# PROCEEDINGS OF THE ICTSZ - INTERNATIONAL CONFERENCE IN INFORMATION AND COMMUNICATIONS TECHNOLOGIES

## Cresta Golfview Hotel, Lusaka – Zambia

## 12TH - 13TH DECEMBER 2018


## Organized by The University of Zambia, Copperbelt University and Mulungushi University

## Key Sponsors

ISBN: 978-9982-70-787-9

# FORWARD

The objective of ICICT2018 conference was to support and stimulate active productive research which could strengthen the technical foundations of engineers and scientists in the continent, develop strong technical foundations and skills and lead to new small to medium enterprises within the African sub-continent. We also seek to encourage the emergence of functionally skilled technocrats within the continent. Tutorials and Sessions which will impact and enhance postgraduate research within the continent were considered. Training Workshops on research software tools such as Matlab, SPSS, Scilab, LINUX, Althium, Genesys, COMSOL were held during the conference. The conference also provides a forum for students to compete for best papers and receive an award. Registration fees for student authors are discounted. ICICT2018 made provision in 2018 for display and demonstration of working software and hardware that are relevant to the African manufacturing base. Demonstrations considered working prototypes from university students, staff, and industry. The Technical Program Tracks included the following.

1. Devices and Mobile Phone Applications and Software
2. Wireless Mobile Communications (3G, 4G, and 5G)
3. Sensors and Remote Sensing
4. ICT, e-Learning, and e-Education
5. ICT for Development
6. Smart Grid and Energy Harvesting
7. Telecommunications
8. Applied Mathematics
9. Computer Networks
10. Scientific Computing
11. Computational Intelligence and Forensics
12. Environmental and Sustainability
13. Radio Frequency and Microwave Engineering
14. Cloud Computing, Distributed Systems, and Virtualisations
15. ICT in Food Security
16. Data Science and Big Data
17. Bioinformatics and Computational Biology
18. Natural Sciences (Physics, Chemistry, Mathematics and Biological Sciences)
19. Engineering (Electrical and Electronics)
20. Library and Information Science
21. GIS and Geospatial Sciences
22. Business Information Systems and Information Management Systems

The Conference received a total of 42 papers which where peer reviewed through the EDAS system. Only 30 papers were accepted and included in the proceedings. This represent 71% acceptance rate. A selected papers have been extended and included in the ***Zambia ICT Journal***, Volume 3, Issue Number 1 (2019) - http://ictjournal.icict.org.zm/index.php/zictjournal/index

# TABLE OF CONTENTS

## Contents

# LOCAL ORGANISING COMMITTEE

| | |
|---|---|
| 1. Prof. Douglas Kunda | Mulungushi University (Chair) |
| 2. Prof. Jameson Mbale | Copperbelt University |
| 3. Dr. Jackson Phiri | The University of Zambia (Secretariat) |
| 4. Dr. Josephat Kalezhi | Copperbelt University (Secretariat) |
| 5. Mr. Bonny Khunga | ZAMREN (Member) |
| 6. Dr. Christopher Chembe | Mulungushi University (Member) |
| 7. Mrs. Monde Kalumbilo-Kabemba | The University of Zambia (Member) |
| 8. Dr. Charles Lubobya | University of Zambia |
| 9. Mr. Ariel Phiri | Zambia Airports Corporation Limited and IEEE Zambia (Member) |

# REVIEW COMMITTEE

1. Prof. Douglas Kunda — Mulungushi University, dkunda@mu.edu.zm
2. Prof. Jameson Mbale — Copperbelt University, jameson.mbale@cbu.ac.zm
3. Prof. Lisa Seymour — University of Cape Town, Lisa.Seymour@uct.ac.za
4. Prof. Hastings Libati — The Copperbelt University, libati@cbu.ac.zm
5. Dr. Josephat Kalezhi — Copperbelt University, kalezhi@cbu.ac.zm
6. Dr. Jackson Phiri — The University of Zambia, jackson.phiri@cs.unza.zm
7. Mr. Bonny Khunga — ZAMREN, khungab@zamren.zm
8. Mr. Ariel Phiri — Zambia Airports Corporation Limited and IEEE Zambia, ariel.phiri@zacl.aero, ahphiri@ieee.org
9. Dr. Christopher Chembe — Mulungushi University, cchembe@mu.edu.zm
10. Dr. Edmore Chikohora — Namibia University of Science and Technology, chikohora@nust.na
11. Dr. David Chisanga — The University of Zambia, david.chisanga@cs.unza.zm
12. Dr. Derrick Ntalasha — Copperbelt University, dbntalasha@gmail.com
13. Dr. Mayumbo Nyirenda — University of Zambia, mnyirenda@unza.zm
14. Dr. Vincent Omwenga — Strathmore University, vomwenga@strathmore.edu
15. Dr. Lighton Phiri — University of Zambia, lighton.phiri@unza.zm
16. Mr. David Zulu — University of Zambia, makadanizulu@gmail.com
17. Dr. Shemi Alice — Copperbelt University, shemiap@gmail.com
18. Mrs. Monde Kalumbilo-Kabemba — The University of Zambia, monica.kalumbilo@cs.unza.zm
19. Dr. Okuthe Kogeda — Tshwane University of Technology, kogeda@gmail.com
20. Dr. Manoj Lall — Tshwane University of Technology, LallM@tut.ac.za
21. Dr. Evans Lampi — University of Zambia, evanslampi@gmail.com
22. Dr. George Mufungulwa — The Copperbelt University, mufungulwac@gmail.com
23. Dr. Francis Mulolani — The Copperbelt University, fmulolani@cbu.ac.zm

# ABSTRACTS

## A Ransomware Classification Framework Based on File-Deletion and File-Encryption Attack Structures

Aaron Zimba (University of Science and Technology Beijing & Mulungushi University, P.R. China); Mumbi Chishimba and Sipiwe Chihana (Mulungushi University, Zambia)

Ransomware has emerged as an infamous malware that has not escaped a lot of myths and inaccuracies from media hype. Victims are not sure whether or not to pay a ransom demand without fully understanding the lurking consequences. In this paper, we present a ransomware classification framework based on file-deletion and file-encryption attack structures that provides a deeper comprehension of potential flaws and inadequacies exhibited in ransomware. We formulate a threat and attack model representative of a typical ransomware attack process from which we derive the ransomware categorization framework based on a proposed classification algorithm. The framework classifies the virulence of a ransomware attack to entail the overall effectiveness of potential ways of recovering the attacked data without paying the ransom demand as well as the technical prowess of the underlying attack structures. Results of the categorization, in increasing severity from CAT1 through to CAT5, show that a lot ransomware exhibit flaws in their implementation of encryption and deletion attack structures which make data recovery possible without paying the ransom. The most severe categories CAT4 and CAT5 are better mitigated by exploiting encryption essentials while CAT3 can be effectively mitigated via reverse engineering. CAT1 and CAT2 are not common and are easily mitigated without any decryption essentials. As such, the proposed framework avails the victim with a deeper understanding of the underlying ransomware attack structure. He can make an informed decision knowing well the consequences. Our objective is limited to providing a victim with full information after an attack has occurred. We do not seek to detect or prevent ransomware attacks. As such, our threat model assumes that the ransomware attack has already occurred which is reasonable enough considering the number of attacks thus far on the Internet.

## Closed-loop Current Control of a Grid-connected Five-level Inverter

Francis Mulolani (The Copperbelt University & Newcastle University, Zambia); Francis Kafata and Esau Zulu (The Copperbelt University, Zambia)

This paper presents a closed-loop current control scheme for a 5-level 3-phase diode-clamped multilevel inverter system. The proposed closed loop current control technique is based on the proportional integral (PI) controller theory, and the modulation technique used is level-shift-carrier sinusoidal pulse width modulation (SPWM). The gain values of PI controller were selected on the basis of trial and error to achieve a good compensating signal. Grid synchronization was achieved by using a phase-locked loop. Matlab/Simulink software has been used to run the simulations. The simulation results show that a 1.17% total harmonic distortion (THD) of the output current was attained.

## Addressing Energy Consumption Problem Using Wi-Fi at A Care Home for the Aged in Zambia

Justine Chilongu (The Copperbelt University, Zambia); Lusungu Ndovi and Josephat Kalezhi (Copperbelt University, Zambia)

The development of smart homes for the aged people has over the recent years come as a result of the development of electronics and Information Communication Technologies (ICTs) in which electronic devices have played a very important role of enabling both the monitoring and controlling of electrical devices. Ordinarily, the existing systems do not allow a user to get a feasibility to actively mitigate the power consumption of home equipment. In the past, the electric home equipment could not be easily controlled and monitored resulting into having huge energy consumption costs. However, at the moment, a wireless Sensor Network (WSN) technology is being used for the controlling and monitoring of electric home equipment far more than in the past. This paper describes the proposed

methodology of a Wi-Fi based home automation system in which two things are considered. The first one is energy consumption and the other is energy control at Mitanda Home for the aged in Zambia. In this work, Wi-Fi is used for monitoring energy consumption of home equipment. Then home sensor collects the energy consumption data and analyzes them for energy approximation and control the home energy utilization schedule to slump the energy cost. Then energy data of home servers evaluates them, and generates useful statistical examination information aggregated by the remote energy management server. The system is intended to be a more efficient means of energy saving and result in home energy cost reduction in the care homes for the aged people in Zambia. The system will add value to Zambian by promoting efficient electrical energy management.

## Assessing the Readiness of Students to Use Mobile Applications in Collaborative Learning Looking for Answers with UTAUT

Phillimon Mumba and Maybin Lengwe (Copperbelt University, Zambia)

To improve student performance and retention rates, higher institutions of learning are constantly researching on the approaches, tools and techniques to use. In recent times, concepts such as mobile learning, electronic learning, collaborative learning, flipped classroom and deep learning have emerged. These describe the different approaches that institutions are using to improve student performance and retention rates. However, the successful implementation of an approach largely depends on the willingness of the users (learners and educators) to use. In this paper we are using the Unified Theory of Acceptance and Use of Technology (UTAUT) to determine the willingness of students at Copperbelt University to use mobile application-aided collaborative learning in their studies.

## Examining Factors Influencing E-Banking Adoption: Evidence from Bank Customers in Zambia

Hillary Chanda (The Copperbelt University, Zambia)

This paper contributes to the electronic banking (e-banking) literature by applying the modified Technology Acceptance Model (TAM) in an under-researched Zambian context. Specifically, it examines the influence of e-banking technology's perceived usefulness, perceived ease of use and trust (safety and credibility) on e-banking adoption. Based on a quantitative correlational design, primary sample data were collected from 222 bank customers from two of Zambia's largest cities. The findings indicate that the modified TAM model is applicable in the Zambian context and that perceived usefulness, ease of use and trust each significantly positively influences attitude to e-banking. In turn attitudes to e-banking influence intention and actual adoption of e-banking services. For scholars, practitioners and policy makers, the study shows that improving perceptions of trust (safety, security and credibility), usefulness and ease of use of e-banking systems would result in increased adoption. This paper is the first to extend the modified TAM model into the under-researched developing country context of e-banking in Zambia.

## The Factors Affecting the Adoption of Electronic & Mobile Banking

Hillary Chanda (The Copperbelt University, Zambia)

This study aimed at investigating the factors affecting the adoption of electronic and mobile banking among students of the Copperbelt University. The study objectives were; to assess the influence of users' perceived risk on adoption of mobile banking, to evaluate the effect of trust in mobile banking by users towards influencing their choice of adopting it, to examine how perceived convenience affects adoption of mobile banking and to explore how relative advantage influences adoption of mobile banking. This study employed a descriptive research design with a sample size 100 students. Data was collected using questionnaires which had a return rate of 92 percent. Data collected was edited, coded and analyzed using statistical package for social sciences (SPSS). Research results indicate that the adoption of mobile banking was not as high as it should have been. The main reasons found to be behind low adoption of mobile banking was risk of loss and fear of system failure. Additionally, customers' perceived risk was found to negatively affect adoption of mobile Banking service. On the other hand, perceived convenience was found

to positively affect adoption of mobile banking. The usefulness presented included; accessibility, saving of time and comfort. In addition, mobile banking services mostly used were account balance inquiry and funds transfer. In regards to trust, the reliability and integrity of mobile banking service providers was found to positively affect adoption of the service. The more reliable the service was found to be, the more adopted it was. Lastly, the mobile banking service was found to possess relative advantages in comparison to traditional banking services. Such advantages included; accessibility, saving of time, less cost, privacy and comfort. Consequently, the researcher recommends that banks undertake an aggressive marketing campaign to make customers aware of their electronic and mobile banking services. It should lay down strategies to leverage the service against other traditional banking services such as highlighting the conveniences and relative advantages presented by the service in order to make it attractive to customers. In addition, the bank should assure customers of the reliability and integrity of the service so that they can trust it and get to adopt it.

## Adoption Intention Of Mobile Banking, The Role Of Perceived Risk: A Comparative Study Between University Students In The United Kingdom And Zambia

Hillary Chanda (The Copperbelt University, Zambia)

The purpose of this study is to investigate the adoption intention of mobile banking, the role of perceived risk. Mobile banking was introduced in many countries, however, in some countries mostly in Africa it has proved to have been very successful, while in others it is in its early stages and in other countries, it has failed or taken very long for adoption and diffusion to take place. In United Kingdom, mobile banking is in its early stages and still developing. In Africa and in Zambia to be specific, mobile banking is quite successful and still developing. This study is a comparative study between students in United Kingdom and Zambia. It will analyse and evaluate the role of perceived risk as well as perceived usefulness, perceived ease of use and trust, on the adoption of mobile banking among university students in these two countries. The study reviewed the available literature of mobile banking technology. The findings of the research show that the students in Zambia (Copperbelt university students) did not only use mobile banking with banks, but they also used it with other companies, while of the students in the United Kingdom (Bournemouth University) that used mobile banking only used it with their Banks. The study also reviewed that Perceived useful, perceived ease of use and perceived risk have positive relationship with mobile banking among both the Bournemouth University respondents and the Copperbelt University respondents.

## Collaboration on the Web: a Review of Web 2.0 Social Software and Wikis

Hellen Syachaba and Douglas Kunda (Mulungushi University, Zambia)

The Web has changed the way in which we work nowadays, it has become more participatory, easier for users to contribute, share and work collaboratively. It offers new interaction possibilities that facilitate desktop-like interfaces; it is also the way in which new software is built by mashing up distributed components using service-oriented communication. The purpose of this article is to present literature review on some the technologies and services that has enabled the web to be more interactive and collaborative for the users. These include Wikis Social Software and Web 2.0

## Performance, Scalability and Quality of Service on Web: Challenges and Open Issues

Evaristo Chishimba and Douglas Kunda (Mulungushi University, Zambia)

The coming of web technology and accessing it through the Internet has led to the creation of a digital society, where almost everything is now connected and is accessible from anywhere. However, despite their widespread adoption, accessing the web sites pause a challenge to the users and the people managing the websites. This is so because the development of the web sites demands a lot of time it's costly and needs skilled power. In this paper, we discuss challenges associated with performance, scalability and quality of service for web. We discuss the solutions which can be employed when developing the web sites found and proposed in literature. Furthermore, we present an outline of Challenges of website performance and issues which can be avoided when developing the web. In addition, we present some of the open issues in web technology.

## Active Learning Environment a Comparative Analysis of Web Services (SOAP, RESTful, WSDL, UDDI)

Ng'andu Mwiiya and Douglas Kunda (Mulungushi University, Zambia)

Attention to Web services has swiftly increased since their introduction. Exchanging information among applications in an acceptable way is the main objective of web services. This communication among applications has brought a need for uninterruptible and continuous web services which is centered on the SOAP and REST standard. Network traffic and processing delays are some of the limitations of SOAP communications. These limitations can be overcome by the use of the RESTful architecture because REST is lightweight. This paper will outline the motivation for the introduction of web services. The concept of an active learning is discussed followed by the use of web services in active learning environment. A brief background of the two significant types of web services is given and comparisons of these two frameworks based on their service discovery, interface, security and general performance. Each of these types of web services has its own value, advantages and disadvantages. Therefore, one has to understand the circumstances in which each of these designs ought to be used.

## Web Engineering: Challenges and Open Issues

Selina Halubanza and Douglas Kunda (Mulungushi University, Zambia)

The web-based systems haven't been developed following laid down software development approaches in the recent past. Managing quality control as well as software quality assurance has been a challenging task. A concern has therefore arisen regarding not only the security of web-based systems but also integrity issues. A systematic approach of software development is hence advocated for through software engineering. Web engineering promotes "the use of sound scientific, engineering and management principles, and disciplined and systematic approaches to development, deployment and maintenance of Web-based systems." This paper hence looks at key web engineering issues and challenges through a review of various similar literatures.

## Active Learning Environment: Web Metrics, Monitoring and Analysis, Open Issues

Joshua Mubila and Douglas Kunda (Mulungushi University, Zambia)

Technology has made the world to become a global village in the area of e-commerce, people in different localities are enabling to buy thing online as if they are right in the shopping store and supplier are enable to sell the product as if they are right in their shop with their customer. This is made possible through web browsers which create a platform for both the supplier and the buyer, it enables the buyer to have access to the supplier's website. The suppliers or web operators may want to know the number of people visiting their website and what the visitors are doing on their website, what pages they visit, how long it takes the visitor on their website before leaving. Web operators may also want to know what makes visitor leave their website very fast, what makes them visit one page and leave, and how are the visitors getting to their website. The web operator must not ignore the needs of their visitors, taking care of the visitors needs will increase traffic coming to their website hence increasing sales. Therefore, this paper will look at Web Metrics, Monitoring and Analysis.

