advanced predictive models based on genetic programming and advanced machine learning algorithms including deep learning (e.g., LSTM) to address issues in the first stream [7] [8]. These models help to interpret patterns of the public sentiments in disseminating the related health information and assess political and economic influences of the spread of the virus. Researchers in the biological and medical domains are handling the second stream and are actively working on the vaccine development [6]. At present, several vaccine candidates are in the market places and several others are in the trial stage. Several researches were published in the third stream in which lexicon-based approaches, machine learning, and topic models are used [3] [9] [10]. These studies revealed the economic and political impacts of the COVID-19 as the most commonly discussed topics, while the risk for the public health and its prevention were the least discussed topics. Another aspect in this stream is the analysis of retweet networks and retweet speed. Beside the scientific researches, the governments in different countries have been taking various preventive measures such as the declaration of lockdown or emergency conditions, the financial support packages for the individual and business supports etc. However, it is quite unclear whether these measures really affected the public lives and sentiments. Moreover, how the public worries towards the health and economy has changed or is changing over time is also not fully explored. In this study, we therefore focus on exploring these issues. For the first one, we will consider two events namely the first declaration of lockdown or emergency (LED) and the economic support package (ESP). In the second case, a supervised machine learning method will be developed to classify public reactions into health and economy worries and their progression over time. Such understanding would enable the largescale opportunities for the prediction of public worries towards the health and economy during future catastrophic events.

3. DATA COLLECTION

3.1. Dataset-1

To test our research questions on the corona pandemic, we collected public tweets of the four countries (Japan, USA, UK, and Australia) for a six-month duration (January 2020 to June 2020). Target countries were selected considering geographical locations and the nature of infection varieties due to corona virus. Tweets were collected using “keyword” based search strategy [1] [11] [12]. Since we are interested in the public reactions in relation to the respective government actions, we used leader’s twitter handle as one of the search terms. For example, to extract public tweets for Japan, we constructed the query-string by concatenating corona related terms with the twitter handle of the corona in-charge in Japan (i.e., Economy Minister Yasutoshi Nishimura; @nishy03) by using logical “AND” operation.

In the keyword-based searching, several keywords (or hashtag keywords) are used for downloading the required tweets. The selection of keywords and their numbers usually depends on the target of projects [3] [13] [14]. Our study aims at comparing public reactions to COVID-19 among three English-spoken and one non-English spoken (e.g., Japan) countries. In this study, we therefore selected seven commonly used keywords and/or hashtags in the selected countries: “corona”, “coronavirus”, “novel coronavirus”, “COVID-19”, “COVID19”, “virus”, and “covid”. During the selection of the above keywords, we were motivated by the Oxford English Dictionary (OED) ranking and several other sources, namely Instagram, Yale Medicine Team, and several online reports [15 - 21]. With the mentioned keywords, we finally extracted 28, 930 public tweets using twitter standard search API for the mentioned four countries. A threshold of 100 maximum tweets per day was set to make the API workable during data collection. The characteristics of our datasets are given in given in Table 1. The number of the collected tweets is approximately proportional to the number of active twitter users (on the corona issue) in each country. Although, the dataset is not very large one, the random selection of the tweets by the API greatly reduces the possible biases on the collected datasets. These datasets will be used for statistical analysis of events. Besides, the dataset for the Japanese public will also be used for the classification of the public reactions in Japan.

Table 1: Information on public tweet datasets

| Group | Total days | Min | Max | Avg | Tweets |
|-----------|------------|-----|-----|-------|--------|
| Japan | 145 | 1 | 100 | 28.08 | 4072 |
| USA | 173 | 1 | 100 | 88.09 | 15240 |
| UK | 163 | 1 | 100 | 63.09 | 8090 |
| Australia | 164 | 1 | 81 | 16.93 | 1528 |

3.2. Dataset-2

It is a small dataset, constructed for Japan based on a different set of keywords than those used for collecting Japanese tweets in Dataset-1. These keywords were selected to directly construct annotated datasets for supervised classification of two dominant classes, i.e., health-worry

“hWorry”), and economy-worry (“eWorry”) as found in the Dataset-1. The translated version of some Japanese keywords for “hWorry” and “eWorry” classes include mask shortage, medical system collapse, lack of testing, bankruptcy, corona recession, damage on tourism etc. To keep consistent with the tweet categories in the Dataset-1, we have created a third class, designated as “other”, by including freely available USA airline review tweets [22]. Therefore, this dataset consists of 146, 349, 300 samples for “hWorry”, “eWorry”, and “other” classes, respectively. Finally, the Dataset-2 is used for training and validation, while the Japanese tweets in Dataset-1 is used as the final testing sets in our study.

4. EVENT-DRIVEN SENTIMENT ANALYSIS

4.1. Overview

In this study, timeseries tweet data is first preprocessed to eliminate URL, punctuations, stop words and rare words. Lemmatization is also performed to remove inflectional endings only and to return the base or dictionary form of a words. Then sentiment parameters are extracted and the date-wise sentiment timeseries are constructed for mean polarity (“meanPol”) and positive-negative count ratio (“pnRatio”) using a rule-based sentiment extraction model, called VADER [23]. To explore whether government actions against corona pandemic has potential effects on the public sentiments, we consider two well-known events, namely the first declaration of lockdown or emergency conditions (LED), and the first declaration of economic support package (ESP) among four countries Japan (JAP), USA, UK, and Australia (AUS). Statistical error analysis and Welch’s *t*-testing has been performed fifteen days before and after each event for each country to justify the significance of the actions taken by each government. A validation study on the event-related tweets has been performed to verify the results of the statistical analysis. Finally, a case study using the Japanese tweets has been performed which employs logistic regression to classify and then to verify the tweets related to “hWorry”, “eWorry”, and the “other” categories, respectively. Results showed the promising performance of the proposed method and analysis. Figure 1 below show the overview of our approach.