## E-government Implementation Models and Challenges: The Case of Zambia

Akabana Kalaluka and Douglas Kunda (Mulungushi University, Zambia)

Since the development of computers, the world has seen many innovations as a result of their extensive use. Information and Communications Technologies (ICT) have made a significant impacted how governments, business and society at large process and interact with each other. For over two decades now, governments the world over have realized that ICT's can facilitate new approaches to service delivery, stakeholder engagement and information access. With the push for e-government in full force, challenges emerged which are common in nature for countries in the developing world.

## Electronic Publishing on the Web: Challenges and Open Issues

Mark Mwale and Douglas Kunda (Mulungushi University, Zambia)

The web as an internet based hypertext system is a good platform for disseminating and sharing information. Electronic books, magazines, newspapers and the development of digital libraries are all examples of electronic publications available on the web. Over the years, electronic publishing on the web has undergone a lot of innovations thereby becoming too complex. These include; editing electronic books, magazines, journals and other documents meant for public consumption on particular areas of concern or interest. In view of these developments, the paper discusses electronic publishing on the web with special interest in challenges and issues.

## Benefits and Challenges in the Use of Cloud Computing in Colleges of Education in Zambia

Phyllis Siyomunji (David Livingstone College of Education, Zambia); Douglas Kunda (Mulungushi University, Zambia)

Cloud computing is a computing technology that allows or enables access to a pool of shared configurable resources. It is involves computer networks, servers, storage, applications and services and these can be quickly provided over the internet without much effort from management. There are several cloud service models that are used and these include Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Mostly these models offer increasing abstraction and they are often seen as layers in a stack. The paper discussed the benefits the education sector and its clientele stood to derive from the usage of cloud computing applications. Benefits such as automated assessments and examinations ,lower long-term costs, instant feedback to students ,creation of digital records of student growth and development, greater storage efficiency increased productivity and low operational variability where discussed. We noted in the paper that there was the association between a set of variables such as IT self-efficacy, perceived cloud ease of use, computer anxiety, and users' perception of the usefulness and effectiveness of cloud computing applications. In this vein, these variables where considered as the possible factors that could cause the adoption and/or not adoption of cloud computing application in public Colleges of Education in Zambia. Finally, the paper discussed the various challenges that come with cloud computing in education. Some of these challenges include security, data privacy, as well as insufficient network. However, it was concluded and recommended that cloud computing technology as a tool can be used and it should be rapidly expanded in public Colleges of Education in Zambia. This is especially so, because cloud computing has the ability to collaboratively share, edit, process, and store huge amounts of data and this have obvious applications within the research and educational communities..

## A Systematic Literature Review of Big Data Analytics Implementation in Health Care

Richard Chellah (Mulungushi University & Copperbelt University, Zambia); Douglas Kunda (Mulungushi University, Zambia)

The implementation and use of Big Data analytics in healthcare has huge potential for improving the quality of care, reducing waste and error, and reducing the cost of care. This systematic review of Big data Analytics literature aims to determine the opportunity of Big Data analytics in healthcare including its applications and challenges in its adoption in healthcare. It also intends to identify the strategies to overcome the challenges. A systematic search of the articles was carried out on four scientific databases: ScienceDirect, PubMed, IEEEXplore and ResearchGate. The articles on Big Data analytics in healthcare published in English language literature from January 2011 to June 2018 were shortlisted and considered for the review. The review selected descriptive articles and usability studies of Big Data analytics in healthcare and medicine. The analyses of these articles found that: researchers lack consensus about the operational definition of Big Data in healthcare; Big Data in healthcare comes from the internal sources within the hospitals or clinics as well external sources including government, laboratories, data aggregators, medical journals etc.; Big Data analytics finds its application for clinical decision support; optimization of clinical operations and reduction of cost of care. The major challenge in adoption of Big Data analytics is non-availability of evidence of its practical benefits in healthcare. This review study unveils that there is a little of information on evidence of real-world use of Big Data analytics in healthcare. This is can be owed to the fact that, the usability studies have considered only qualitative approach which describes potential benefits but does not take into account the

quantitative study. It should also be noted that, the majority of the studies were from developed countries which brings out the need for promotion of research on Healthcare Big Data analytics in developing countries.

## Assessment of the Impact of Social Networking Sites Usage on Students' Academic Performance: A Systematic Review

Beauty Lweendo and Douglas Kunda (Mulungushi University, Zambia)

The implementation and adoption of contemporary Internet services in Zambia has seen the greatest number of citizens originating from Zambia spend their time online on social networking sites (SNSs). Social networking sites are online platforms that provide individuals with an opportunity to manage their personal relationship and remain updated with the world. This study will focus on assessing the impact of social networking sites usage on the students' academic performance in the selected Colleges of Education in Zambia. This is with the view of establishing whether the use of these SNSs has positive or negative impact on students' academic performance in the modern education arena. The study will be based on a research model with six hypotheses. Quantitative approach will be used. The sample size will be 500 students from Colleges of Education and the researcher will use stratified random sampling to select participants because of the uneven enrolment at Colleges of Education. Data for the study will be collected through self-administered questionnaire and data will be analysed using Pearson's correlation and SPSS tool will be used. The six hypotheses will be tested using Pearson's correlation and Cronbach's alpha will be used to test validity and reliability of the instruments. The general objective of this study is to assess the use of SNSs by college students in selected Colleges of Education in Zambia with the view of determining if the use of SNSs has impacted upon their academic performance and if so, how. It is hoped that the information to be generated from this study will first and foremost help stakeholders like colleges to re-assess the importance and significance of the use of SNSs by students in their educational studies and help not only students but also other individuals throughout the country on how to use SNSs responsibly and for the development of their knowledge and skills.

## Identity Management Based on Frontal Facial Recognition for Voters Register in Zambia

Lubasi Musambo and Jackson Phiri (The University of Zambia, Zambia)

biometric technology offers a great opportunity to identify individuals, authenticate individuals and separate individuals. Using these advantages, an election or voting model can be developed to perform elections for a country such as Zambia. Zambia currently uses a manual based voting or election model that heavily relies on paper presented documents that must be physically verified and or matched to existing prior collection information before an individual is allowed to participate in an election or a voting system. This paper proposes a frontal facial election based biometric model that can be used to rid the current election system of redundancy and introduce paperless, accurate and efficient identification, authentication and voting process. A baseline study conducted shows that biometric authentication based our model improves a work related process such as a voting system. We start by introducing the elements that make a biometric model ideal, we then give an insight into the Zambian based election system and then we review various biometric technologies available and then finally introduce our biometric model.

## The Effects of Testing Data Size on Isolated Word Recognition

Zita Lifelo, George Mufungulwa and Tracy Chisanga (The Copperbelt University, Zambia)

In this paper, features for automatic speech recognition (ASR) are based on the contribution of short time energy (STE) and zero-crossing rate (ZCR) as a combined approach on one hand, and STE only on the other. The combined (STEZCR) approach emphasizes important short term frames in a signal while alleviating most noisy and silent frames. Little is known about speech feature extraction algorithms in Zambia. This paper aims to ascertain whether the vocabulary size of the testing data set as a ratio of the vocabulary size of trained data has an effect on automatic speech recognition. In addition, the paper seeks to consider whether gender of subjects does influence system performance in terms of recognition accuracy. The study applies Mel frequency cepstrum coefficients (MFCC) and linear prediction coding (LPC) feature extraction techniques respectively. The vocal frequency is reduced from 3 to 1 over the MFCC and LPC respectively. Testing results of the proposed combined speech features on isolated

Japanese speech phrases was done in a clean environment. The average recognition accuracy improvement of 7.83% and 10.28% is achieved in word accuracy when the vocal frequency is reduced from 3 to 1 over the MFCC and LPC respectively. The results clearly confirm that the vocabulary size of the recognizer has an effect on automatic speech recognition. In addition, the combined STEZCR approach performs better on female subjects with the recognition accuracy improvement of between 0.20% and 1.5% while STE performs better on male subjects at 5.50% and 9.0% on MFCC and LPC respectively. However, the results show that gender of subjects has no direct influence on system performance but that a combination of the feature extraction and VAD techniques influences ASR system performance.

## Technology Paradigm Shift: A Case of Ethical and Unethical Hackers and Their Subtle Tools

Raphael Banda and Jackson Phiri (The University of Zambia, Zambia); Mayumbo Nyirenda (University of Zambia & Hokkaido University, Zambia); Monde M Kalumbilo-Kabemba (The University of Zambia, Zambia)

Paradigm shift implies change of rules or change of the way we do same things. You don't actually change the way you have been things thins but the way you shift camp but still doing the same things you know better. In doing so some old ways of doing things do not stop working but will just get into one's way. There are three major types of hackers that we can identify although it is not easy to draw a line between them. The three hackers I will discuss are the Black hat the grey hat and the white hat. The black hats are the bad people, the grey hats are the intermediate ones somehow on the fence and the white hats are the good people. The good people may be referred to as the ethical hackers; the black people could be called the bad people. These people make people's lives difficult and can prove to be costly to the company. We will also look at some of the tools the black hat hackers have produced or developed to hack into selected systems. Why do they do it? For many its mostly to show off their computer skills and prowess or to gain some access to private data. Some of the hacked tools developed elsewhere have found their way to Zambia through the internet and other similar media like local area networks, flash, and CD media and the internet. In Zambia it is illegal to use illegal materials but due to the prices that are normally high for an ordinary Zambian it's difficult to prevent such dark business. If its wrong to use hacked software is it possible to devise a way of preventing such cybercrime.

## Developing an Automated Fall Army Worm (Faw) Identification and Early Warning and Monitoring System Based on Artificial Neural Networks Techniques

Francis Chulu (University of Zambia, Zambia); Jackson Phiri (The University of Zambia, Zambia); Phillip Nkunika (University of Zambia, Zambia); Mayumbo Nyirenda (University of Zambia & Hokkaido University, Zambia); Monde M Kalumbilo-Kabemba and Miyanda Miyanda Moonga (The University of Zambia, Zambia)

Since its reported presence in Africa in 2016, the fall army worm (FAW-Spodoptera frugiperda) has caused major damage to a good number of plant species including maize which is a stable food for most African countries. Their presence in Africa pauses a challenge to the food security in many African countries contributing to the already existing food problem that the continent has been facing. This possess a challenge to stakeholders such as FAO, governments, Universities and other stakeholders involved in research to come up with precise and proactive methods of monitoring and controlling the FAW pest. This paper therefore, proposes a study to develop an automated fall army worm identification and early warning and monitoring tool for the Zambian species based on Artificial Neural Network (ANN). The study will aim to address current challenges that entomologists are facing when using the pheromone traps as a way of monitoring the occurrence of FAW pests in Zambia. We will modify pheromone traps and automate them with sensors for automatic data collection. We will develop an algorithm based on ANN for identifying the FAW moth, then we develop web and mobile applications integrated with geographic information system (GIS) technology. The developed system will be able to provide some near real time FAW occurrences in Zambia. The. The tool will improve the accuracy and efficiency of FAW monitoring and reduce manual data collection thereby reducing the aspect of human intervention. In addition, it will act as a source of data that can be used by all stakeholders ranging from FAO personnel, government, small scale and commercial farmers in making good and informed decisions.

## An Application of Machine Learning Algorithms in Automated Identification and Capturing of Fall Armyworm (FAW) Moths in the Field

Simon H. Chiwamba (The University of Zambia & MVCL, Zambia); Jackson Phiri (The University of Zambia, Zambia); Phillip Nkunika (University of Zambia, Zambia); Mayumbo Nyirenda, Monde M Kalumbilo-Kabemba and Philemon H. Sohati (The University of Zambia, Zambia)

As farmers' struggles with unpredictable rainfall patterns, the horror of emerging crop pests that will likely affect the quality and quantity of their harvest does not spare them therefore it is very important to protect the crop by monitoring the pests. While the Southern African Development Community (SADC) experienced normal rainfall in the 2016/2017 farming season, the regions' food security was threaten by the outbreak of the fall armyworm (FAW), a pest that affects crops including maize, the staple food, in DRC, Botswana, Malawi, Namibia, Swaziland, South Africa, Zambia and Zimbabwe. In Zambia alone, attempts to control the FAW pest that had affected approximately 130,000 hectares of crops costed about US$ 3 million. This led SADC in collaboration with Food and Agriculture Organization (FAO) to classify the outbreak as a threat to food and nutrition security, and livelihoods of smallholder farmers in the region. This has posed a number of challenges to the stakeholders such as the University of Zambia (UNZA) School of Agricultural Sciences and the Department of Biology Science to come up with more precise and proactive measures for monitoring and controlling the pest. This paper proposes a study to develop an automated system for identifying and capturing images of FAW moths in the field using machine learning (ML) in Zambia. The proposed study will aim to address current challenges that the UNZA School of Agricultural Sciences and the Department of Biology are facing in monitoring the FAW using the pheromone traps. The pheromone trap will be modified to include a raspberry PI, vision sensor and motion sensor to be used for data collection. The system will further integrate GPRS and 3G/4G connectivity as a means of sending data to the cloud server for further analytics.

## A Comparative Analysis of Web Searching and Information Discovery Techniques: Systematic Literature Review

Darius Bwalya and Douglas Kunda (Mulungushi University, Zambia)

Never before has there been so much information explosion than in the 21st century. With this increase comes, a great challenge in accessing the actual desired content that is relevant information out of superfluous data links that are often harvested at every search, which often leads to more time being wasted in sifting unwanted links with ultimate delays. Just like how efficient is achieved when there is systematic and standard orderly arrangement of files along with efficient access method of this information, there has been even a greater need for standardized formatting of information and for technically devising efficient methods of accessing information in the cyber space. This has given rise to information formatting and search technology that are technically advanced namely Web search engines, Web crawling, Web indexing, Page and site ranking, Spam detection, Content ranking, Collaborative filtering, Social recommendation, Personalization, Social tagging. In this paper these web formatting and search technologies are analyzed and compared in order to give insight when it comes to user assisting search methods to use in the design of specific purpose web sites and applications by software developers, e-marketers and advertisers and many other purposes too numerous to mention.

## Challenges of Identity Management Systems and Mechanisms - A Review of Mobile Identity

Raphael Banda and Jackson Phiri (The University of Zambia, Zambia); Clayton Sikasote (University of Zambia, Zambia)

Digital identity and management lays the groundwork necessary to guarantee that the Internet infrastructure is strong enough to meet basic expectations for security and privacy. "Anywhere, anytime" mobile computing is becoming real; in this ambient intelligent world, the choice of identity management mechanisms will have a large impact on social, cultural, business, and political aspects of our lives. From their point of view, Privacy is a human need, and all of society would suffer from its demise; people have hectic lives and cannot spend all their time administering their digital identities. Most systems do not fulfill several of these tests; they are particularly deficient in fine-tuning the access control over identity to minimize disclosure of data. This paper looks the procedure for a defence mechanism against MITM (Man in the Middle). We have mainly centred on how best we can create

passwords that cannot be easily decoded by middlemen. We have surveyed how the requirements for user-centric identity management and their associated technologies have evolved, with emphasis on federated approaches and user centricity. Second, we have focused on related standards XRI and LID, as well as platforms, mainly ID-WSF, OpenID, CardSpace, Sxip, and Higgins and finally, we have looked at identity management in the field of mobility with focus on the future of mobile identity management.

## A Review of System Intrusion Prevention Techniques and Tools in Developing Countries

Yvonne N Akende (University of Zambia, Zambia); Jackson Phiri (The University of Zambia, Zambia)

The area of intrusion detection and prevention is the central concept in overall network and computer security architecture. With the expansion of the internet and e-commerce over the years, we see individuals, governments and businesses having online presence and creating huge investments onto online platforms in developing countries. In the recent past, we witness these critical infrastructures becoming prime targets and more vulnerable to major cyber-attacks than ever before. Over a million compromises of electronic data occur annually worldwide, ranging from simple system intrusions to more sophisticated and malicious attacks resulting in huge losses for businesses. This paper outlines the review of the current situation as it pertains to System Intrusions and Prevention from the developing countries perspective, pin pointing the general issues and concerns in leveraging the full benefits of system intrusion prevention techniques and mechanisms discussed in this paper. The analysis of these issues can be utilized to suggest the future direction of System Intrusion Prevention and cyber security as a whole.

## A Review of Major Local Area Network Security Challenges

Jimmy Katambo (University of Zambia, Zambia); Mayumbo Nyirenda and Jackson Phiri (The University of Zambia, Zambia)

Cloud computing, file sharing, social networking and other computing trends increase the vulnerability of networks to security threats and challenge network resources. The shared computational resources and operations increase the vulnerability of these networks to intrusions and security threats such as spams, backdoors, proxy server intrusions and denial of service (DoS) attacks, which may jeopardize the confidentiality, accuracy and accessibility of shared data. Based on the method of deployment, Intrusion Detection Systems (IDS) can be classified as Host Based IDS (HIDS) and Network Based IDS (NIDS). Network intrusion detection is a dynamic research area as intruders or attackers have increased attacks on all kinds of networking set-ups. An intrusion can be caused by an insider or an outsider. Security policy is a main mechanism of information security management. While there are a lot of security-related standards and guidelines which specify requirements for high-level security policies, implementation of network security policy still depends on interfaces provided by network security solutions. The need for tools to help network administrators in the network management process is increasing. Access Control Lists (ACLs) refer to security rules associated to network equipment, such as routers, switches and firewalls. Firewalls provide a mechanism for protecting enterprises from the less secure internet over which customers or collaborating partners transfer packets destined for the corporate network.

## The Major Wireless Network Security Challenges - A Review

Jimmy Katambo (University of Zambia, Zambia); Mayumbo Nyirenda and Jackson Phiri (The University of Zambia, Zambia)

Wireless networks can be broadly categorized into two major classes namely wireless ad hoc networks and cellular networks. The main difference between these two is whether a fixed infrastructure is present. While cellular networks require fixed infrastructures to support the communication between mobile nodes and deployment of the fixed infrastructures is essential, Wireless ad hoc networks do not require a fixed infrastructure; thus it is relatively easy to set up and deploy a wireless ad hoc network. Security Protocols for Sensor Networks are a family of security protocols, which were specially designed for low end devices with severely limited resources, such as sensor nodes in sensor networks. This paper reviews a number of papers in the areas of wireless network security. To allow routers to automatically discover new routes and maintain their routing tables, routers exchange routing information periodically. Wireless Sensor Networks are not like Wired Sensor Networks or other types of wireless

networks, and it is easier for the Wireless Sensor Networks to be attacked and more challenging to ensure the security of the Wireless Sensor Network. As a result, the security of Wireless Sensor Networks has been widely studied and many wonderful security policies have been proposed

## Security, Privacy and Integrity in Internet of Things

Chalwe Musonda (University Of Zambia, Zamabia & Chingola Secondary School, Chingola, Zambia); Monde M Kalumbilo-Kabemba (The University of Zambia, Zambia); Mayumbo Nyirenda(University of Zambia & Hokkaido University, Zambia); Jackson Phiri (The University of Zambia, Zambia)

This paper addresses the Data Security, Privacy and Integrity in Internet of Things (IoTs) it borrows and characterized by heterogeneous technologies, which concur to the provisioning of innovative services in various application domains. In this scenario, the satisfaction of security and privacy requirements plays a fundamental role in information system security. Such requirements include data confidentiality and authentication, access control within the IoT network, privacy and trust among users and things which are connected on the internet and the enforcement of security, privacy and integrity policies which exist as at now. Traditional security issues countermeasures cannot be directly applied to IoT technologies due to the different standards and communication stacks involved in the technology. Moreover, the high number of interconnected devices arises scalability issues; therefore a flexible infrastructure is needed so that its able to deal with security threats in such a dynamic environment. It is the job of this paper to highlight few issues among them security, privacy and Integrity in IoTs. Different vision of this Internet of Things paradigm are reported and taken researched and reviewed. And what comes out is the major issues which shall be faced by the research community. There is more research to be done in this area especially of distributed intelligence for smart objects are just the most relevant.

## A Review of Identity Attribute Metrics Modeling Based on Distance Metrics

Felix Kabwe (University of Zambia, Zambia); Jackson Phiri (The University of Zambia, Zambia)

The growth in the use of services on the World Wide Web has brought about in the proliferation of online fraud. This is hinged on the fact that cyber fraudsters and criminals would hide their online identities to steal services and other valuables. Work has been done in the past on strengthening of identity management systems as a way to arrest this growing problem. This study considers past work on the subject matter and builds on developing the metrics models in order to provide quantitative analysis to quantify the credential identity attributes in online services. Metrics models will be explored that would quantify the credential identity attributes which will help in uniquely identifying the real owners of the digital identities before services and other assets could be issued to requesters of the same. A review of literature in this area of interest has been done and therefore, the identified area of interest adds value to the resolution of the said problem.

## Assessing the Readiness of Students to Use Mobile Applications in Collaborative Learning Looking for Answers with UTAUT

Phillimon Mumba and Maybin Lengwe (Copperbelt University, Zambia)

To improve student performance and retention rates, higher institutions of learning are constantly researching on the approaches, tools and techniques to use. In recent times, concepts such as mobile learning, electronic learning, collaborative learning, flipped classroom and deep learning have emerged. These describe the different approaches that institutions are using to improve student performance and retention rates. However, the successful implementation of an approach largely depends on the willingness of the users (learners and educators) to use. Even the best approaches or techniques cannot yield fruitful results if the users are not willing to use them. This paper assesses the willing of students at Copperbelt University to use mobile application-aided collaborative learning in their studies. In this paper, we have identified that students are very confident in collaborative activities. Higher learning institution should incorporate learning activities requiring collaboration among students. This will help the students learn how to work with their peers and to encourage them to take charge of the learning process.

## Web Design Tools: Challenges and Open Issues

Raphael Gondwe and Douglas Kunda (Mulungushi University, Zambia)

We are living in this world where technology has taken a center stage in all human endeavours. Accessing and sharing of information is just by the finger tips and this is facilitated by the web (Internet). The web in this scenario imply the way of exchanging information between computers on the Internet, tying them together into a vast collection of interactive multimedia resources. It is for this reason that web designers should carefully look at open issues and challenges of web design as regards to Hypertext, Hypermedia, Markup languages and XML related technologies. If only designers and developers of the web can embrace and address these issues and challenges pertaining to web design, then when can realise the potential merits that they may bring on board. Thus, the prime objective of this paper is to provide systematic theoretical discussion on web technologies in terms of open issues and challenges.

## Web Based Monitoring and Detection of Copper Cable Cuts in Fixed Access Networks

Ndiwa Mutemwa, Moris Moris Matoomana and Makaita Masaita (Copperbelt University, Zambia)

Telecommunication service providers relying on copper cables to deliver services to their subscribers are facing massive cable thefts in Zambia leading to unreliable service delivery, loss of revenue and high network maintenance costs. The theft of copper cables has been necessitated by the rise in price and demand for copper globally. Current methods employed to monitor cable cuts in fixed access networks are not only inefficient but are also costly while replacing the copper cables with optic fibre cables might not always be appropriate. This paper therefore discusses a proposal to not only monitor the status of copper cables, but also to detect cable cuts in fixed access networks using cable capacitance and sending the captured cable status information to a centralized remote monitoring centre in real time. A copper cable exhibits capacitance between two conductors which are insulated from each other and this capacitance is directly proportion to the length of the cable. Hence by determining changes in cable capacitance, the change in cable length can be calculated. In the proposed system design a microcontroller calculates the cable distance using the cable capacitance changes. Upon detection of a cable cut in the fixed access network, the system automatically updates the web server at the remote monitoring centre with information indicating the distance from termination to the cable cut through a gateway. It also alerts personnel in the remote monitoring centre by sending an alert signal to the web server thus enabling appropriate intervention measures to be taken to avoid the cable theft and reduce outage time

## Towards Increased Online Visibility of Scholarly Research Output in Zambia

Lighton Phiri (University of Zambia, Zambia)

Scholarly research and publication forms an integral part of the core functions of Higher Education Institutions (HEIs). It is generally standard practice for HEIs to deposit scholarly output into publicly accessible Institutional Repositories (IRs). While Zambia has seen a rise in the number of HEIs, with a total of six Public HEIs and 60 Private HEIs, there is little online visibility of scholarly output generated by these HEIs. A bibliometric analysis, focused on electronic theses and dissertations (ETDs), was conducted by harvesting scholarly publications from HEIs IRs, in order to demonstrate the low online visibility of scholarly research output in Zambia. We also outline technological initiatives, by using case examples from The University of Zambia, that can be employed to potentially increase the online visibility of HEIs scholarly output. Specifically, we illustrate how subject repositories and downstream aggregate services can be utilised to increase the visibility of scholarly output. The study shows that only two HEIs have established IRs, with noticeably low scholarly publications by academic staff. In addition, there is a noticeably long delay between the publication date of the ETDs and the ingestion date into the IRs. In addition, while not comprehensive, the proposed initiatives demonstrate technological initiates that could be employed to increase the visibility of scholarly research output.

ISBN 978-9982-70-787-9


THIS PAGE IS INTENTIONALLY LEFT BLANK

# Towards Increased Online Visibility of Scholarly Research Output in Zambia

Lighton Phiri
*Department of Library and Information Science*
*University of Zambia*
Lusaka, Zambia
lighton.phiri@unza.zm

***Abstract*—Scholarly research and publication forms an integral part of the core functions of Higher Education Institutions (HEIs). It is generally standard practice for HEIs to deposit scholarly output into publicly accessible Institutional Repositories (IRs). While Zambia has seen a rise in the number of HEIs, with a total of six Public HEIs and 60 Private HEIs, there is little online visibility of scholarly output generated by these HEIs. A bibliometric analysis, focused on electronic theses and dissertations (ETDs), was conducted by harvesting scholarly publications from HEIs IRs, in order to demonstrate the low online visibility of scholarly research output in Zambia. We also outline technological initiatives, by using case examples from The University of Zambia, that can be employed to potentially increase the online visibility of HEIs scholarly output. Specifically, we illustrate how subject repositories and downstream aggregate services can be utilised to increase the visibility of scholarly output. The study shows that only two HEIs have established IRs, with noticeably low scholarly publications by academic staff. In addition, there is a noticeably long delay between the publication date of the ETDs and the ingestion date into the IRs. In addition, while not comprehensive, the proposed initiatives demonstrate technological initiates that could be employed to increase the visibility of scholarly research output.**

***Index Terms*—bibliometrics, digital libraries, OAI-PMH, repositories**

## I. Introduction

Higher Education Institutions (HEIs), in Zambia, play the crucial role of providing training towards the attainment of advanced degrees such as Masters and Doctoral degrees. In addition, HEIs conduct research that is aimed as solving many of society's pressing problems, with a key output being scholarly research publications.

The Higher Education Authority (HEA)[1] of Zambia—through the Higher Education Act of 2013 [1]—has been given the legal mandate to register Private HEIs and, more importantly ensure that HEI quality is not compromised. With the increasing demand of higher education, Zambia has seen a steady increase in the number of HEIs: there are a total of six Public HEIs [2] and 60 Private HEIs [3].

While there has been an increase in the number of registered HEIs and, corresponding enrolment rates of postgraduate students, the online visibility of scholarly research output is still noticeably low in Zambia. This paper outline the extent of the low online visibility of HEIs scholarly research output. In addition, the paper describes initiatives currently being undertaken to facilitate the increased online visibility of HEIs research output.

This paper contributes the following: (1) Empirical evidence showing the low online visibility scholarly research output generated by HEIs in Zambia. (2) Demonstration of initiatives that can potentially increase the visibility of scholarly research output generate by HEIs in Zambia.

The remainder of this paper is organised as follows: Section II discusses literature related to this work. Section III describes the methodology, while the results and discussion are presented in Section IV. Section V outlines initiatives that could potentially lead to increased online visibility of scholarly output and, finally, Section VI concludes the paper and outlines potential future work.

[1]http://www.hea.org.zm

## II. Related Work

### A. Bibliometric Analysis of HEIs in Zambia

Kalyambanino's dissertation examined faculty productivity at University of Zambia (UNZA) by analysing their research and publications. A mixed-methods approaches, involving questionnaires and interview guides with 251 participants, was used to gather data. The study suggests a low publication output, with 19.5% and 39% academic staff indicating having published books and articles in the previous two years, respectively [4].

Akakandelwa analysed publications authored by academic staff at UNZA by downloading publications papers authored between 2002 and 2007 from the Thomson Scientific database [5]. The publications were analysed in order to determine authorship patterns and collaboration. The average publication count was 36.7, with a highest publication count of 63, recorded in 2006.

Ahmed et al. conducted a mapping of postgraduate research in the School of Medicine, at UNZA, in order to explore research characteristics of the Master of Medicine programme [6]. A desk review of the Master of Medicine programme dissertations was conducted by reviewing manuscripts that had been published between 1986 and 2009 and, deposited in the Special Collections of The UNZA Library. In contrast, this work is focused analysing the online visibility of ETDs that have been deposited on the UNZA institutional repository (IR).

### B. Scholarly Research Visibility

In an attempt to explore alternative aspects for measuring the impact of The Medical Journal of Zambia[2], Kanyengo et al. reviewed online and hard-copy literature. Their online review was an online visibility assessment of the journal on platforms such as Google Scholar[3], ResearchGate[4] and academic databases such as Africa Journals Online[5]. Abrahams et al. state that the higher education sector in Southern Africa is, in part, dependant on universities' capacity to produce, communicate and use research output for educating future generations. However, they note that research output in the majority of Southern African universities is not visible [7].

Czerniewicz and Wiens conducted a study to assess the online visibility of poverty alleviation research in South Africa. Their analysis of indexed research on Google Scholar indicated relative online invisibility of research in the area [8]. In another study aimed at exploring the potential role of digital affordances in knowledge production and dissemination, Czerniewicz et al. observe that while Southern climate change researchers have a discoverable online presence, it is uneven and typically restricted to social media [9].

One of the the findings of SCAP was that Southern African research is marginal invincible in the global context [10]. Interestingly, another SCAP finding was that most universities typically have the technology required for effective scholarly communication.

### C. Software for Increased Online Visibility

Scholarly publications are generally organised in collections referred to as Digital Libraries (DLs) [11]. There are a wide variety of open source DL software tools and services that are used for storing and, making available scholarly research output. HEIs generally use such DL tools for implementing IRs [12]–[14] and, increasingly, electronic journals [15].

While the DL tools have varying implementations [16], they offer generic services for facilitating core DL features like searching and browsing. More importantly, their implementations are standards based, integrating protocols for effecting ingestion and discovering of content. For instance, protocols such as the OAI-PMH are effective at facilitating the harvesting of metadata from external repositories.

## III. Methodology

This section outlines the in-depth analysis conducted to explore the online visibility of scholarly publications for HEIs in Zambia. As a first step, the domains for the six Public HEIs were crawled to determine if their scholarly output is visible only and, specifically, to determine if they have established IRs. Digital objects from HEIs with IRs were then harvested and analysed.

[2]https://www.mjz.co.zm
[3]https://scholar.google.co.za
[4]https://www.researchgate.net
[5]https://www.ajol.info

### A. Harvesting Digital Objects

Open source Digital Library Systems (DLSes) that are used for setting up IRs are standards-based and implement interoperability protocols for effective storage and retrieval of digital objects. Digital Objects are generally composed of bitstreams—the digital resource consumed by end-users—and metadata—textual description of digital objects that provide for context about the digital resource.

The Open Archives Initiatives Protocol for Metadata Harvesting (OAI-PMH) [17] was used to harvest metadata from HEIs IRs, using the LibreCat Catmandu data processing toolkit [18]. The harvesting was done using the Dublin Core [19] metadata format—`metadataFormat=oai_dc`. In addition to the SetSpec field of the harvested metadata, the `Identifier`, `Date` and `Type` Dublin Core elements were used during the analysis stage, as outlined in Section III-B. Resources associated with each digital object were harvested using the Open Archives Initiative Object Reuse Exchange (OAI-ORE) standard [20]—`metadataFormat=ore`.

### B. Processing Harvested Digital Objects

The harvested digital objects were analysed in order to classify the different types of digital objects and, additionally, determine when the digital objects were published and ingested into the IRs.

*1) Metadata Processing:* Metadata elements were processed in order to determine hierarchies the digital objects were associated with, the designated classification of the digital objects and, publication and ingestion dates associated with the digital objects. Specifically, the following metadata elements were analysed.

- `SetSpec`—Indicates the various hierarchical structures within which digital objects are nested.
- `Subject`—Indicates the research topics associated with the digital objects.
- `Creator`—Indicates the authors of the digital resource associated with the digital object.
- `Contributor`—Indicates the entities that contributed towards the creations of the resource associated to the digital object.
- `Description`—Indicates additional contextual overview of the digital resource.
- `Date`—Indicates the publication and ingestion dates.
- `Type`—Indicates whether the document is an ETD, preprint or any other specified resource types.

*2) Bitstream Processing :* The digital object resources—PDF documents—were processed in order to determine if an ETD was Masters dissertation or Doctoral thesis. While analysing the `SetSpec` and `Type` metadata elements helped with the initial classification digital resources, parsing and processing the actual digital resource bitstream provided more comprehensive details. In addition, the processing acts as a mechanism for validating the data contained in the metadata. Furthermore, processing the digital resource helped determine the format of the digital resource—whether it was born digital or digitised; the latter adversely affects full-text searching.

As earlier stated, OAI-ORE was used to harvest digital resources (PDF documents). For each PDF document, the pdftk[6] utility was used to extract the first page of the document—the cover page—and, thereafter, the pdftotext[7] utility was used convert the PDF page to plain text. Finally, the resulting text document was analysed for useful information such as the ETD classification: Masters or Doctoral; additionally, the result plain text document was used to determine if the PDF was born digital or digitised.

## IV. Results and Discussion

This section presents results from the analysis of digital objects harvested from HEIs IRs. Table I indicates that out of the six Public HEIs, only Copperbelt University (CBU) and UNZA have IRs. The results also suggest that the 168 digital objects in the CBU IR are exclusively ETDs. One obvious observation is the low count of digital objects, especially that both CBU and UNZA graduate relatively large number of Masters and Doctoral students. This is especially the case for CBU which only has 168 digital objects in its IR.

An important point worth noting is that the UNZA IR also consists of final year students' capstone project reports, scanned copies of past examinations and digital objects from external research institutes. For simplicity and consistency, subsequent analyses presented in this paper are restricted to pre-prints and ETDs, published between 2010 and 2017.

TABLE I
Scholarly Publications Audit for Public HEI IRs

| Institution | Repository | Output | Items |
|---|---|---|---|
| University of Zambia | DSpace@ UNZA† | ETD | 3070 |
| | | Pre-print | 253 |
| | | Capstone | 1110 |
| | | Exams | 356 |
| | | External | 74 |
| Copperbelt University | DSpace@ CBU‡ | ETD | 168 |
| Chalimbana University | — | — | 0 |
| Kwame Nkhruma University | — | — | 0 |
| Mulungushi University | — | — | 0 |
| Mukuba University | — | — | 0 |

†http://dspace.unza.zm:8080/xmlui
‡http://dspace.cbu.ac.zm:8080/jspui

### A. Analysis 1. Digital Object Ingestion

In order to understand and better explain the low publication count, the dates the digital objects were published and their corresponding ingestion dates were analysed.

Figure 1 shows the HEI publications by year for the CBU and UNZA. The CBU IR only has digital objects published between 2011 and 2014, suggesting that nothing has been ingested into the IR since 2014. While the UNZA IR seems to be regularly updated with publications, there are obvious inconsistencies in the rate of ingestion. The pattern suggests that digital objects are ingested in batches as opposed to when when are published.