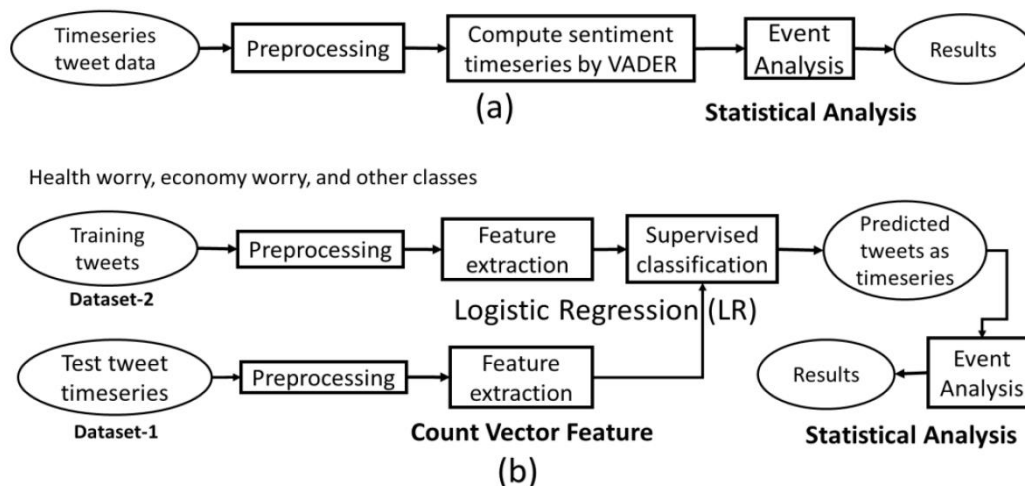


Fig. 1. (a)Even-driven Processing of twitter sentiment timeseries (b) Classification of public reactions to health worry, economic worry, and others.

4.2. Polarity Extraction

A rule-based model, entitled **Valence Aware Dictionary and Sentiment Reasoner (VADER)**, is adopted for public sentiment analysis from the twitter data. This model combines lexical features

with five general rules that embody grammatical and syntactical conventions for emphasizing sentiment intensity and showed effectiveness in social media data analysis. It tells not only about the positivity and negativity score but also tells us about how positive or negative a sentiment is. It outperforms individual human raters and captures better contexts compared to eleven benchmarks including ANEW, SentiWordNet, and machine learning oriented techniques relying on Naive Bayes, Maximum Entropy, and Support Vector Machine (SVM) algorithms. Please refer to [23] for detail information.

4.3. Sentiment Parameters

We have computed several sentiment parameters using VADER. These are mean polarity (“meanPol”) and the positive-negative polarity count ratio (“pnRatio”) in our analysis. In VADER, emotion intensity or sentiment score is measured on a scale from -4 to +4, where -4 is the most negative and +4 is the most positive. The midpoint 0 represents a neutral sentiment. VADER can map emoticons that appear in the social media texts like tweets. In our study, we only consider positive and negative scores of each sentiment bearing word. The above parameters were computed based on compound polarity score as defined by Eq. 1.

$$\text{Compound Score} = \frac{x}{\sqrt{x^2 + \alpha}} \quad (1)$$

where x is the sum of the sentiment scores of the constituent words of the sentence or sentences and α is a normalization parameter that we set to a default value 15. This gives us a normalized score between -1 (most extreme negative) and +1 (most extreme positive). Please refer to (Hutto et al., 2014) for more information.

4.4. Classification of Public Reactions

Understanding the health and economic related public reactions is very important in decision making by the leaders and government of a country. This could help proactive policy decisions especially during disastrous situation like corona pandemic. In this study, we have developed a classification algorithm using logistic regression to classify public reactions as “hWorry” and “eWorry” using Japanese tweets as a case study.

After preprocessing, the count-vectorizer are applied to construct numerical feature matrix for classification [24]. Classical machine learning models usually perform better than the advanced deep learning models with relatively small datasets having short-length text data. After an investigation with Dataset-2 using four machine learning models namely the Naïve Bayes (NB), linear support vector machine (LSVM), logistic regression (LR), and Random Forest (RF), we have finally selected the count-vector feature and the LR model for our analysis. The collection of Japanese tweets for six-month duration (Please refer to Section 3.1) was used as the final testing set. Japanese tweets are translated into English using “Googletrans”, a python library that implemented Google Translate API (“Googletrans”, a python library). Classification performance is computed using well-known evaluation metrics as defined below [25].

$$\text{PREC} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{SN} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{SP} = \frac{TN}{TN + FP} \quad (4)$$

$$ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)}, \quad (5)$$

where TP, FN, FP and TN represent the number of true positives, false negatives, false positives and true negatives, respectively.

5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1. Polarity Timeseries Analysis

With the collected data, we have generated “meanPol” and “pnRatio” timeseries for four countries (Figs. 2 and 3). Figure 2 shows that the Japan, UK, and Australia have large variations in the mean polarity before March 15, 2020, while the mean polarity for USA varies wildly throughout the six-month duration. The mean polarity for the Australia also varies widely after the 2nd week of April 2020. The most negative polarity was observed for UK and AUS public, while the most positive polarity was found for Japanese people. An approximately similar trend was observed in the pnRatio timeseries as shown in Fig. 3.