[6]http://www.pdflabs.com/tools/pdftk-the-pdf-toolkit
[7]http://manpages.ubuntu.com/manpages/bionic/man1/pdftotext.1.html

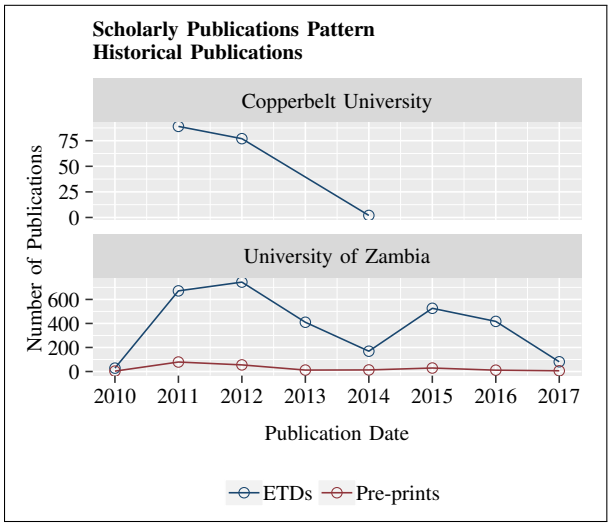



Fig. 1. Scholarly Publications by Year

The batch ingestion assumption for the UNZA IR is supported by Figure 2, which shows that the vast majority of publications were ingested between 2015 and 2016.

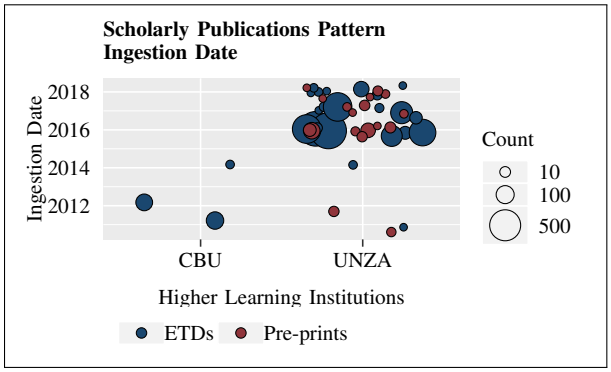



Fig. 2. Scholarly Publication Distribution by Ingestion Date

A further look at the publication distribution in Figure 2 is illustrated by Figure 3. The bubble plot indicates that most of the digital objects published are only ingested into the IR more a year after they are published, clearly affecting the online visibility of the resource. The long period has implications on not only the citation count of the digital object, but, more importantly, on other researchers potentially building up on related work—if content is not visible online, it becomes difficult for other researchers to realise this. In the case of ETDs, an argument could be made that this could ultimately result in the duplication of research conducted in various HEIs in Zambia.

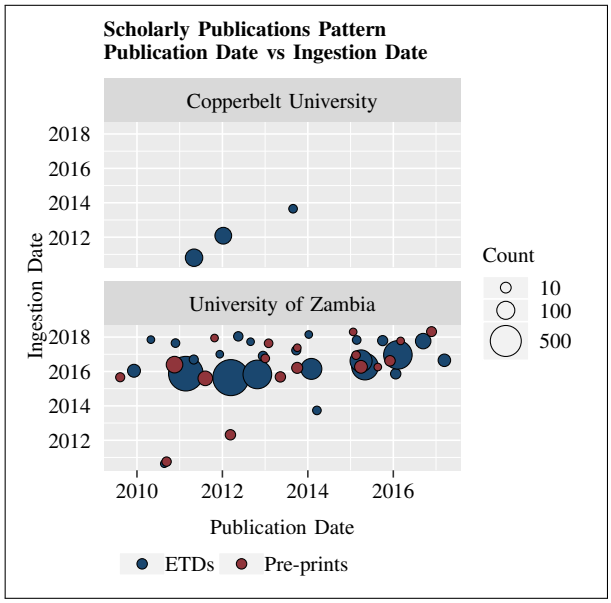



Fig. 3. Scholarly Publication Distribution by Ingestion Date

### B. Analysis 2. Quality of Metadata

Listing 1. Descriptive Metadata for a Sample ETD From The UNZA IR

```
<oai_dc:dc>
 <dc:title>
  Evaluation of [...] networks (ZAMREN)
 </dc:title>
 <dc:creator>Mwiinga, Jervas</dc:creator>
 <dc:subject>
  High performance computing--Zambia
 </dc:subject>
 <dc:subject>
  Research education networks--Zambia
 </dc:subject>
 <dc:description>
  THESIS M.ENG
 </dc:description>
 <dc:description>
  [...]
 </dc:description>
 <dc:date>2018-07-23T13:00:50Z</dc:date>
 <dc:date>2018-07-23T13:00:50Z</dc:date>
 <dc:date>2017</dc:date>
 <dc:type>Thesis</dc:type>
 <dc:identifier>
  oai:dspace.unza.zm:123456789/5275
 </dc:identifier>
 <dc:language>en</dc:language>
 <dc:format>application/pdf</dc:format>
 <dc:publisher>
  The University of Zambia
 </dc:publisher>
</oai_dc:dc>
```

Digital object metadata provides descriptive information about the digital object resource. While metadata comes in different variations—administrative metadata, structural metadata and descriptive metadata—descriptive metadata plays the crucial role in facilitating the effective browsing and searching of digital objects. DLSes, in part, index metadata elements to facilitate the discovery of digital objects. The specific issues observed are explained below, by making reference to Listing 1.

*1) Controlled Vocabularies:* Metadata elements such as `Subject`, `Creator` and `Type` are vital for facilitating effective browsing and as such, require the use of controlled vocabularies. However, analysing the metadata harvested from the UNZA IR suggests otherwise. For instance Listing 1 clearly indicates that the two `Subject` elements do not make use of a controlled vocabulary. Incidentally, most popular DLSes like DSpace provide browsing features based on `Subject`, `Date` and `Author`, illustrating the importance of controlling the vocabulary used to populate these fields.

*2) Missing Metadata Elements:* One key observation made during the analysis of metadata was that crucial metadata elements was missing. Most ETDs did not have the supervisor/advisor field included—this is generally included using the `dc.contributor.supervisor` or `dc.contributor.advisor` qualifiers. This is especially important because downstream services such as OATD[8] harvest ETDs from IRs and, in certain instances, crosswalk them to different metadata schemes like ETD-ms[9].

### C. Analysis 3. Bitstreams

Analysing the `PDF` documents yielded some interesting results. Only 38.66% of the digital objects were classified into their respective degree. Approximately 5.60% of the digital resources are suspected to be digitised since the resulting text file for the cover page had no content.

Interestingly enough, there were inconsistencies in the textual content on the cover pages of the ETDs. Further analyses would have to be conducted to determine if the inconsistencies are associated with publication dates for the ETDs.

## V. Tooling for Online Visibility

This section describes some initiatives that the author is involved with, which are aimed at increasing the online visibility of scholarly output at The UNZA.

### A. Electronic Journals

The UNZA presently publishes seven official journals [21] and, additionally, three journals run by the Directorate of Research and Graduate Studies, aimed at publishing postgraduate research output [22]. In addition to these official journals, there are departmental journals that are subject specific. While there are a few journals such as the Journal of Preventive and Rehabilitative Medicine[10] and the Journal of Library and

[8]https://oatd.org
[9]http://www.ndltd.org/standards/metadata
[10]http://medicine.unza.zm/research/journal

Information Science[11] that have transitioned into electronic format, the vast majority of journals are still print-based.

The author is involved in institutional initiatives that aim to migrate print-based journals to electronic platforms, using Open Journal Systems journal management system. Migrating the print-based journals into electronic format is certain to increase the visibility of scholarly publications.

### B. Subject Repositories

One of the reasons why there is a large time gap between the publication dates and ingestion dates of digital objects analysed in Section IV-A is possibly because the submission workflow is solely handled by the UNZA Library. A potential solution is to decentralise the process, using subject repositories, enabling authors to electronically submit their work. This could potentially ensure that the correct and appropriate metadata elements are submitted before the digital objects are actioned into the IR.

The OAI-PMH protocol can be used to integrate the subject repositories with the IR, as shown in Figure 4. The subject repositories could be school-specific or department-specific. Fundamentally, the subject repository acts as a data provider, enabling the IR to easily harvest metadata and bitstreams. Ongoing work is being conducted to assess the feasibility and effectiveness of this approach.

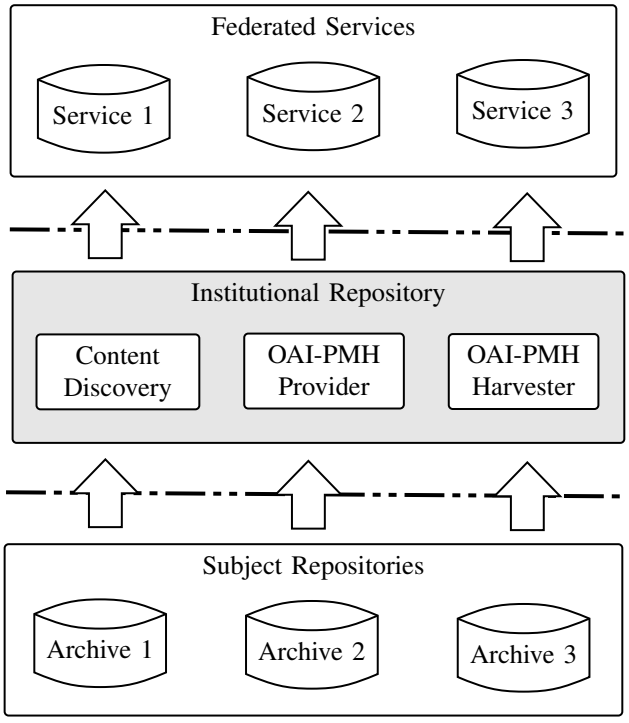



Fig. 4. Decentralised Architecture for Increased Visibility of Publications.

### C. ETD Harvester

Harvester services typically take advantage of OAI-PMH protocol to collect and aggregate digital objects into a central portal, enabling end-users perform centralised searching and browsing of content. Popular portals include national initiatives such as the South African National Electronic and Dissertation portal[12] and the global Networked Digital Library of Theses and Dissertations Union Catalog[13]. Similarly, a Zambian National ETD portal[14] has been set up to aggregate ETDs from the various HEIs in Zambia. Figure 5 shows a screenshot of the portal with ETD metadata harvested from CBU and UNZA IRs.

### D. Summary

This section has outlined some practical and actionable technology-centric approaches that relevant stakeholders can be undertake to work towards increasing the visibility of scholarly output. In particular, the case examples discussed illustrate the feasibility of these technological initiatives.

## VI. Conclusion and Future Work

This paper illustrates the extent of the low online visibility of scholarly research output in Zambia. HEIs IRs were empirically analysed by extracting digital object resources and corresponding metadata. Due to the noticeably low numbers of pre-prints in the IRs, emphasis was placed on ETDs. The findings highlight the low visibility of research and, additional factors that might ultimately affect visibility of research. The paper also describes technological initiatives that could potentially lead to increased visibility of scholarly output. While technology is a major contributing factors for increased visibility of research, working towards increasing the online visibility of research requires a multi-faceted approach that should also involve changes in institutional culture and research communication practices [10].

Ongoing work the author is involved with includes understanding barriers associated with electronic publishing and, the potential effectiveness and feasibility of using subject repositories. As part of future work, machine learning and crowdsourcing could be potentially employed to automatically verify, validate and re-classify digital objects that are not properly tagged.

## References

[1] National Assembly of Zambia, "The Higher Education 2013," 2013. [Online]. Available: http://www.parliament.gov.zm/node/3097

[2] Higher Education Authority, "Public HEIs." [Online]. Available: http://www.hea.org.zm/index.php/public-heis

[3] Higher Education Authority, "Registered Private HEIs." [Online]. Available: http://www.hea.org.zm/index.php/registered-private-heis2

[4] C. Kulyambanino, "Faculty Productivity at The University of Zambia: Exploring Research and Publication," Ph.D. dissertation, University of Zambia, 2016. [Online]. Available: http://dspace.unza.zm:8080/xmlui/handle/123456789/4899

[5] A. Akakandelwa, "Author Collaboration and Productivity at the University of Zambia, 2002-2007," *African Journal of Library Archives and Information Science*, vol. 19, pp. 13–23, 2009.

[6] Y. Ahmed, C. Kanyengo, and A. Akakandelwa, "Mapping Postgraduate Research at the University of Zambia: a review of dissertations for the Master of Medicine Programme." *Medical journal of Zambia*, vol. 37, no. 2, pp. 52–57, 2010.

[11] https://zajlis.unza.zm

[12] http://www.netd.ac.za

[13] http://union.ndltd.org

[14] http://lis.unza.zm/portal

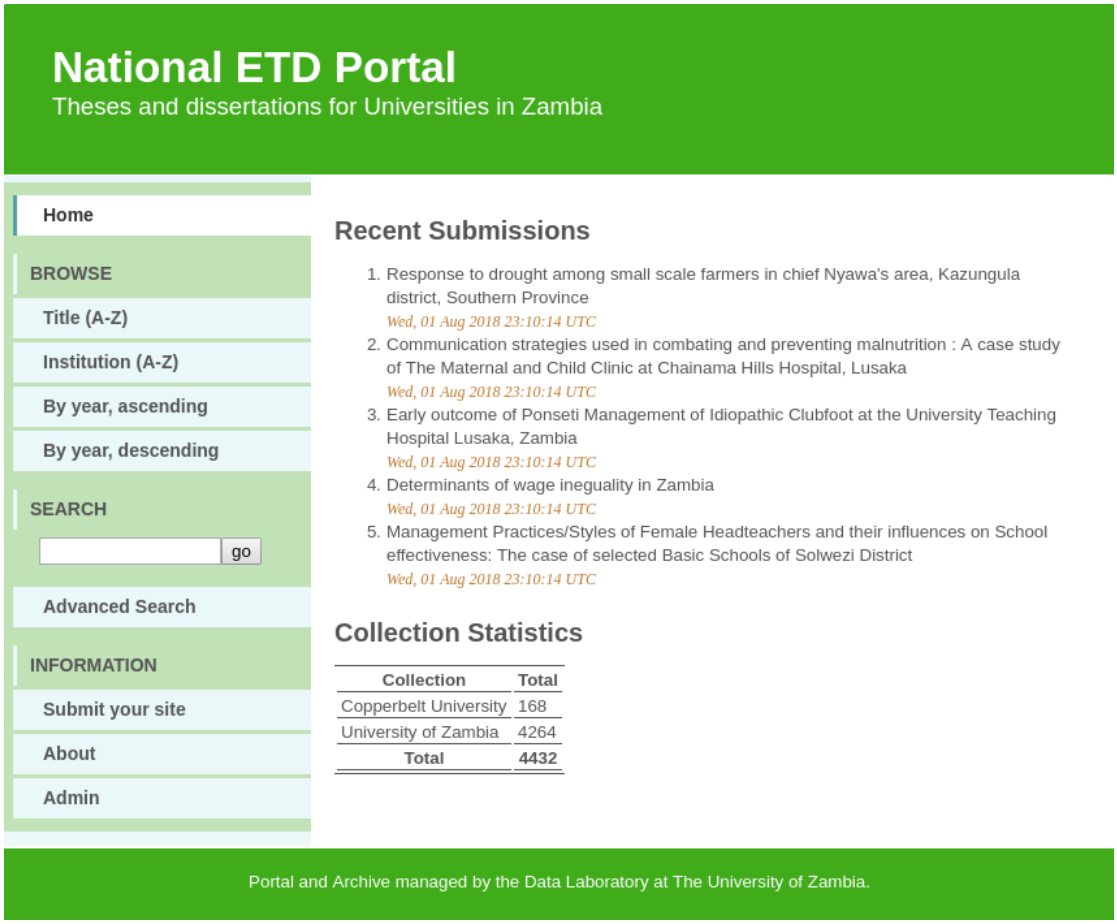



Fig. 5. OAI-PMH Downstream Service for Collecting and Disseminating Electronic Theses and Dissertations in Zambia.


[7] L. Abrahams, M. Burke, and J. Mouton, "Research Productivity-Visibility-Accessibility and Scholarly Communication in Southern African Universities," *The African Journal of Information and Communication*, no. 10, pp. 20–36, 2010.