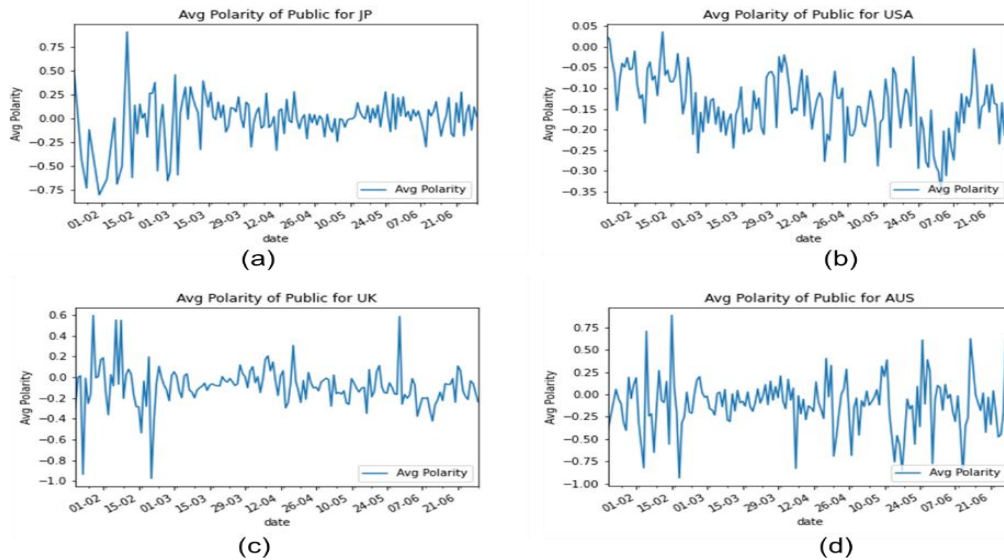


Fig.2. Plots for average polarity for (a) Japan, (b) USA, (c) UK, and (d) Australia

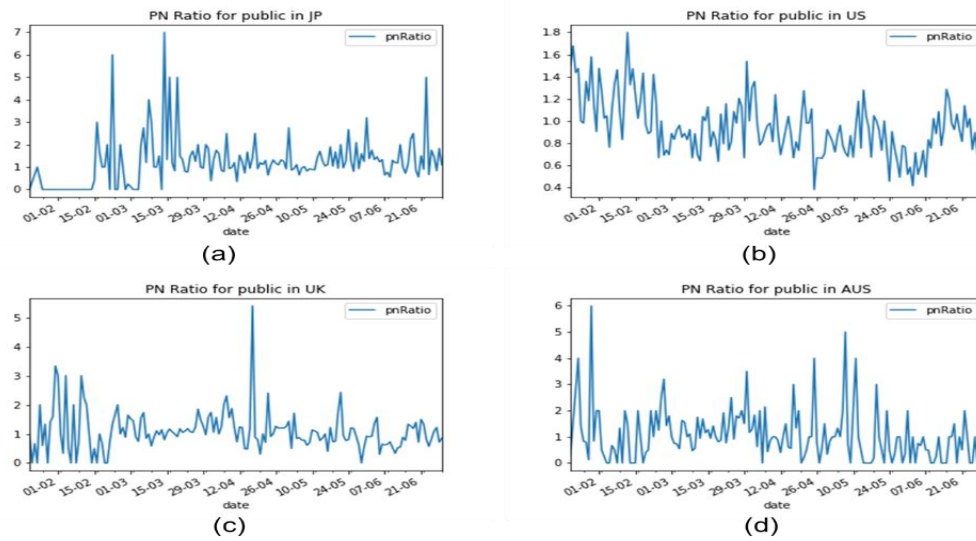


Fig.3. Plots for pnRatio for (a) Japan, (b) USA, (c) UK, and (d) Australia

5.2. Event Analysis

To address the research question, we have considered two events: (i) the first declaration of lockdown or emergency (LED) and (ii) the first declaration of economic support package (ESP).

LED: The starting date at which some restrictions on the human and/or business activities at public places were first imposed for the sake of ensuring public and national safety in a special occasion (e.g., violence, pandemic etc.). In case of the corona pandemic, such event is known as lockdown or the state of nation-wide emergency declaration. In this study, we considered this date as an important event.

ESP: On the eve of disaster situations, the national government usually declares financial support to protect its citizens from financial and mental crisis. During COVID-19, most of the affected countries declared such packages at different point of time. In our study, the first occurrence of this kind of support is considered as an event for analysis.

Table 2. MeanPol and pnRatio before and after LED

| | MeanPol | | pnRatio | |
|-----------|-----------|----------|-----------|----------|
| | BeforeLED | AfterLED | BeforeLED | AfterLED |
| Japan | 0.039362 | 0.003026 | 1.430606 | 1.282577 |
| USA | -0.16379 | -0.14241 | 0.876282 | 0.925935 |
| UK | -0.06723 | 0.013028 | 1.045158 | 1.429418 |
| Australia | -0.10858 | -0.01557 | 1.073405 | 1.536419 |

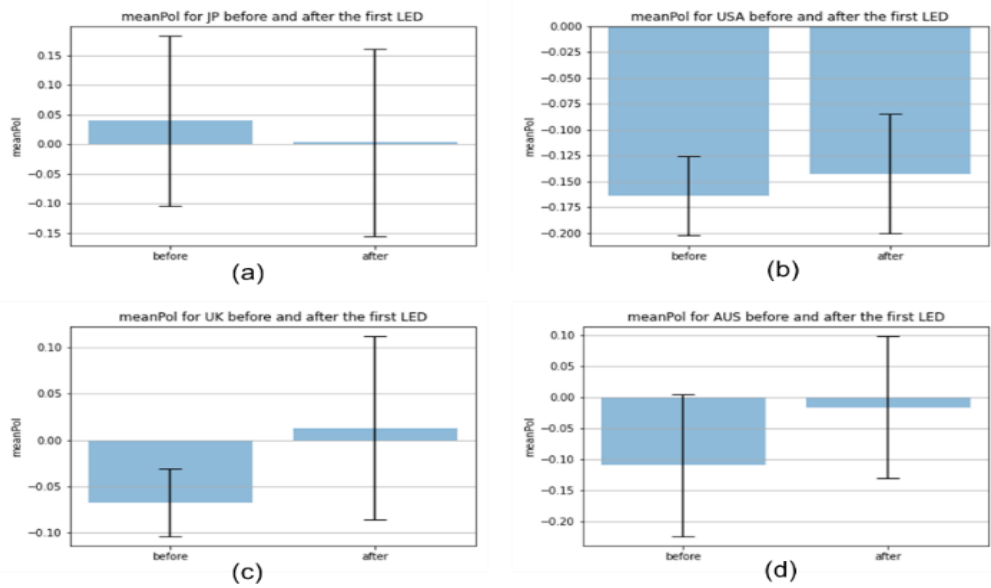


Fig.4. Error plots of meanPol for LED. (a) Japan, (b) USA, (c) UK, and (d) Australia