[8] L. Czerniewicz and K. Wiens, "The online visibility of South African knowledge: Searching for poverty alleviation," *South African Journal of Information and Communication*, no. 13, 2013. [Online]. Available: http://hdl.handle.net/10539/19274

[9] L. Czerniewicz, S. Goodier, and R. Morrell, "Southern knowledge online? Climate change research discoverability and communication practices," *Information, Communication & Society*, vol. 20, no. 3, pp. 386–405, mar 2017. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/1369118X.2016.1168473

[10] H. Trotter, C. Kell, M. Willmers, E. Gray, and T. King, *Seeking Impact and Visibility: Scholarly Communication in Southern Africa*. African Minds, 2014. [Online]. Available: https://open.uct.ac.za/handle/11427/2310

[11] W. Y. Arms, *Digital Libraries*, 2nd ed. Cambridge, Massachusetts: The MIT Press, 2000. [Online]. Available: http://www.cs.cornell.edu/wya/DigLib/

[12] C. Gutteridge, "GNU EPrints 2 Overview," in *11th Panhellenic Academic Libraries Conference*, 2002. [Online]. Available: http://eprints.ecs.soton.ac.uk/6840/

[13] R. Tansley, M. Bass, D. Stuve, M. Branschofsky, D. Chudnov, G. McClellan, and M. Smith, "The DSpace institutional digital repository system: current functionality," in *2003 Joint Conference on Digital Libraries, 2003. Proceedings.* IEEE Comput. Soc, 2003, pp. 87–97. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1204846

[14] C. Lagoze, S. Payette, E. Shin, and C. Wilper, "Fedora: an architecture for complex objects and their relationships," *International Journal on Digital Libraries*, vol. 6, no. 2, pp. 124–138, dec 2005. [Online]. Available: http://www.springerlink.com/index/10.1007/s00799-005-0130-3

[15] J. Willinsky, "Open Journal Systems," *Library Hi Tech*, vol. 23, no. 4, pp. 504–519, dec 2005. [Online]. Available: https://www.emeraldinsight.com/doi/10.1108/07378830510636300

[16] M. Kökörčen\'y and A. Bodnárová, "Comparison of digital libraries systems," in *Proceedings of the 9th WSEAS international conference on Data networks, communications, computers*. World Scientific and Engineering Academy and Society (WSEAS), 2010, pp. 97–100. [Online]. Available: http://www.wseas.us/e-library/conferences/2010/Faro/DNCOCO/DNCOCO-16.pdf

[17] C. Lagoze, H. Van de Sompel, M. Nelson, and S. Warner, "Open Archives Initiative-Protocol for Metadata Harvesting-v. 2.0," 2002. [Online]. Available: http://www.openarchives.org/OAI/openarchivesprotocol.htmlhttp://www.nii.ac.jp/irp/archive/translation/oai-pmh2.0/OpenArchivesProtocol.htm

[18] LibreCat, "LibreCat/Catmandu data processing toolkit." [Online]. Available: http://librecat.org

[19] S. Weibel, J. Kunze, C. Lagoze, and M. Wolf, "Dublin Core Metadata for Resource Discovery," 1998. [Online]. Available: http://www.hjp.at/doc/rfc/rfc2413.html

[20] C. Lagoze, H. Van de Sompel, M. L. Nelson, S. Warner, R. Sanderson, and P. Johnston, "Object Re-Use & Exchange: A Resource-Centric Approach," *0804.2273*, apr 2008. [Online]. Available: http://arxiv.org/abs/0804.2273

[21] The University of Zambia Press, "UNZA Press." [Online]. Available: https://www.unza.zm/units/press

[22] Directorate of Research and Graduate Studies, "Dissemination of Research Findings." [Online]. Available: http://graduate.unza.zm/research/dissemination-of-research-findings

# A Ransomware Classification Framework Based on File-Deletion and File-Encryption Attack Structures

Aaron Zimba
*Department of Computer Science and Information Technology*
*Mulungushi University*
*Kabwe*
*Zambia*
azimba@mu.ac.zm

Mumbi Chishimba
*Department of Information Technology*
*National Institute of Public Administration (NIPA)*
*Lusaka, Zambia*
chishimba.mumbi@gmail.com

Sipiwe Chihana
*Center for Information and Communications Technology*
*Northrise University*
*Ndola, Copperbelt*
*Zambia*
Sipiwechihana09@gmail.com

***Abstract*— Ransomware has emerged as an infamous malware that has not escaped a lot of myths and inaccuracies from media hype. Victims are not sure whether or not to pay a ransom demand without fully understanding the lurking consequences. In this paper, we present a ransomware classification framework based on file-deletion and file-encryption attack structures that provides a deeper comprehension of potential flaws and inadequacies exhibited in ransomware. We formulate a threat and attack model representative of a typical ransomware attack process from which we derive the ransomware categorization framework based on a proposed classification algorithm. The framework classifies the virulence of a ransomware attack to entail the overall effectiveness of potential ways of recovering the attacked data without paying the ransom demand as well as the technical prowess of the underlying attack structures. Results of the categorization, in increasing severity from CAT1 through to CAT5, show that many ransomwares exhibit flaws in their implementation of encryption and deletion attack structures which make data recovery possible without paying the ransom. The most severe categories CAT4 and CAT5 are better mitigated by exploiting encryption essentials while CAT3 can be effectively mitigated via reverse engineering. CAT1 and CAT2 are not common and are easily mitigated without any decryption essentials.**

***Keywords-ransomware; file-deletion; file-encryption; attack structure; data recovery***

## I. INTRODUCTION

Since the invention of the Internet, cyber-crime has continued to grow [1] with attackers employing more innovative ways to attain proceeds of cyber-crime. Since the motivation behind most cyber-crime is monetary gain (excluding cyber espionage and hacktivism), the challenge mainly has been to seamless collect the associated monetary proceeds without a trace. The invention of Bitcoin seems to be a dream come true for cyber criminals due to the anonymity provided by the Bitcoin system [2]. As such, attackers eschewing data exfiltration attacks for less tedious attacks with a high turnover. One such attack is ransomware where the attacker takes hostage of the victim's data without the need to exfiltrate it at all. In a ransomware attack, the attacker uses robust and resilient encryption to make the target data inaccessible without the appropriate decryption keys [3]. Furthermore, the attacker demands a ransom in Bitcoins and usually the victim is left with a binary option of whether to pay or not to. This has seen some victims part away with over a million dollars in a single attack [4]. As such, the ransomware business model is a multi-billion lucrative industry in the cyber-crime landscape which is growing each day [5] with criminal business concepts such as Ransomware-as-a-service [6]. The popularity of ransomware is echoed by Interest Over Time (IOT) as shown in Figure 1 below.

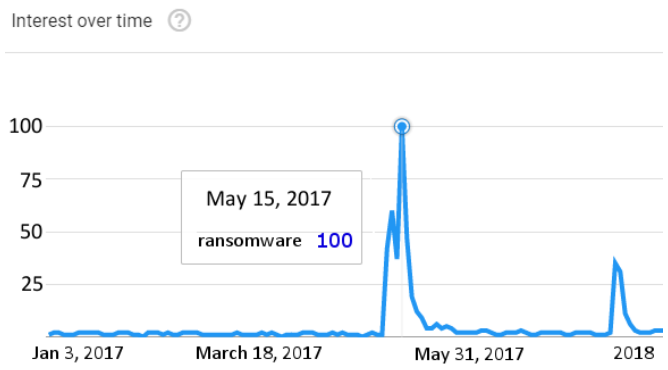



Figure 1. Ransomware attacks IOT. [7]

Sadly, the myths and inaccuracies around ransomware continue to deepen. This has caused victims to make uninformed decisions upon a ransomware attack. Depending on the underlying attack structures, some ransomware attacks can be mitigated and the data recovered without paying the ransom. Unfortunately, some victims have had to pay ransom demands when data could be recovered without honoring the ransom demand [8], as was with the major ransomware attack of 2017 depicted in Figure 1. As such, knowledge of a ransomware's attack structure is vital to the mitigation thereof.

In light of the aforesaid, this paper evaluates attack methodologies of a ransomware attack: the underlying file-deletion and file-encryption attack structures. In the former, we uncover the data recovery-prevention techniques and in the latter, we uncover the associated cryptographic attack models. The deeper comprehension of potential flaws and inadequacies exhibited in these attack structures form the basis of the overall objective. This enables the provision of enough technical information before making a hasty decision to pay a ransom which might result into not only financial loss but loss of access to the attacked files if decryption is not possible by the attacker. We present a threat and attack model which is

representative of a typical ransomware attack process from which we derive the ransomware categorization framework based on a proposed classification algorithm. The framework classifies the virulence of a ransomware attack to entail the overall effectiveness of potential ways of recovering the attacked data without paying the ransom demand as well as the technical prowess of the underlying attack structures.

The rest of the paper is organized as follows: Section II discusses the taxonomy and the threat model with the associated attack structures while Section III presents the proposed classification framework. The methodology and approach are presented in Section IV while classification results and the analysis thereof are brought forth in Section V. The conclusions of the paper are drawn Section VI.

## II. TAXONOMY, THREAT MODEL AND ATTACK STRUCTURES

### A. *Ransomware Attacks taxonomy*

There are several factors that affect the categorization of ransomware attacks. We categorize the attacks based on the following characteristics; target platform, cryptosystem used, severity of loss of data and attack structure. This categorization is not dependent on the underlying infection vectors. The taxonomy is based on the classifications common in the cryptovirology landscape [26]. The result of the categorization is shown in Figure 2 below.

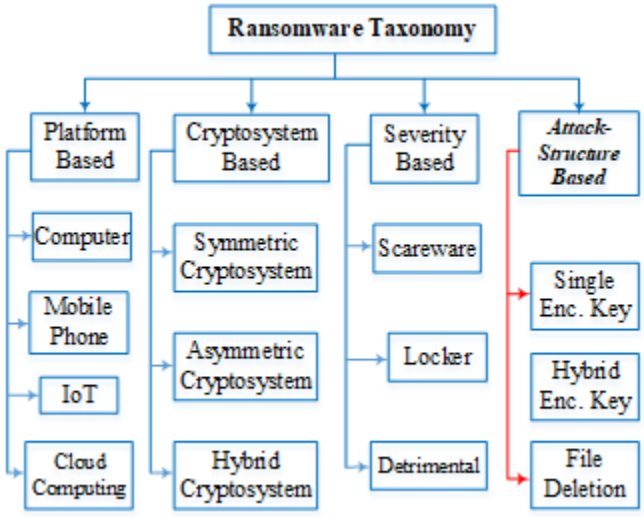



Figure 2. A taxonomy of ransomware attacks

Based on the target platform, ransomware can be made to target computers, mobile phones, Internet of Things devices or cloud computing. Computer based ransomware is the most prevalent owing to wide attack surface and ease of implementation. Android-based ransomware dominates the mobile phone landscape [9] whilst Windows-based IoT devices have been found to be the most susceptible. Cloud computing is an emerging niche for ransomware attacks [10] but the most effective ransomware attack in this domain have been targeted attacks [11] and not indiscriminate. Owing to the disparities in the target platform, we do not use it as a basis for formulating our proposed classification framework. Regarding the cryptosystems used in ransomware today, they can be classified as those that use symmetric, asymmetric or hybrid cryptosystem. In symmetric cryptosystem, the same key is used to encrypt and decrypt the target data. As such the attacker has the classic challenge of secure key management [12]. To overcome the challenge of key management, asymmetric crypto systems are used where the public key is used to encrypt the target data whereas the corresponding private key retained by the attacker is used for decryption. The challenge in asymmetric encryption is that it is slow. Hybrid encryption is adopted as an alternative as it utilizes the speed of symmetric cryptosystems and the resilience of asymmetric cryptosystems. We use some facets of encryption as one of the bases of our framework considering that encryption is at the core of the ransomware business model. In terms of attack structures, ransomware attacks can be categorized into those that use a single key (either a symmetric key or a public), hybrid key (use of both symmetric and public key) and file deletion which either overwrites the original file after encryption or primitively deletes it. We incorporate this characteristic in our framework as it is pivotal to data recovery after an attack. In terms of severity of the damaged caused, scareware type of ransomware doesn't damage or delete the files, it just obfuscates them in one form or another. Locker ransomware usually locks the system login or the boot menu. As such, offline mitigation is effective against this attack category. Detrimental ransomware is one that both encrypts the target files and deletes the remnant original files after encryption.

### B. *Threat Model and Attack Structures*

We now evaluate the threat model and attack structures based on the two selected categories from the taxonomy. The diagram below in Figure 3 depicts the resultant threat model and attack structures.

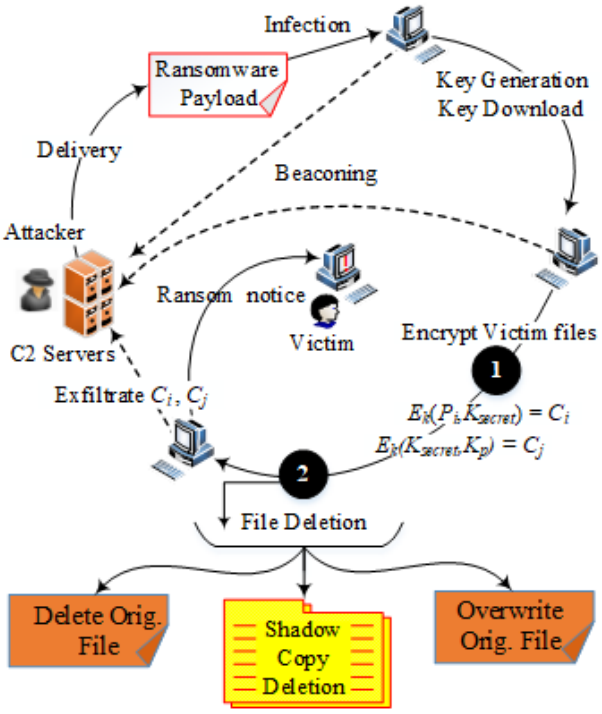



Figure 3. Threat model and attack structures

The model comprises the attacker with command and control (C2) server resources who seeks to victim a host. The C2 might house the ransomware or associated encryption keys depending on the attack model. The attacker uses any of the discussed cryptosystems for encryption and chooses an effective infection vector [13]. Depending on the ransomware variant, the attacker might embed the encryption key in the ransomware payload or the ransomware would have to beacon to the C2 upon infection to download the necessary encryption

keys. To effectuate an effective ransomware attack, the ransomware carries out two major tasks; (1) encrypt the target files and (2) delete the original files after encryption. Encryption of the target files is denoted by $E_k(P_i, K_{secret}) = C_i$. Some attack structures generate a symmetric key using the victim's operating system *CryptoAPI* [14]. After this key completes encrypting the target files, it is further encrypted by the embedded public key which is denoted by $E_k(K_{secret}, K_p) = C_j$. The resultant ciphertext $C_j$ is exfiltrated to the C2 server. In the case of single key attack model, the encryption process is denoted as:

$$\{m_i(target\_payload)\}_{K_{pub}} \rightarrow C_i \quad (1)$$

$$\{m_i(target\_payload)\}_{K_{secret}} \rightarrow C_i \quad (2)$$

Equation (1) is an implementation of an asymmetric cryptosystem whilst Equation (2) a symmetric cryptosystem. In the case of a hybrid key attack model, the encryption attack process is denoted as:

$$\{m_i(data_1)\}_{K_{sym}} \rightarrow C_i \quad (3)$$

$$\{K_{secret}(data_2)\}_{K_{pub}} \rightarrow C_j \quad (4)$$

Equation (3) denotes the first stage of the encryption process where $m_i$ is the plaintext message (targeted files) $data_1$ and $C_i$ is the resultant ciphertext. Equation (4) is the second stage of the encryption process where $data_2$ is the symmetric key $K_{secret}$ used in Equation (3) and $K_{pub}$ is the public key while $C_j$ is the resultant ciphertext.

After completion of the encryption, the ransomware proceeds to delete the remnant files after encryption and the volume shadow copies. The volume shadow copies are usually deleted via vssadmin.exe while the remnant files are either deleted primitively by erasing directories structures and meta-data information of the files. The other way of deleting the files is by overwriting it with random data which corrupts the file and make it unreadable. If the files are deleted via meta-data information and directories structures, they are easily recoverable via third party software and utilities. On the other hand, overwriting the target file with random data makes recovery very difficult.

Considering the above attack structures (file encryption and file deletion), we propose a ransomware classification framework that is based on evaluation of the underlying attack structures. This is helpful because poorly implemented attack structures make recovery of data possible regardless of the resilience of the used crypto-algorithms.

## III. PROPOSED CLASSIFICATION FRAMEWORK

We propose a ransomware classification framework that is based on the attack structures depicted in the threat model in Figure 3 and on the characteristics from the Cryptosystem-Based and Attack Structure-Based categories in the taxonomy in Figure 2. The classification framework is shown in Table 2. We use this categorization framework to formulate a classification algorithm that classifies a ransomware given its attack structures. The algorithm is shown in Figure 4. The abbreviations of the parameters of the algorithm are shown in Table 1. Our framework expresses the severity of a ransomware in terms of file encryption and file deletion. As such, it shows how challenging and time consuming it will be to mitigate a given ransomware attack using the classical methods of static and dynamic analysis [15]. The virulence depicted in the framework is flexible, i.e. a ransomware can move up or down the category list depending on newly discovered properties.

Algorithm 1: Ransomware Classification

```
Input: Encryption & deletion attack structures
Output: Ransomware category
1. if SKc2emb=SKPemb=SKlocalgen=
       HKc2emb=HKPemb=HKlocalgen=no then
2.      malware ← CAT1
3.  else
4.    if delShdCpy=ovrFile=no then
5.       malware ← CAT2
6.    else
7.      if SKc2emb=SKPemb=SKlocalgen=no then
8.         malware ← CAT5
9.      else
10.       if SKc2embsym = SKPembsym =
              SKlocalgensym = yes then
11.          malware ← CAT3
12.       else
13.         malware ← CAT4
14.       end if
15.     end if
16.  end if
17. end if=0
```

Figure 4. Ransomware virulence classification algorithm

**Table 1**. Abbreviations of the parameters of Algorithm 1

| Feature | Term | Code |
|---|---|---|
| Hybrid cryptosystem | C2 download | HKc2emb |
| | Payload embedded | HKPemb |
| | Local key Generation | HKlocalgen |
| Single Key Cryptosystem | C2 download | SKc2emb |
| | Payload embedded | SKPemb |
| | Local key Generation | SKlocalgen |
| Delete Volume Shadow Copies | delShdCpy | |
| Overwrite & Delete Original File | ovrFile | |

Since the framework categorizes the ransomware in ascending order, it is clear that ransomware CAT1 is easier to mitigate than CAT5. As such, CAT5 is the most virulent whilst CAT1 is the least virulent where recovery of data does not require any decryption keys. Since the first sub-category of CAT1 does not implement any of the earlier discussed attack structures, the severity is negligent and it's thus categorized as Scareware as depicted in the taxonomy in Figure 2. Examples of such malware include AnonPop [16].

Other sub-categories in CAT1 implement some of file deletion but not encryption. Thus, recovery of data is possible via third-party software such as Recuva or Photorec [17]. In all these cases, data is recoverable without the need to honor the ransom demand. As such, mitigation measures should never focus on

key retrieval as there's no key in the attack structure. CAT2 ransomware employs only the file encryption attack structures. The key can be download from the C2 or it can come embedded in the payload. This is an example of a poorly implemented ransomware as was the case with Bad Rabbit [18].