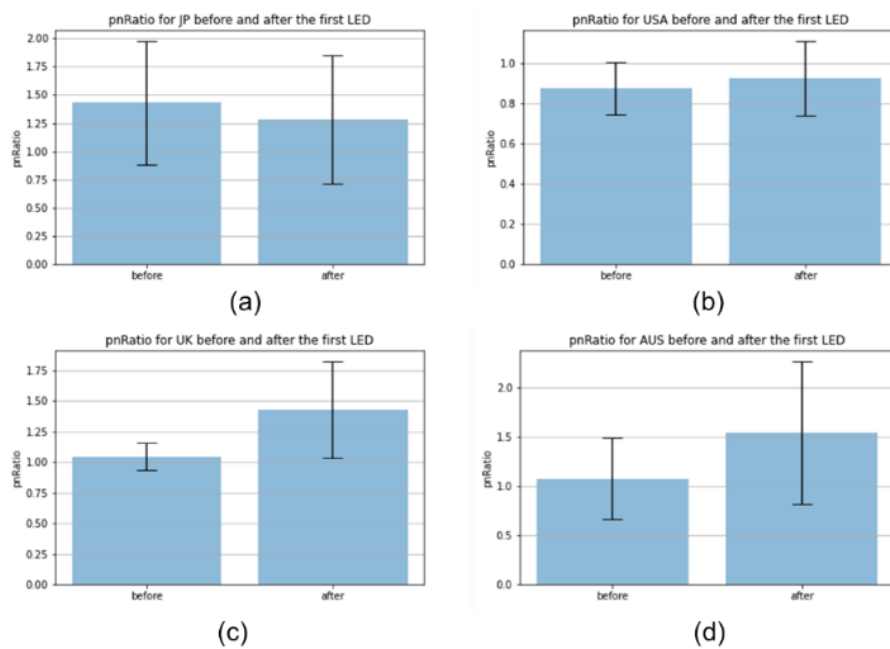


Fig.5. Error plots of *pnRatio* for LED. (a) Japan, (b) USA, (c) UK, and (d) Australia

Table 2 and Figs. 4 and 5 show the meanPol and pnRatio parameters 15 days before and after the LED event. These figures clearly show that except for Japan, the negative meanPol reduces after the LED event for Australia, UK, and USA. On the other hand, pnRatio increases after the LED event for the same three countries (Australia, UK, USA) except Japan. This indicates that the public in those three countries somewhat accepted the LED as a positive step. While in Japan, it showed the opposite behavior indicating that the Japanese people are not much satisfied with the LED implementation.

Table 3 and Figs. 6 and 7 show the meanPol and pnRatio parameters 15 days before and after the ESP event. Figures show that except for Japan and Australia, the negative meanPol reduces after the ESP event for UK, and USA. On the other hand, pnRatio increases after the ESP event for the same two countries (UK, USA) except the Japan and Australia, which showed the opposite behavior. This indicates that the public in the USA and UK accepted the ESP as a positive step to some extent. While the public in Japan and Australia are not much satisfied with the ESP implementation.

Table 3. MeanPol and pnRatio before and after ESP events

| | MeanPol | | pnRatio | |
|-----------|-----------|----------|-----------|----------|
| | BeforeESP | AfterESP | BeforeESP | AfterESP |
| Japan | 0.039362 | 0.003026 | 1.430606 | 1.282577 |
| USA | -0.16379 | -0.09534 | 0.876282 | 1.044415 |
| UK | -0.06723 | -0.00656 | 1.045158 | 1.292268 |
| Australia | -0.10858 | -0.04912 | 1.073405 | 1.292818 |

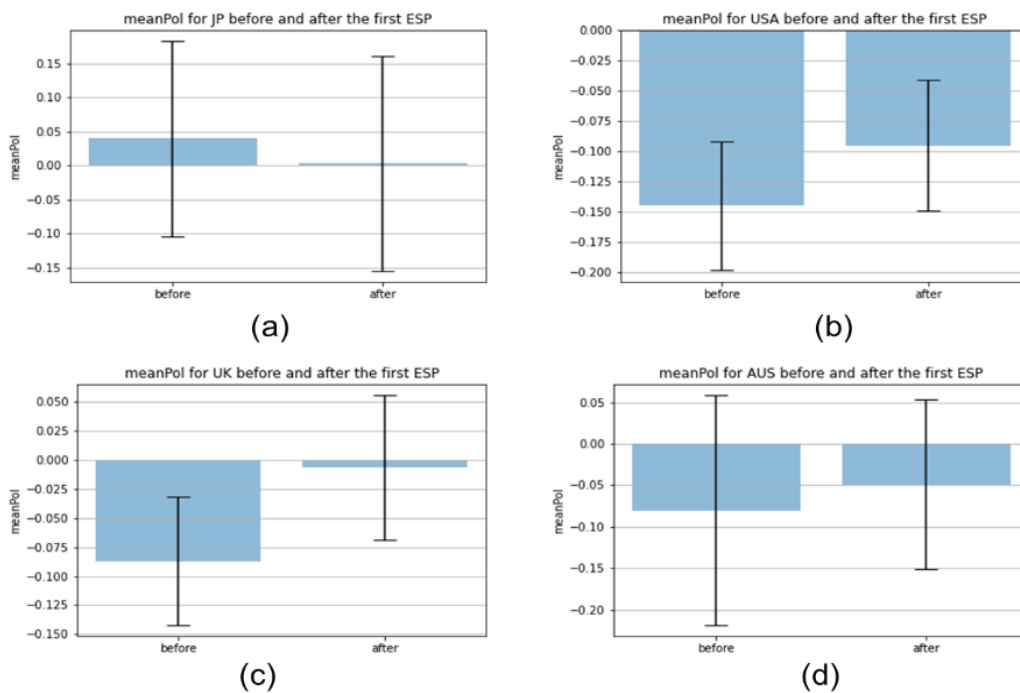


Fig.6. Error plots of MeanPol for ESP. (a) Japan, (b) USA, (c) UK, and (d) Australia

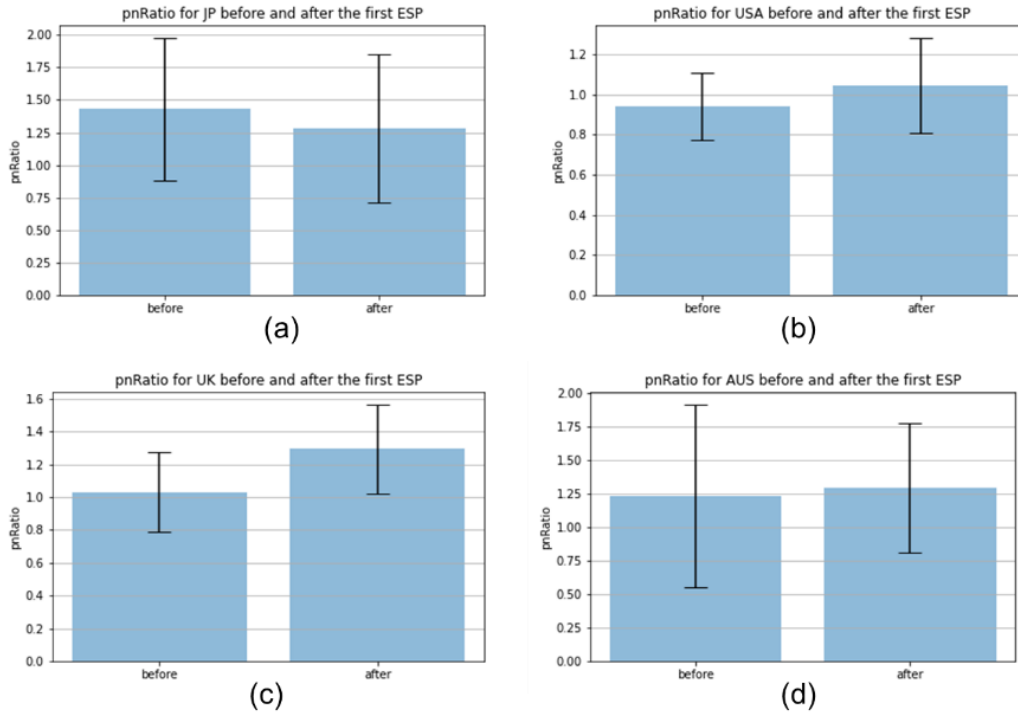


Fig.7. Error plots of *pnRatio* for ESP. (a) Japan, (b) USA, (c) UK, and (d) Australia

To justify the above observations more closely, we have performed Welch's *t*-testing on the 15-days sentiments before and after the mentioned two events. Table 4 shows the test results. Results showed significant differences of the MeanPol and *pnRatio* parameters over the LED event for UK and Australia, while they showed significant differences for USA and UK on the FSP event. These testing outcomes are quite congruent with the results from the statistical error analysis (Figs. 6 and 7) except for the USA on the LED event. However, the Japanese tweets does not show any significant differences for both events. In conclusion, we can say that UK and USA people are somewhat satisfied with the government's ESP event, while the UK and Australian people are satisfied with the LED event.

Table 4. Welch's *t*-test Results for Event Sentiments Analysis

| Country | Parameters | LED | | ESP | |
|-----------|------------|----------------|----------------|----------------|----------------|
| | | <i>MeanPol</i> | <i>pnRatio</i> | <i>MeanPol</i> | <i>pnRatio</i> |
| Japan | t | 0.6399 | 0.7023 | 0.6399 | 0.7023 |
| | p | 0.5274 | 0.4882 | 0.5274 | 0.4882 |
| | dof | 27.744 | 27.948 | 27.744 | 27.948 |
| USA | t | -1.155 | -0.8212 | -3.860 | -2.334 |
| | p | 0.2590 | 0.4192 | 0.0007 | 0.0291 |
| | dof | 24.262 | 25.188 | 25.116 | 21.845 |
| UK | t | -2.845 | -3.503 | -3.151 | -3.160 |
| | p | 0.0108 | 0.0029 | 0.0045 | 0.0052 |
| | dof | 17.748 | 16.173 | 22.669 | 18.529 |
| Australia | t | -2.152 | -2.074 | -1.448 | -1.292 |
| | p | 0.0401 | 0.0497 | 0.1587 | 0.2070 |
| | dof | 27.999 | 22.296 | 27.656 | 27.416 |

To justify the above observations, we have manually investigated the top 50 most negative tweets before and after each event. Negative tweets are selected based on the observations that the overall MeanPol is mostly negative for all countries except Japan which is slightly positive (Please refer to the Tables 2 and 3). Figures 8 and 9 show the results of investigation before and after the LED and ESP event, respectively.