**Table 2**. Ransomware Classification Framework

| ***CATEGORY (Severity)*** | **ENCRYPTION ATTACK MODEL** | | | | | | **DELETION ATTACK MODEL** | |
|---|---|---|---|---|---|---|---|---|
| | *Hybrid cryptosystem* | | | *Single Key Cryptosystem* | | | *Delete Volume Shadow Copies* | *Overwrite & Delete Original File* |
| | *C2 download* | *Payload embedded* | *Local key Generation* | *C2 download* | *Payload embedded* | *Local key Generation* | | |
| ***CAT1*** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | *NO* | *NO* |
| | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | *YES* | *NO* |
| | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | *NO* | *YES* |
| | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | *YES* | *YES* |
| ***CAT2*** | ✓‖ ✓‖ ✓ | | | ✗ | | | *NO* | *NO* |
| | ✗ | | | ✓‖ ✓‖ ✓ | | | *NO* | *NO* |
| ***CAT3*** | ✗ | ✗ | ✗ | $\checkmark(K_{enc} = K_{sym})$ | | | *YES* | *YES* |
| ***CAT4*** | ✗ | ✗ | ✗ | ✓‖ ✓‖ ✓ | | | *YES* | *YES* |
| ***CAT5*** | ✓‖ ✓‖ ✓ | | | ✗ | ✗ | ✗ | *YES* | *YES* |

Since there's no deletion of volume shadow copies, data can be recovered via system restore utilities or third-party software. CAT3 represents earlier and uncommon types of ransomware that are based on single key attack structures. In this category, the ransomware comes with an embedded symmetric key in the payload. The key can be simply retrieved using reverse engineering. In the event that the key is deleted from the payload, data deletion recovery techniques discussed in the preceding categories can be used to recover the key. However, if the embedded key is a public key from an asymmetric cryptosystem, it is of no value to extract the public key since it cannot decrypt the data. This is representative of CAT4. Another instance of CAT4 is where the key is downloaded from the C2 server. The key can be symmetric or asymmetric as was with the case of CryptoWall [19]. In the case of the latter, it is very difficult to mitigate the attack since there are no residual encryption essentials on the victim. CAT5 represents the current generation of ransomware. The attack structures implement all the deletion techniques and use hybrid cryptosystems. A typical example is Wannacry which deletes not only the volume shadow copies but the remnant files as well. Further, it comes with an embedded master RSA public key and uses the operating system's CryptoAPI to generate an RSA sub-key pair and AES keys. Each of the unique AES keys is used to encrypt a unique target file. The embedded RSA master public key is used to encrypt the private key from the generated RSA sub-key pair. The public key of the generated RSA sub-key pair is used to encrypt the unique AES keys. As such, to decrypt the data, the victim needs the AES which is encrypted by the private key of the generated RSA sub-key pair. This can only be decrypted by the corresponding generated RSA private sub-key pair, but then, it has been encrypted by the embedded RSA master public key. What the RSA master key has encrypted can only be decrypted by the corresponding RSA master private key. which is in the domain of the attacker. The attacker thus demands a ransom to release the decryption key.

## IV. Methodology and Approach

To evaluate the feasibility and effectiveness of our framework, we analyzed 20 ransomware samples and applied the algorithm of the framework for categorization purposes as shown in Figure 5. We use IDA Pro and Ollydebug for static analysis.

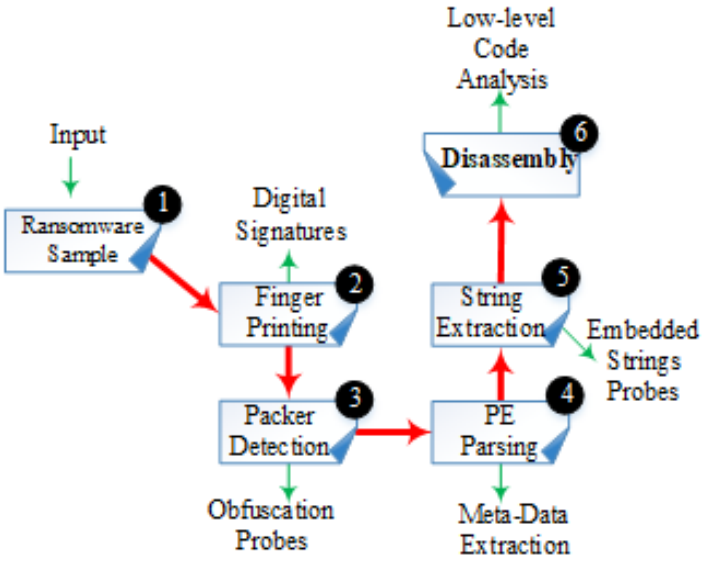



Figure 5. Static analysis workflow

We analyze and extract encryption and deletion features from the ransomware using two approaches: static code analysis and behavioral analysis. The workflow for static analysis is shown in Figure 5. In step 1, we choose the ransomware binaries from reputed malware sources not limited to ReverseIT, VirusTotal and Malware Byte's Malwr. We verify the malware's fingerprints by computing corresponding

cryptographic hashes in step 2. We check whether the malware is obfuscated via packing in step 3 and then parse it for meta-data extraction in step 4. We probe any embedded strings in step 5 and finally analyze the low-level code in step 6. We mainly use IDA Pro and Ollydebug complemented with PEView.

Furthermore, we complement static analysis with behavioral analysis. We add this extra step because there are some ransomware attack-structure features that only appear when the malware is actively executing. We use Cuckoo sandbox as our execution environment. The Cuckoo server runs on Kali Linux whist the virtual hosts through which we actively monitor the malware's activities run Windows on VirtualBox. Though we run the test-bed in *host-only-adapter*, we simulate the Internet by sink-holing it with the *FakeDNS* utility.

By thoroughly running through the processes of static and behavioral analysis, we extract attack structure features which we use as input for the algorithm in Figure 4. We implement the algorithm in the WEKA machine learning workbench. As such, we classify the virulence of the ransomware according to the formulated categories. The results of the classification are presented in the next section.

## V. Classification Results and Analysis

The results from the analyzed samples and their corresponding categories are shown in Table 3 below which we obtained from current ransomware reports [25]. The *Name* column denotes the name of a the particular strain of the malware. We stick to the initial name associated with ransomware and not its subsequent versions. The *Year* depicts the period during which the given sample version existed. It is worth noting that most of the ransomware activities span a couple of years and during this period, newer versions appear with enhanced capabilities. For example, the earliest version of DMA-Locker was seen in 2015 and it used a single symmetric key (AES-256 in ECB mode) to encrypt all the targeted file. And since it deleted remnant files and volume shadow copies, it falls into CAT4. However, in 2016, the enhanced version appeared which used a separate AES key for encryption of each file [24]. Furthermore, the used AES keys were encrypted by an embedded RSA public key. Based on our classification framework, this essentially moves DMA-Locker from CAT4 to CAT5. As such, the *Year* column in table 3 depicts the particular year in which the associated malware variant was seen. The *Paid Ransom* column attaches the monetary value associated with each ransomware campaign. The null entries depict unavailability of verified data. Furthermore, the monetary value associated with each campaign might be more because some ransomware variants are known to use multiple Bitcoin addresses. Some other variants have been known to accept payments in other forms of digital money other than Bitcoins [21] implying that the cumulative value is more than those we traced from BlockChain Info [22]. The *Platform* column denotes the targeted operating system of which 85% represents Windows. The *Category* shows the class in which the corresponding ransomware falls. As can be seen from Table 3, most of the ransomware variants fall into CAT4 and CAT5 accounting for 35% each. CAT1 is not common and is usually the work of script-kiddies whilst CAT2 and CAT3 encompasses early unmatured variants of ransomware such as AIDS [23]. Furthermore, ransomware in CAT4 and CAT5 account for the highest values in paid ransoms. This can be attributed to the difficulty in mitigating such ransomware categories owing to the complex encryption and deletion attack structures.

**Table 3**. Classification of notable ransomware incidents

| Name | Year | Paid Ransoms | Platform | Category |
|---|---|---|---|---|
| AnonPop | 2016 | - | Windows | *CAT1* |
| Cerber | 2016 | > $500,000 | Windows | *CAT5* |
| Bad Rabbit | 2017 | - | Windows | *CAT2* |
| CryptoDefense | 2014 | > $65,000 | Windows | *CAT4* |
| CryptoLocker | 2014 | > $ 3 million | Windows | *CAT4* |
| CryptoWall | 2015 | $18 million | Windows | *CAT4* |
| DMA-Locker | 2015 | > $180,000 | Windows | *CAT4* |
| Jigsaw | 2016 | > $2,000 | Windows | *CAT3* |
| Erebus | 2017 | > $1.04 million | Linux | *CAT5* |
| NotPetya | 2017 | > $10,000 | Windows | *CAT3* |
| KeRanger | 2016 | > $5,000 | Mac OS | *CAT4* |
| Linux.Encoder | 2015 | - | Linux | *CAT3* |
| Locky | 2016 | > $ 1.3 million | Windows | *CAT4* |
| AIDS | 1989 | - | Windows | *CAT3* |
| Petya | 2016 | > $30,000 | Windows | *CAT5* |
| SamSam | 2018 | > $850,000 | Windows | *CAT5* |
| TeslaCrypt | 2015 | > $80,000 | Windows | *CAT4* |
| VenusLocker | 2016 | > $6,500 | Windows | *CAT5* |
| WannaCry | 2017 | > $140,000 | Windows | *CAT5* |
| ZCryptor | 2016 | - | Windows | *CAT5* |

The evolution of ransomware file-deletion and file-encryption characteristics has seen an increment in the emergence of new resilient ransomware mostly falling in CAT5 as depicted in ransomware-attack statistics in Figure 6 [20]. The surge in ransomware attacks represent a 229% increment most of which are CAT4 and CAT5 ransomware.

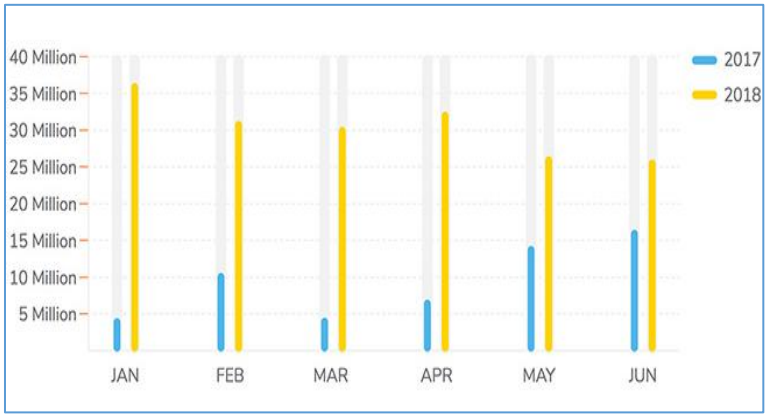



Figure 6. Global ransomware attacks volume for 2017/2018

The changes in ransomware attack structures are echoed in Figure 7 below based on our data-set. Poorly designed

ransomwares are neglected over the years as was the case with CAT2 ransomware which appear only in 2017 and not prior or after. This is a common characteristic erroneously or poorly implemented ransomware variants. The years 2014 – 2016 see a steady appearance of CAT4 ransomware which is followed by a steady appearance of CAT5 ransomware from 2016 – 2018. Other ransomwares are resilient and span several years.

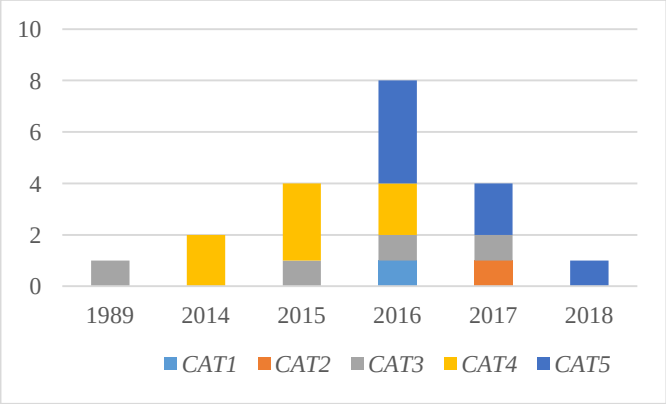



Figure 7. Distribution of ransomware categories over the years

This is the case with CAT3 ransomware except for the AIDS ransomware of 1989. By volume, it is evident from Figure 7 that CAT4 and CAT5 are the most common followed by CAT3. CAT3 and CAT4 ransomware can be mitigated effectively via reverse engineering (static analysis) provided the key used is symmetric. CAT5 can be mitigated if the encryption attack structure uses hybrid encryption essentials from the victim.

## VI. CONCLUSIONS

In this paper, we have presented a ransomware classification framework based on file-deletion and file-encryption attack structures. Based on these two attack structures, we have presented a thorough taxonomy of ransomware attacks which formed the basis of our classification framework. We have defined the threat model and attack structures to characterize a detailed overview of the ransomware attack process not only in terms of encryption but data deletion as well. The threat and attack models are representative of a typical ransomware attack process from which we derived the ransomware categorization framework based on a proposed classification algorithm. The framework classifies the virulence of a ransomware attack to entail the overall effectiveness of potential ways of recovering the attacked data without paying the ransom demand as well as the technical prowess of the underlying attack structures. The algorithm classifies the virulence of a ransomware in increasing severity from CAT1 through to CAT5. Results show that the most recent ransomware strains are CAT4 and CAT5 which are better mitigated by exploiting encryption essentials. CAT1 and CAT2 ransomware are not common in the wild whilst CAT3 and CAT4 ransomware can be mitigated effectively via reverse engineering (static analysis) provided the key used is symmetric. CAT5 can be mitigated if the encryption attack structure uses hybrid encryption essentials from the victim.

## REFERENCES

[1] Hunton P. The growing phenomenon of crime and the internet: A cybercrime execution and analysis model. Computer Law & Security Review. 2009 Nov 1;25(6):528-35.

[2] Khalilov MC, Levi A. A Survey on Anonymity and Privacy in Bitcoin-like Digital Cash Systems. IEEE Communications Surveys & Tutorials. 2018 Mar 26.

[3] Weckstén M, Frick J, Sjöström A, Järpe E. A novel method for recovery from Crypto Ransomware infections. In Computer and Communications (ICCC), 2016 2nd IEEE International Conference on 2016 Oct 14 (pp. 1354-1358). IEEE.

[4] Baek S, Jung Y, Mohaisen A, Lee S, Nyang D. SSD-Insider: Internal Defense of Solid-State Drive against Ransomware with Perfect Data Recovery. In 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS) 2018 Jul 2 (pp. 875-884). IEEE.

[5] Srinivasan CR. Hobby hackers to billion-dollar industry: the evolution of ransomware. Computer Fraud & Security. 2017 Nov 30;2017(11):7-9.

[6] Young AL, Yung M. Cryptovirology: The birth, neglect, and explosion of ransomware. Communications of the ACM. 2017 Jun 26;60(7):24-6.

[7] Google Trends Explore. Ransomware. (2018). [Online] Available: https://trends.google.com/trends/explore?q=ransomware

[8] Mukesh SD. An Analysis Technique to Detect Ransomware Threat. In 2018 International Conference on Computer Communication and Informatics (ICCCI) 2018 Jan 4 (pp. 1-5). IEEE.

[9] Andronio N, Zanero S, Maggi F. Heldroid: Dissecting and detecting mobile ransomware. I nInternational Workshop on Recent Advances in Intrusion Detection 2015 Nov 2 (pp. 382-404). Springer, Cham.

[10] Bhattacharya S, Kumar CR. Ransomware: The CryptoVirus subverting cloud security. InAlgorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET), 2017 International Conference on 2017 Feb 16 (pp. 1-6). IEEE.

[11] Chaurasia R. Ransomware: The Cyber Extortionist. In Handbook of Research on Information and Cyber Security in the Fourth Industrial Revolution 2018 (pp. 64-111). IGI Global.

[12] Chandramouli R, Iorga M, Chokhani S. Cryptographic key management issues and challenges in cloud services. InSecure Cloud Computing 2014 (pp. 1-30). Springer, New York, NY.

[13] Zimba A, Wang Z, Chen H. Reasoning crypto ransomware infection vectors with Bayesian networks. InIntelligence and Security Informatics (ISI), 2017 IEEE International Conference on 2017 Jul 22 (pp. 149-151). IEEE.

[14] Palisse A, Le Bouder H, Lanet JL, Le Guernic C, Legay A. Ransomware and the legacy crypto API. InInternational Conference on Risks and Security of Internet and Systems 2016 Sep 5 (pp. 11-28). Springer, Cham.

[15] Tzermias Z, Sykiotakis G, Polychronakis M, Markatos EP. Combining static and dynamic analysis for the detection of malicious documents. InProceedings of the Fourth European Workshop on System Security 2011 Apr 10 (p. 4). ACM..

[16] Bajpai P, Sood AK, Enbody R. A key-management-based taxonomy for ransomware. In2018 APWG Symposium on Electronic Crime Research (eCrime) 2018 May 15 (pp. 1-12). IEEE.

[17] Emm D. Cracking the code: The history of Gpcode. Computer Fraud & Security. 2008 Sep 1;2008(9):15-7.

[18] Mohanta A, Hahad M, Velmurugan K. Preventing Ransomware: Understand, prevent, and remediate ransomware attacks. Packt Publishing. 2018.

[19] Cabaj K, Gawkowski P, Grochowski K, Osojca D. Network activity analysis of CryptoWall ransomware. Przeglad Elektrotechniczny. 2015;91(11):201-4.

[20] Ransomware back in big way, 181.5 million attacks since January. (July 13, 2018). [Online] Available: http://vinransomware.com/latest-news/ransomware-back-in-big-way-181-5-million-attacks-since-january.

[21] Huang DY, Aliapoulios MM, Li VG, Invernizzi L, Bursztein E, McRoberts K, Levin J, Levchenko K, Snoeren AC, McCoy D. Tracking ransomware end-to-end. In2018 IEEE Symposium on Security and Privacy (SP) 2018 May 20 (pp. 618-631). IEEE.