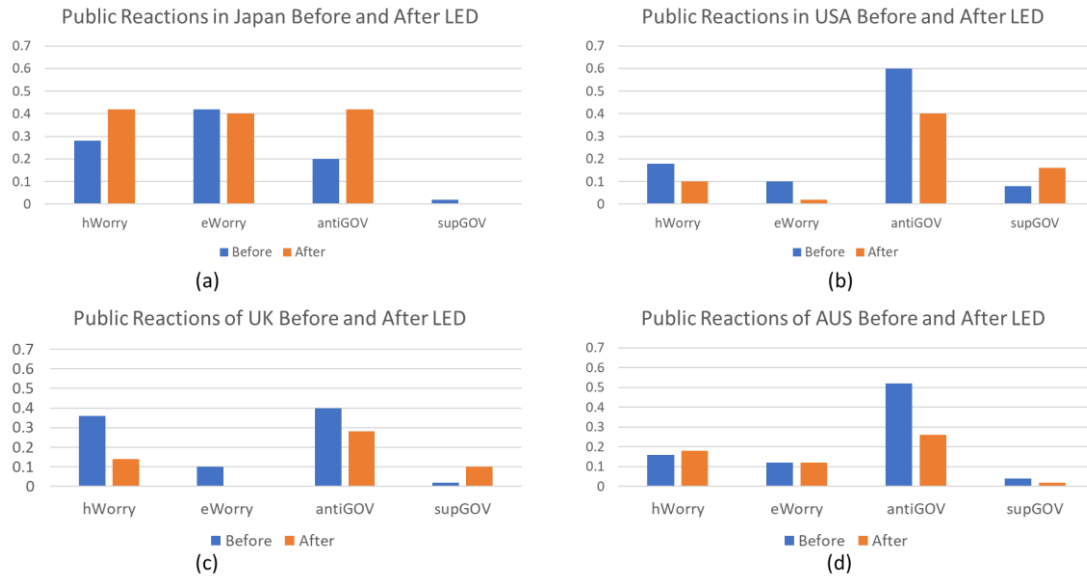


Fig.8. Number of tweets for LED. (a) Japan, (b) USA, (c) UK, and (d) Australia

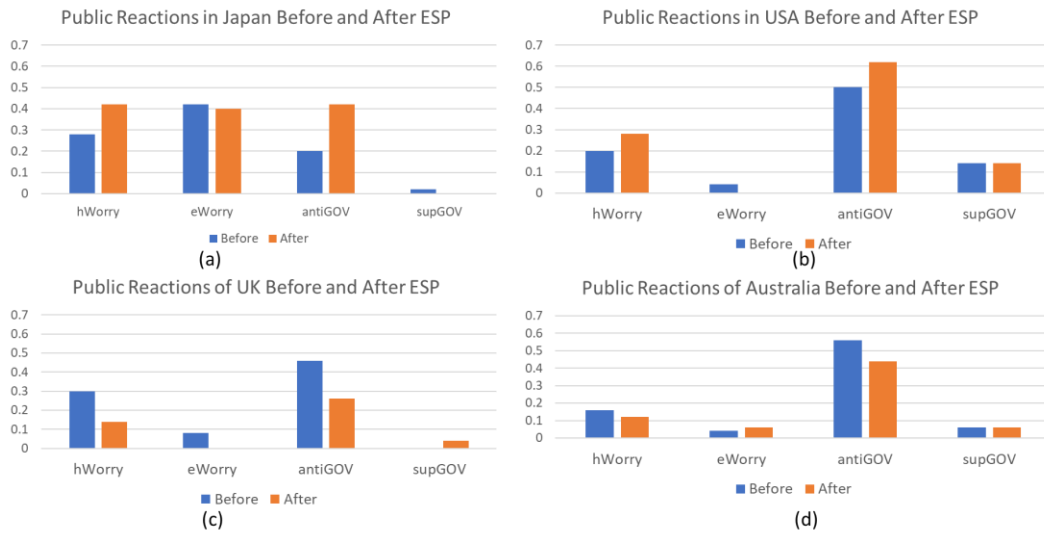


Fig.9. Number of tweets for ESP. (a) Japan, (b) USA, (c) UK, and (d) Australia

The depicted results in the above figures agree with the results of statistical analysis as explained above. The Fig. 8 (b-d) for the Australia, UK, and USA show similar tendency with decreasing worries for health and economy in case of the LED event, while Japan shows the opposite behavior. Note that Australia has a small perturbation for the health worry case. This is perhaps a side effect for not including all negative tweets in manual verification. Similar findings were observed in case of the ESP event (Fig. 9). In this case, the USA is not in agreement with the results of statistical analysis. Again, this perturbation may happen because of the mentioned

reason as above as well as the limitation due to human annotation. However, the overall outcomes are approximately in agreement with the results that we obtain from statistical analysis. In addition, we can go more insight with the validation results in terms of anti-government public sentiments. The anti-government sentiments of the Japanese people have increased after each event, while the same for the most other countries have decreased except the USA with the ESP event. Results also showed the highest anti-government sentiments in USA and Australia compared to other countries. Note that both the LED and ESP events were declared in the same date in Japan. Note also that the Japanese people gave stronger reactions on the health and economic issues compared to other countries (Fig. 8 and 9 (a)).

5.3. Classification of Public Reactions

Based on the above observations, we have decided to perform a classification study on the reactions of Japanese public as a case study. In this case, supervised classification method has been developed as explained in the section 4.4. **Table 5** below shows the results of classification using count vectorizer with LR classification model. With the three classes (“hWorry”, “eWorry”, and “other”), we have obtained 83.11% average classification accuracy (macro average) with the high percentage of specificity (87.33%) and reasonable precision (74.78%) and sensitivity (74.66%). Note that the classification accuracy for the “hWorry” class is relatively low compared to the “eWorry” class, indicating the effects of our unbalanced training set. However, we hope to improve our results by adopting proportional weighting scheme or advanced deep learning model with larger dataset and more appropriate numerical feature model. Results also indicates that we can successfully classify high-level concepts like the worries for health and economy using the properly designed training sets.

Table 5: Classification Performance on the test set using LR with count-vector feature

| | PREC | SN | SP | ACC |
|---------------|--------|--------|--------|--------|
| hWorry | 0.767 | 0.690 | 0.895 | 0.827 |
| eWorry | 0.735 | 0.860 | 0.845 | 0.850 |
| other | 0.742 | 0.690 | 0.880 | 0.817 |
| Macro average | 0.7478 | 0.7466 | 0.8733 | 0.8311 |

*Test set: 4072 tweets, verified 100 tweets from each class

Figure 10 shows per day classified tweets as the line plot. These plots showed clear dominance of economy worry of the Japanese people compared to their health worry especially between the beginning of April 2020 to the first week of May 2020. This was, in fact, the peak devastation period at which the Japanese government declared the emergency condition and the financial support package for the first time. Since the test-set has more than 4000 tweets, it is very time consuming to label all tweets manually. We therefore manually annotated the first 300 tweets (100 tweets per class) using the resultant tweet-groups after classification.