[22] Bitcoin Charts & GraphsBlockchain. (2018). [Online] Available: https://www.blockchain.com/charts

[23] Gazet A. Comparative analysis of various ransomware virii. Journal in computer virology. Springer-Verlag. 2010 Feb 1;6Vol. (1): pp. 77-90.

[24] Conti M, Gangwal A, Ruj S. On the Economic Significance of Ransomware Campaigns: A Bitcoin Transactions Perspective. arXiv preprint arXiv:1804.01341. 2018 Apr 4.

[25] A History of Ransomware Attacks: The Biggest and Worst Ransomware Attacks of All Time. [Online] Available: https://digitalguardian.com/blog/history-ransomware-attacks-biggest-and-worst-ransomware-attacks-all-time#4

[26] Al-rimy BA, Maarof MA, Shaid SZ. Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions. Computers & Security. 2018 Jan 10.

# *Closed-loop Current Control of a Five-level Grid-connected Inverter*

Francis Kafata
franciskafata@gmail.com

Francis Mulolani
fmulolani@cbu.ac.zm

Esau Zulu
esau.zulu@cbu.ac.zm

Copperbelt University, School of Engineering, Electrical Department,
PO Box 21692, Kitwe, Zambia.

***Abstract* — This paper presents a closed-loop current control scheme for a 5-level 3-phase diode-clamped multilevel inverter system. The proposed closed loop current control technique is based on the proportional integral (PI) controller theory, and the modulation technique used is level-shift-carrier sinusoidal pulse width modulation (SPWM). The gain values of PI controller were selected on the basis of trial and error to achieve a good compensating signal. Grid synchronization was achieved by using a phase-locked loop. Matlab/Simulink software has been used to run the simulations. The simulation results show that a 1.17% total harmonic distortion (THD) of the output current was attained.**

**Keywords — Closed Loop current Control, Level-Shift-Carrier-SPWM, Diode-Clamped Multilevel Inverter, Proportional Integral controller (PI). Introduction (*Heading 1*)**

## I. INTRODUCTION

Multilevel inverters are power electronics converters that convert direct current (DC) power to alternating current (AC) power at the required output voltage and frequency. The output current of multilevel inverter is compared with reference waveform and the error is used by the PI controller to generate the gating signals. Multilevel inverters topologies have some advantages over the conventional two level inverters such as increased power ratings, improved harmonic performance, and reduced electromagnetic interference (EMI) emission [1].

These topologies have also the characteristic of increasing the output voltage without increasing the voltage rating of switching devices; as a result they are used for direct connection of Photovoltaic (PV) systems to the grid system without the use of transformers hence reducing the cost. They can also be applied in medium- or even low-power/voltage applications apart from the high power/voltage because they allow operations with lower voltage-rated devices, which is potentially a better performance/economical features [2].

The diode clamped multilevel inverter, is one of the most commonly used multilevel inverter in which the diodes are used as the clamping devices to clamp the dc bus voltage so as to achieve steps in the output voltage. It also uses capacitors in series to divide up the dc bus voltage into a set of voltage levels. Therefore, to produce n-levels of the phase voltage, an n-level diode-clamp multilevel inverter needs n-1 capacitors on the dc bus. It also requires 2(n-1) main switches and 2(n-1) diodes on the main switches while the clamping diodes are 2(n-2) [8]. In general the voltage across each capacitor for an n-level diode clamped inverter at steady state is $V_{dc}/(n-1)$. Thus, each active switching device is only required to block a voltage of $V_{dc}/(n-1)$, where $n$ the number of levels.

The performance of the diode clamped multilevel inverter is also determined by the modulation and control schemes used. The carrier based pulse width modulation (PWM) schemes are classified into two categories, namely; (i) Phase shifted carrier modulation (ii) Level shifted carrier modulation [3]. The level shifted modulation has a lower Total Harmonic Distortion (THD) compared to phase shifted carrier modulation. However, the level shifted carrier modulation is mostly used due to its flexibility and simplicity in implementation than the Phase shifted carrier modulation [4].

This paper presents a closed-loop current control scheme for a five-level three-phase diode-clamped dc–ac inverter using the level-shift carrier sinusoidal pulse width modulation (SPWM) based on proportional integral (PI) controller. The current control has been performed by the comparison of the reference currents with the actual measured current. The phase locked loop is implemented for synchronisation. Simulation results from Matlab/Simulink are presented to demonstrate the operation of the five level diode clamped multilevel inverter system.

This paper is arranged or organised as follows: In section II, the overview of the grid connected system is given, in section III, the details of the control scheme are given and explained, in section IV the method used is given, while in section V, the simulation results are presented and discussed, and section VI gives the conclusion.

## II. SYSTEM OVERVIEW

The single line diagram for a Grid connected five level diode clamped multilevel inverter with a DC-Link is shown in Fig. 1.

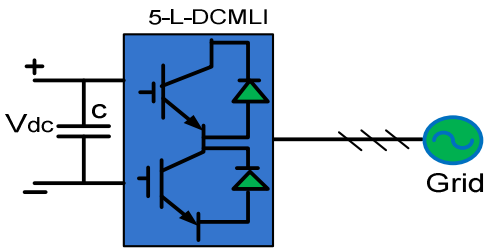



Fig.1: Single line diagram for a grid connected five Level-DCMLI

The detailed circuit diagram for a 5 level diode clamped multilevel inverter is shown in Fig.2. For simplicity and easy reference only one leg has been shown. The circuit consists of eight Insulated Gate Bipolar Transistors, (IGBT), six clamping diodes and four capacitors. This topology was proposed by Nabae, Takahashi, and Akagi in 1981 and it is one of the mostly used topology for grid connected photovoltaic systems [5].

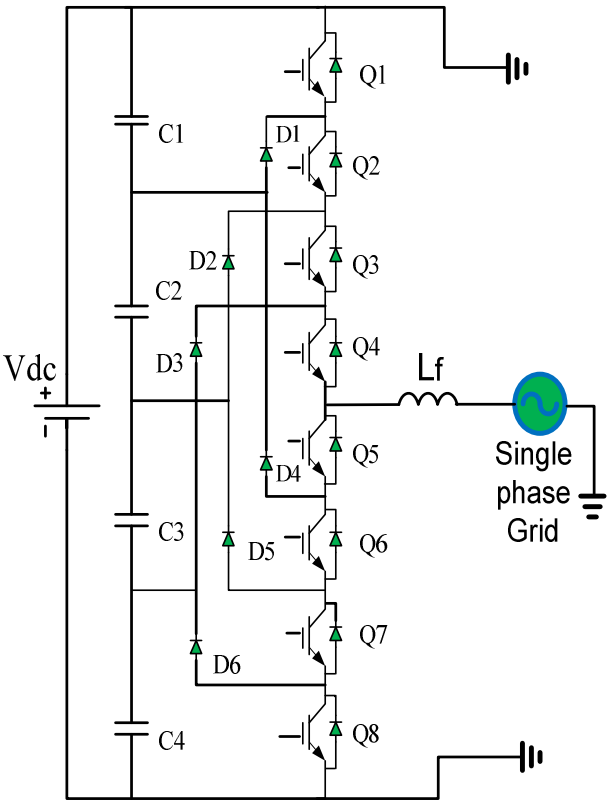



Fig.2. Detailed circuit of a Single Phase five level Diode Clamped Multilevel Inverter Grid connected

## III. CONTROL SCHEME

The five level three phase diode clamped multi-level inverter is used for grid connection. To optimise the multi-level efficiency and operation, various schemes of modulation and control techniques have been employed or implemented. Pulse Width Modulation (PWM) is one of the most popular methods of controlling the output voltage by adjusting the On/Off period of power inverter switches. The Pulse width modulation (PWM) techniques requires a sinusoidal reference signal and triangular carrier signal to generate the required modulating signals to drive the inverter for the desired output [6]. The most common PWM technique for diode clamped multilevel inverter is the sinusoidal pulse width modulation and in this paper the level-shift-carrier sinusoidal pulse width modulation (LS-SPWM) has been applied. The output of the inverter is synchronized to the grid supply by the application of the Phase Locked Loop (PLL). The control system of the five level three phase diode clamped grid connected is shown in Fig.5.

### *A. Phase Locked Loop (PLL)*

The PLL technique has been implemented in this paper for grid synchronization due to its robustness in various grid conditions. It is simple to implement, more effective, and it is reliable [7].

The major role played by the PLL is to provide Grid synchronization which is a very important function in grid connected inverters. The Phase Locked Loop (PLL) will detect or (sense) the frequency and phase (phase angle) from the grid system, where the phase angle is used as the reference for the two-phase transformation from abc to dq. The PLL has three major components namely; phase detector (PD), loop filter, and a voltage controlled oscillator (VCO) [8]. PLL block diagram is shown in Fig.3.

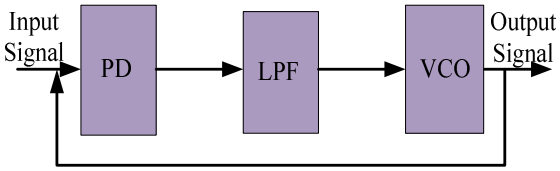



Fig.3: Block diagram of Phase Locked Loop

### *B. Proportional Integral (PI)Controller*

The PI controller has been widely used as current controller in variety of inverter applications because the design and implementation of this controller are quite simple, low computational burden of this control algorithm makes it also easier to implement the whole control system and also its stability makes it more useful as a current controller in most inverter topologies [9], [10], [11]. The proportional integral (PI) current control technique is employed for grid connected inverter to keep the output current sinusoidal and to have high dynamic performance under rapidly changing atmospheric conditions and to maintain the power factor at near unity [12]. The transfer function for the PI controller is given in (1).

$$G_{PI} = K_P + \frac{K_I}{S} \qquad (1)$$

### *C. Output L- Filter*

The L-filter has been implemented to reduce the ripple in the output current. This topology has the fewer number of components among the filter topologies [13]. The other two filter topologies are LC-filter and the LCL-filter. The most critical objective of the output filter is to achieve a reduced ripple in the output current. Fig. 4 shows the L-filter where $V_{inv}$ is the inverter voltage, $L_f$ is the filter inductor and $V_g$ is the grid voltage.

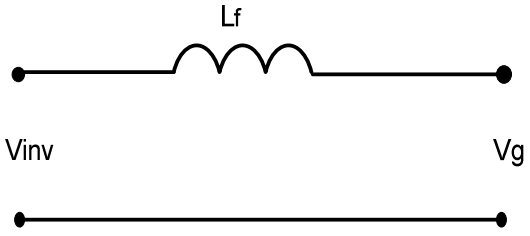



Fig.4 L-filter

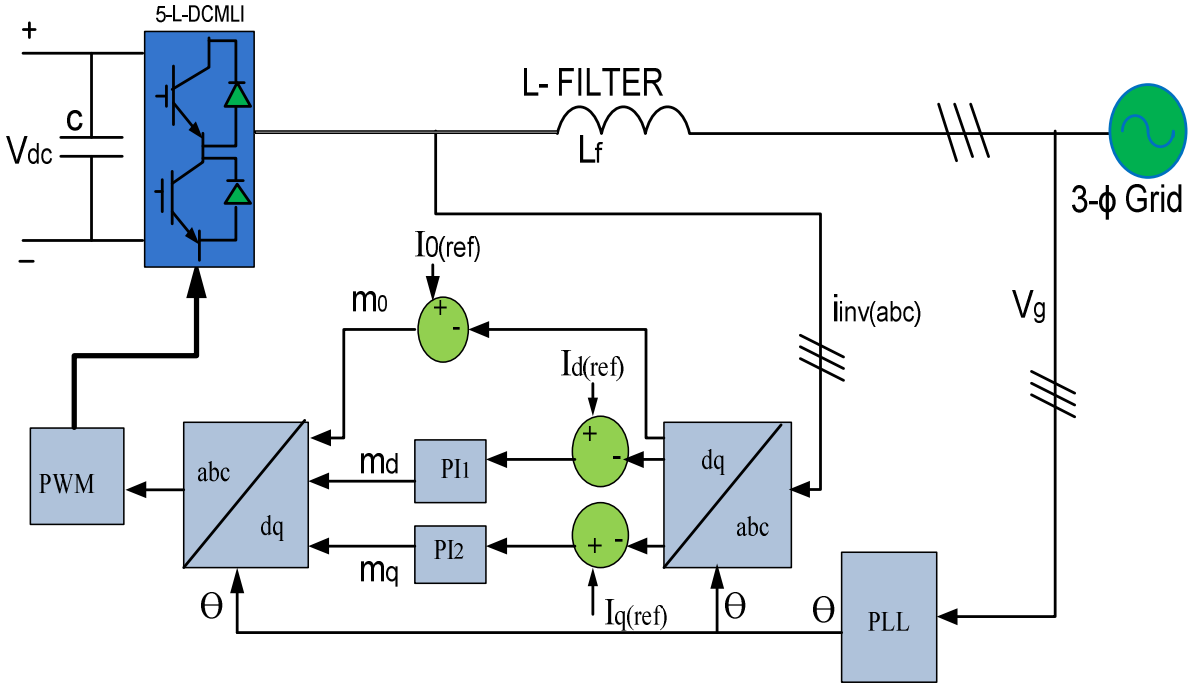



Fig.5: Closed loop current Control scheme of the five Level 3-Phase Diode Clamped Multilevel Inverter

The inverter is connected to the grid through an L-filter. The inverter voltage is related to the grid voltage in (2).

$$v_{inv}^{abc} = v_g^{abc} + R_g i^{abc} + (L_g + L)\frac{d}{dt} i^{abc} \qquad (2)$$

Where inverter output voltage space vector is $v_{inv}^{abc}$, the grid voltage space vector is $v_g^{abc}$, and the output current space vector is $i^{abc}$. The harmonic attenuation capability of the filter can be increased by increasing the inductance value for a certain frequency.

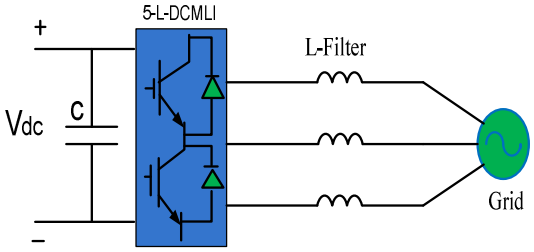



Fig.6: Grid connected five level 3-phase Diode Clamped Multilevel Inverter

The equivalent circuit is shown in Fig.7.

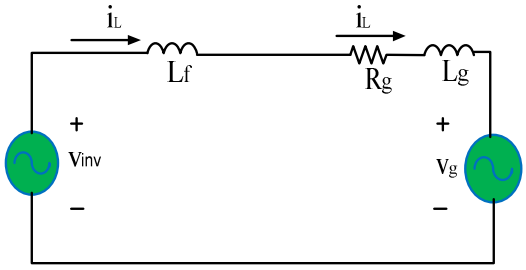



Fig.7. Single phase equivalent circuit

## IV. METHODOLOGY

The Five Level-Three-Phase Diode-Clamped Multilevel Inverter and its control scheme are modelled in MATLAB/Simulink. The DC capacitor is selected so that it minimizes the ripple in the DC voltage and maintains the voltage during transient operation while the parameters of the L-filter are selected to reduce the inverter current ripple, and also the grid current harmonics. Table 1 gives the parameters of the modelled system.

A number of simulations are run in MATLAB/Simulink using the values shown in the table while varying the PI control parameters until the best output voltage waveform is attained.

The harmonic content in the output waveforms is also closely monitored so that the optimal values are obtained while the control parameters are varying.

TABLE I. PARAMETERS USED IN THE MODEL FOR SIMULATION

| Parameter | Value |
|---|---|
| DC voltage, $V_{dc}$ | 800 V |
| Grid voltage, $V_g$ | 400 V |
| Grid frequency, $f_0$ | 50 Hz |
| DC capacitor, $C_{dc}$ | 100 µF |
| Switching frequency, $f_{sw}$ | 2 kHz |
| Filter inductance, $L_f$ | 0.2 mH |
| Load inductance, $L_L$ | 10 mH |
| Load resistance, $R_L$ | 5 Ω |

## V. SIMULATION RESULTS

The simulation results are presented and discussed in this section. The grid three phase voltages are shown in Fig. 8 below. From Fig. 8 it can be shown that in steady state the grid voltage is stable.

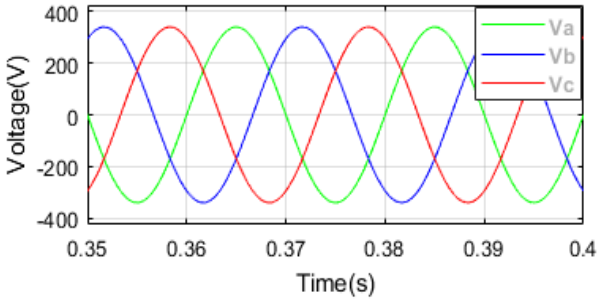



Fig.8: Grid voltage (V).

The 5 level 3 phase diode clamped multilevel inverter output voltage is shown in Fig. 9 below. In Fig.9, the main contribution is to show the 5 level inverter waveform with respect to the operation of the multilevel inverter. Hence from the waveform it shows that the 5 level inverter was successfully designed.

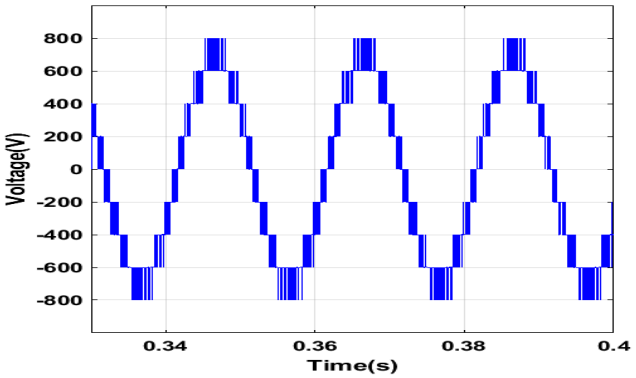



Fig.9: Inverter output voltage (V).