To justify the classification results against the “LED” and “ESP” events, the predicted tweets 15 days before and after each event were inspected manually. Results were shown in the Table 6 which shows that the counting and percentage of the “hWorry” tweets has increased after the LED and ESP events, while the “eWorry” tweets has slightly decreased after the declaration of each event. These results are consistent with the statistical analysis, shown in Figs. 8 (a) and 9(a). This observation once again validates that the Japanese people were not very happy with the government corona measures. Many inspected tweets also revealed that the people were unhappy with the delay in implementing the necessary health and economic measures of the government as well as its bureaucratic attitude.

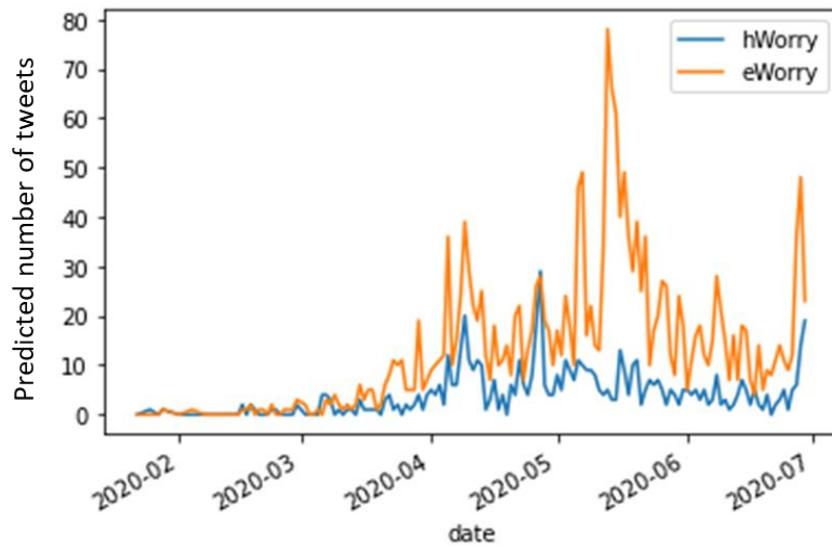


Fig.10. Distribution of the predicted hWorry and eWorry tweets over time

Table 6: Public reactions on the test set using LR with count-vector feature

| No. of tweets | Before LED or ESP | | | After LED or ESP | | |
|---------------|-------------------|--------|-------|------------------|--------|-------|
| | hWorry | eWorry | Other | hWorry | eWorry | Other |
| Tweets | 87 | 43 | 217 | 165 | 42 | 358 |
| % Tweets | 0.25 | 0.123 | 0.625 | 0.292 | 0.074 | 0.633 |

*Events (LED and ESP) dates for Japan were same

5.4. Discussion

We have already discussed the strength and limitations of our method throughout the section 5.

- (1) The subseries length, i.e., 15 days before and after each event (LED, ESP) is currently selected by trial and error method. An automated method for such selection will be developed in our future studies.
- (2) Although our current results were based on a dataset having approximately 30,000 tweets, we believe in the principle that the more the tweets the better will be the analysis results. Therefore, our current dataset will be augmented with more samples in our future study.
- (3) Another point is that the training set we currently used for the classification study are somewhat small and specific for the Japanese tweets. More generalized training set or dictionary and the more sophisticated learning algorithm will be developed in the future studies so that the proposed method can be scaled for more countries.
- (4) In addition, we plan to extend our work to the prediction framework which can predict public sentiments during the similar catastrophic situations in future.
- (5) The expected realistic comparison of the event effects among multiple countries is troublesome because public reactions in a country depends on its social, economic, political, and cultural conditions, which are usually different in different countries. However, the degree of worries or sensitivity of fundamental issues like economy, health, and politics are well understood among multiple countries considered in this study.

6. CONCLUSIONS

We have introduced a method for the event-driven analysis of the public sentiments from corona tweets. This method incorporates a rule-based lexicon for the construction of sentiment timeseries and a machine learning model for the classification of Japanese public sentiments into the “health-worry” and “economy-worry” classes. To answer how the government responses to the corona pandemic affected the public sentiments, we considered two polarity timeseries (i.e., meanPol and pnRatio) and two events namely LED and ESP. Statistical error analysis on the meanPol and pnRatio sub-series showed a decrease of the negative sentiment after the LED and ESP events for all countries except Japan. Welch’s t-test on the above sub-series also showed significant positive impacts on the UK and Australian people by the LED and the USA and UK people by the ESP event, respectively. Manual validation with the relevant tweets approximately showed an agreement with the statistical analysis indicating that the people in different countries have differently affected by the government responses based on their socio-economic and political situations. The proposed logistic regression-based approach classified Japanese public tweets into “health-worry”, “economic-worry” and the “other” classes with 83.11% accuracy (on average). Results showed higher number of tweets on the economy worry, when compared with those with the health worry. An analysis with the classified tweets around each event also re-confirmed the results made by the statistical analysis. Note that the training set (Dataset-2) used for the classification study was extracted independently from the test set (Dataset-1) using health and economy related Japanese keywords. However, more generalized training sets can be developed, which will extend the classification task for multiple countries. Advanced deep learning-based algorithm can also be developed in the future study to perform fine-grained sentiment analysis. A natural extension of our method is to develop algorithm for the prediction of public worry for health and economy during similar disastrous situations in future.

ACKNOWLEDGEMENTS

We would like to thank all members of the policy data lab of TKFD, who had extended their cordial cooperation during the work in progress.

REFERENCES

- [1] Barbera, P. et al., (2019) “Who Leads? Who Follows? Measuring Issue Attention and Agenda Setting by Legislators and the Mass Public Using Social Media Data”, *APSR*, Vol. 113, No. 4, pp. 883-901.
- [2] Medford, R. J. et al., (2020) “An ‘Infodemic’: Leveraging High-Volume Twitter Data to Understand Public Sentiment for the COVID-19 Outbreak”, *Open Forum Infectious Diseases*, Vol. 7, No. 7, pp. 1-9.
- [3] Abd-alrazak, A. et al., (2020) “Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study”, *J Med Internet Res*, Vol. 22, No. 4, pp. 1-16.
- [4] Ansari, M.Z., Aziz M. B. et al., (2020) “Analysis of Political Sentiment Orientations on Twitter”, *Procedia Computer Science*, Vol.167, pp.1821–1828.
- [5] Shurafa, C. et al., (2020) “Political Framing: US COVID19 Blame Game”, In *Proc. SocInfo 2020*, pp. 1-15, virtual (<https://kdd.isti.cnr.it/socinfo2020/>).
- [6] Zhou, P. et al., (2020) “Research progress and challenges to coronavirus vaccine development”, *Journal of medical virology*, pp. 1-14, doi: 10.1002/jmv.26517.
- [7] Salgotra, R. et al., (2020) “Time Series Analysis and Forecast of the COVID-19 Pandemic in India using Genetic Programming”, *Chaos, Solitons, and Fractals*, Vol. 138, pp. 1-15.
- [8] Binti Hamzah, FA. Lau, C. et al., (2020) “CoronaTracker: Worldwide COVID-19 Outbreak Data Analysis and Prediction. [Preprint]”, *Bull World Health Organ*. E-pub (2020). doi: <http://dx.doi.org/10.2471/BLT.20.255695>.

- [9] Chaudry, R. et al., (2020) “A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes”, *EclinicalMedicine*, Vol. 25, pp. 1-15.
- [10] Samuel, J. et al., (2020) “COVID-19 Public Sentiment Insights and Machine Learning for Tweets Classification”, Submitted to *Journal Not Specified*, 1 – 23.
- [11] Josemar, Caetano, A. et al., (2018) “Using sentiment analysis to define twitter political users’ classes and their homophily during the 2016 American presidential election”, “*Journal of Internet Services and Application*”, pp.1-15.
- [12] Hatchard, J. L., (2019) “Tweeting about public health policy: Social media response to the UK Government’s announcement of a Parliamentary vote on draft standardized packaging regulations”, *PLOS ONE*, pp.1-12.
- [13] Wang, Y. et al., (2015) “Should we use the sample? Analyzing datasets sampled from twitter stream API”, *ACM Trans.* pp. 1-15.
- [14] Beauchamp, N. et al.: Predicting and interpolating state-level polls using twitter textual data. *American Journal for Political Science* 61(2), (2016). DOI: 10.1111/ajps.12274
- [15] Samuel, J. et al.: COVID-19 Public Sentiment Insights and Machine learning for Tweets Classification. www.mdpi.com/journal/notspecified
- [16] Oxford Corpus: <https://public.oed.com/blog/corpus-analysis-of-the-language-of-covid-19/>
- [17] Yale University Medicine Experts: <https://www.yalemedicine.org/stories/covid-19-glossary>
- [18] Instagram suggested keywords: <https://ingramer.com/instagram-hashtag/pandemic/>
- [19] Best-hashtags on corona: <http://best-hashtags.com/hashtag/coronavirus/>
- [20] Most-used-hashtags on corona: <https://www.thenational.ae/arts-culture/what-s-trending-during-coronavirus-pandemic-a-definitive-guide-to-the-most-used-hashtags-1.996208>
- [21] Seven trending hashtags: <https://www.mediaupdate.co.za/social/148423/seven-trending-hashtags-about-covid-19-on-social-media>
- [22] Airlines review tweets: <https://www.kaggle.com/crowdflower/twitter-airline-sentiment>, last accessed 2020/12/07
- [23] Hutto, C. J. et al., (2014) “VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text”, In *Proc. Eighth International AAAI Conference on Weblogs and Social Media (ICWSM 2014)*, pp. 1-10.
- [24] Chaudhary, M., “TF-IDF Vectorizer scikit-learn: <https://medium.com/@cmukesh8688/tf-idf-vectorizer-scikit-learn-dbc0244a911a>, last accessed 2020/12/19.
- [25] Sokolova, M., et al., (2009) “A systematic analysis of performance measures for classification tasks”, *Information Processing and Management*, Vol. 45. pp. 427–437.

AUTHORS

Md. Khayrul Bashar received bachelors in electrical and Electronic Engineering from Bangladesh University of Engineering and Technology (BUET), Master of Technology (M. Tech.) in Communication Engineering from Indian Institute of Technology Bombay, and PhD in Information Engineering from Nagoya University in 1993, 1998, and 2004, respectively. After his Ph. D, he served Nagoya University, the University of Tokyo, and Ochanomizu University as a Researcher and faculty member until March 2020. Currently, he is working as a data scientist with Tokyo Foundation for Policy Research. Dr. Bashar published about 20 research articles in the peer-reviewed journals in the fields of image data analysis and algorithms with the application of machine learning algorithms. His research interests include data analytics, applied machine learning, and social computing.

AUTHOR INDEX

| | |
|----------------------------------|----|
| <i>Blaise Hanczar</i> | 37 |
| <i>David A. Noever</i> | 53 |
| <i>Guillaume Ghyselinck</i> | 37 |
| <i>Ken D. Nguyen</i> | 77 |
| <i>Kenil Manishkumar Patel</i> | 61 |
| <i>Maha Aldosary</i> | 21 |
| <i>Md. Khayrul Bashar</i> | 87 |
| <i>Muhammad Asadur Rahman</i> | 77 |
| <i>Norah Alqahtani</i> | 21 |
| <i>Orcun Oruc</i> | 01 |
| <i>Samantha E. Miller Noever</i> | 53 |
| <i>Shahid Ali</i> | 61 |
| <i>Tien Tai Doan</i> | 37 |