The performance of the closed loop current control 5 level diode clamped multilevel inverter based on proportional integral (PI) controller is shown in Fig.10. From Fig.10, the current waveforms show stability in steady state. Hence the system shows a good performance under grid connection.

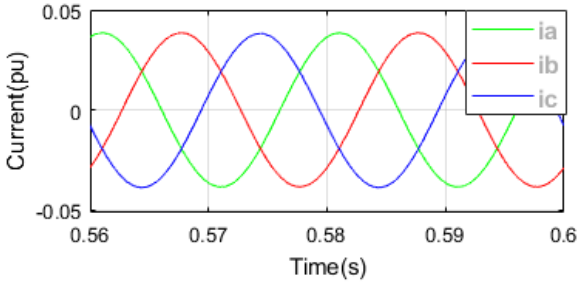



Fig.10: Output current waveform

Fig.11 shows the response of the measured direct current (Id) to the reference current (Iref.). It shows that the measured direct current, (Id) was able to follow the reference. The waveforms show that the PI controller was successfully tuned.

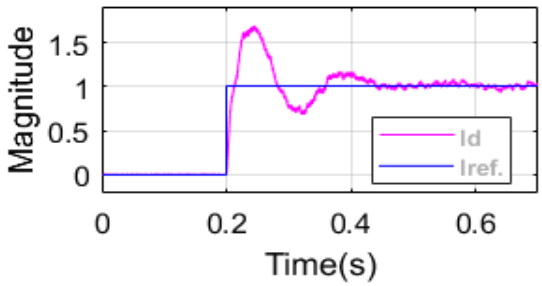



Fig.11: Measured direct current, (Id) and reference current (Iref.)

Fig.12 shows the response for measured quadrature current, (Iq) and reference. From Fig.12, the results show the performance of the PI controller for the reference, (Iref) and the quadrature current, (Iq) which shows also that the tuning of the controller was successful as the measured current was able to follow the reference achieving the error of zero.

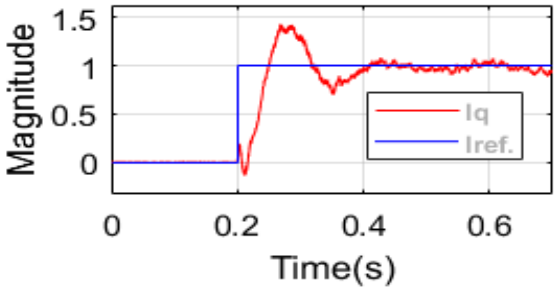



Fig.12: Response of measured quadrature current, (Iq) and reference, (Iref.)

The dynamic response of the output current is shown in Fig. 13. The response of the output current shows that the system stabilizes within a short period after the step response. The system performance under dynamic operation performed as expected when the system is stable.

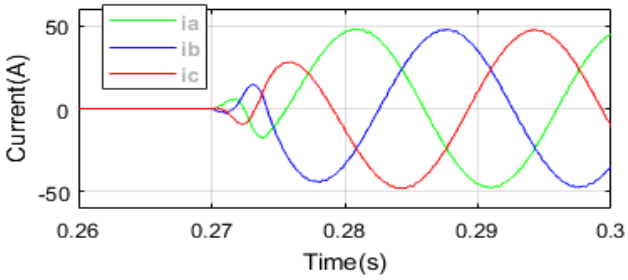
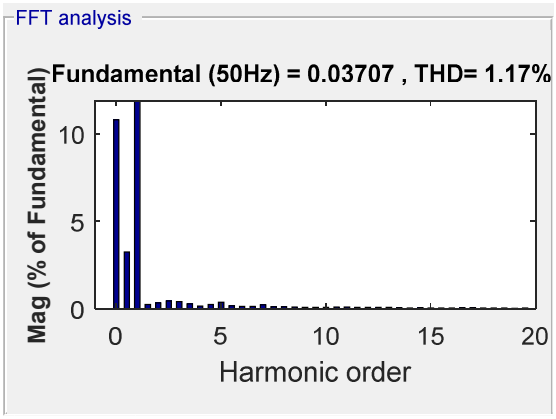



Fig.13: Dynamic response of output current

The percentage total harmonic distortion (%THD) for the output current is shown in Fig.14 with a 1.17% THD. The 1.17% THD results are within the compliance of the IEEE for the allowable current harmonics injected into the grid [14].

## VI. CONCLUSION

The closed loop current control of the 5 level 3 phase diode clamped multilevel inverter grid connected based on PI controller has been presented with the level-shift-carrier sinusoidal pulse width modulation (SPWM).




A controller has been designed successfully for closed loop current control based on PI controller such that the system output current performance and the overall system under steady state and dynamic response shows better system performance with 1.17% total harmonic distortion (THD) . The gain values of PI controller were selected on the basis of trial and error to achieve a good compensating signal. Grid synchronization is achieved using a phase-locked loop. Matlab/Simulink software was used to run the model. The simulation results show that the output current remained stable in both steady state and dynamic responses with a better THD of 1.175% which is within the compliance requirement of IEEE for grid connection of multilevel inverter.

## REFERENCES

[1] N.Das Jaya Raj, P. Bala Krishna and K. Manoz Kumar Reddy, "Analysis of Different Multilevel Converter Topologies for Photovoltaic Applications", International Journal for Modern Trends in Science and Technology, Volume: 03, Issue No: 09, September 2017.

[2] Sergio Busquets-Monge, José Daniel Ortega, Josep Bordonau, José Antonio Beristáin, and Joan Rocabert, "Closed-Loop Control of a Three-Phase Neutral-Point- Clamped Inverter Using an Optimized Virtual-Vector-Based Pulse width Modulation". IEEE Transactions on Industrial Electronics, VOL. 55, No. 5, May 2008.

[3] M. Sudhakaran and R. Seyezhai "A Review of Various Carrier Pwm Techniques for Trinary Cascaded Multilevel Inverter" Middle-East J. Sci. Res., 24 (Recent Innovations in Engineering, Technology, Management & Applications): 81-89, 2016.

[4] Mohit Jain et al. Comparative Analysis and Simulation of Diode Clamped & Cascaded H-Bridge Multilevel Inverter using SPWM Technique". International Journal of Engineering Research and Applications, Vol. 5, Issue 1(Part5), January 2015, pp.95-102.

[5] Bimal K. Bose, "Multi-Level Converters". Electronics 2015, 4, 582-585.

[6] G.Pavana Jyothi, P and PM.Bhagya Lakshmi, "Implementation of New PWM Method for Diode Clamped Multilevel Inverter", International Journal of Scientific Engineering and Applied Science (IJSEAS) - Volume-1, Issue-7,October 2015.

[7] Mostafa Ahmadzadeh, Saeedollah Mortazavi and Mohsen Saniei, "Applying the Taguchi Method for Investigating the Phase-Locked Loop Dynamics Affected by Hybrid Storage System Parameters", Energies 2018, 11, 226; doi:10.3390/en11010226.

[8] Golestan, S.; Monfared, M.; Freijedo, F.D.; Guerrero, J.M. Design and tuning of a modified power-based PLL for single-phase grid-connected power conditioning systems. IEEE Trans. Power Electron. 2012, 27, 3639–3650.

[9] K. Smriti Rao and Ravi Mishra, Comparative study of P, PI and PID controller for speed control of VSI-fed induction motor". International Journal of Engineering Development and Research, Volume 2, Issue 2, 2014.

[10] Bouzid, A.M.; Guerrero, J.M.; Cheriti, A.; Bouhamida, M.; Sicard, P.; Benghanem, M. A survey on control of electric power distributed generation systems for microgrid applications. Renew. Sustain. Energy Rev. 2015,44, 751–766.

[11] Ngoc Bao Lai and Kyeong-Hwa Kim, "An Improved Current Control Strategy for a Grid-Connected Inverter under Distorted Grid Conditions". Energies 2016, 9, 190; doi:10.3390/en9030190.

[12] Jeyraj Selvaraj and Nasrudin A. Rahim, "Multilevel Inverter For Grid-Connected PV System Employing Digital PI Controller", IEEE Transactions on Industrial Electronics, Vol. 56, No. 1, January 2009.

[13] Samuel Vasconcelos Araújo, Alfred Engler, Benjamin Sahan and Fernando Luiz Marcel Antunes, "LCL Filter design for grid-connected NPC inverters in offshore wind turbines", The 7thInternational Conference on Power Electronics, October 22-26, 2007 / EXCO, Daegu, Korea.

[14] A. Bhowmik, A. Maitra, S.M. Halpin, J.E. Schatz, "Determinationof allowable penetration levels of distributed generation resources based on harmonic limit considerations," IEEE Trans Power Del.,vol.18, no.2, pp. 619- 624, April 2003.

# Addressing Energy Consumption Problem Using Wi-Fi At A Care Home For The Aged In Zambia

Justine Chilongu[1], Lusungu Ndovi[2]
Department of Electrical Engineering
School of Engineering
Copperbelt University
Kitwe, Zambia
[1]chilonguj@yahoo.com, [2]lusungu.ndovi@gmail.com

Josephat Kalezhi
Department of Computer Engineering
School of Information and Communication Technology
Copperbelt University
Kitwe, Zambia
kalezhi@cbu.ac.zm

***Abstract*— The development of smart homes for the aged people has over the recent years come as a result of the development of electronics and Information Communication Technologies (ICTs) in which electronic devices have played a very important role of enabling both the monitoring and controlling of electrical devices. Ordinarily, the existing systems do not allow a user to get a feasibility to actively mitigate the power consumption of home equipment. In the past, the electric home equipment could not be easily controlled and monitored resulting into having huge energy consumption costs. However, at the moment, a wireless Sensor Network (WSN) technology is being used for the controlling and monitoring of electric home equipment far more than in the past. This paper describes the proposed methodology of a Wi-Fi based home automation system in which two things are considered. The first one is energy consumption and the other is energy control at Mitanda Home for the aged in Zambia. In this work, Wi-Fi is used for monitoring energy consumption of home equipment. Then home sensor collects the energy consumption data and analyzes them for energy approximation and control the home energy utilization schedule to slump the energy cost. Then energy data of home servers evaluates them, and generates useful statistical examination information aggregated by the remote energy management server. The system is intended to be a more efficient means of energy saving and result in home energy cost reduction in the care homes for the aged people in Zambia. The system will add value to Zambian by promoting efficient electrical energy management.**

***Index Terms*-- Smart Home, Remote Control, Sensor, Wi-Fi Module.**

## I. INTRODUCTION

These days there is a rise in the need for monitoring and controlling environmental parameters almost in all old peoples' homes because they do not regularly control the home electrical devices such as bulbs, fans, heaters etc. Today's technological advancements in electronics and the Internet of Things (IoT) can therefore be deployed as means of reducing the costs that arise from the failure to electrical devices in older people's homes. This could be achieved by monitoring and controlling using environment monitoring system. However there are challenges with some of the measuring components. There is therefore a need of an environment monitoring system which is accurate, easily operated, simple in working, uncomplicated construction, cost-effective, comfortable to carry and lightweight [1]. The aim of this research is to design and develop a system which fulfills these requirements.

The rest of the paper is structured as follows: section II introduced the problem. Section III reviews the related literature. The methodology appears in section VI. The proposed system is provided in section V. The results are provided in section VI. Finally the conclusion appears in section VII.

## II. PROBLEM STATEMENT

Mitanda home for the aged doesn't have the means of monitoring and controlling the electrical devices and this leads to increased cost of electricity. According to Mitanda home management, the latest energy bill for December 2017 is over fourteen Thousand Kwacha, (K14 000 or 1 400 US Dollars which according to the Zambian economy is very high. Mitanda home of the aged in Ndola does not have a means of monitoring and controlling home appliances used by the inhabitants [2]. This leads to increased electricity usage and increased energy costs. In order to lessen the energy consumption problem, this research proposes the adoption of information and communications technologies (ICTs), Internet of Things (IOT) [3] to address the energy consumption problem Home for the aged.

The questions that need to be answered are:

- What are the electrical energy consumption needs of care homes of the aged people in Zambia which when they are addressed will improve their lives?
- What are the relevant technologies in existence today that will enable us to automate the home of the aged people?
- How long can an automated home be sustained?

According the Zambian energy profile, the Global Village Energy Partnership (GVEP), in association with the department of energy and ministry of energy and water development have been working on a mechanism to increase access to reliable, affordable and environmentally sustainable energy services as a means of enhancing economic and social development [16]. This means that the Zambian energy profile has got an economic and social impact on old people's homes also.

In this work, the researchers intend to identify the energy monitoring and controlling technologies in existence that can be a solution to the energy consumption problem at Mitanda home by a way of monitoring and controlling the appliances used by the elderly residents.

The objective of this work is to measure amount of energy consumed when there's no automation involved and the energy consumed when using automation. A comparison of the two situations helps to determine how much energy is wasted when there's no control. The measurement is done by the use of the existing energy measuring equipment.

In the past two approaches to the energy management have been applied, namely smart grid and Home networks both of which focus on the energy management and control of the devices [4]. The proposed research aims to deploy wireless sensor networks to monitor and control the devices.

## III. LITERATURE REVIEW

Elderly people are an important and growing segment in the world population. The statistics show that the percentage of older people is continuously growing due to many reasons, in particular, the declining of birth rates and the reduction of women fertility. According to Quynh et al. [5], in Australia with increased life expectancy, the proportion of population of 65 years and older is expected to more than double by 2050 with the greatest rate of growth of 85+ group [5] A United Nations report [6] estimated that the life expectancy was 65 years in 1950, and 78 years in 2010 and it will continue to rise to 83 years in 2045 [6]. Zambia's population projection in the sub Saharan region is as shown in the table 1 below:

TABLE I. PERCENTAGE OF POPULATION IN ORDER OF AGES BY REGION, 2008, 2020, 2040 [7]

| Region | 65 years and over | 75 years and over | 80 years and over |
|---|---|---|---|
| **Sub Saharan Africa** | | | |
| 2008 | 3.0 | 0.9 | 0.3 |
| 2020 | 3.3 | 1.0 | 0.4 |
| 2040 | 4.2 | 1.4 | 0.6 |

A United Nations report on Zambia further estimated that the population by age group of those above 60 years was 4.4 % in 1950 and rise to 6.7 in 2050[6].

There are few old people's homes in Zambia among which include Mitanda Home of the aged in Ndola, Chibolya in Mufulira, Mwandi and Maramba in Livingstone [8].

The Salvation Army has over the years partnered with the Government Republic of Zambia in providing various social services. With its motto of 'Hand to Man and Heart to God' the Salvation Army has touched many lives and continues to support vulnerable communities in Zambia one of which is Mitanda Home of the aged, located in the heart of Ndola City on the Copperbelt. The Salvation Army, in partnership with the Government Republic of Zambia has provided a Home for vulnerable Senior Citizens. Senior citizens are one of the most vulnerable groups in our Zambian society. They have limited regenerative abilities and are in the majority of cases victims of poverty, disease and homelessness. The HIV/AIDS pandemic has added to their plight. Most Senior Citizens are left without much needed family support systems on account of most young people dying early [2].

Running a Home for the aged comes with many challenges. On a monthly basis, It costs a lot of funding. Each resident requires good health care, good accommodation, meals and social amenities. There's a clinic but the clinic faces challenges in meeting the medical needs and the resources are scarce [2]. The obvious electrical energy consumption need in the Zambian care homes for the aged such as Mitanda home in Ndola is the lack of a systematic way of monitoring and controlling the energy consuming devices. This research has revealed that Mitanda home has got no means of monitoring and controlling the devices and that that the energy bills are a very big source of concern. This research seeks to address this problem in order to reduce the energy consumption costs.

The electrical equipment includes Stoves, TV sets, Fridges, laundry machines, fans, irons etc. Although the environment at Mitanda is clean, maintaining hygiene at the old people's Home is a demanding responsibility both financially and materially. For instance, the care takers have to constantly wash clothes and blankets for the old people. The constant washing needs laundry machines that are in good condition. The current laundry machines have outlived their lifespan and experience constant break down [2]. The Zambian family arrangement strongly believes in supporting the elderly people within their respective communities. However, situations arise in which some old People are denied that support. This is the reason for the establishment of homes like Maramba in Livingstone in the southern Province of the republic of Zambia. Over the years, the Department of Social Welfare saw the need for a community participation that will assist in the effective operations of the Home for improved service delivery to older persons in the institution. As such the community participation is needed [8]. Therefore, it is clear that old people' homes in Zambia face similar challenges.

Smart homes, intelligent homes, home automation, domestics and others are all synonyms that describe, according to the Smart Home Automation of Netherlands, the "integration of technology and services through home networking for a better quality of living" [9]. The UK Department of Trade and Industry defines the smart house as "a dwelling incorporating a communications network that connects the key electrical appliances and services, and allows them to be remotely controlled, monitored, or accessed" [10]. Bilal Ghazal and Khaled Al-Khatib have suggested that the smart home system is not a new science terminology but it is still away from people's vision. This is due to the fact that the technologies are available; they are not being fully employed

in solving problems. In fact, the majority of home appliances are somehow automated but the integration of these technologies, the inter-corporation of automated various appliances in an affordable design, and the ease of deployment due to distant communication provides peace of mind and convenience [10]. Right now, WSNs are useful platforms in IoT-based smart home applications, and data transmission in these networks is usually done through the wireless medium [11]. Some WSN technologies include Wi-Fi, Bluetooth, Zigbee etc. as shown in table 2 below.

TABLE II. WIRELESS SENSOR NETWORK TECHNOLOGIES

| WSN Standard | Wi-Fi | Bluetooth | Zigbee |
|---|---|---|---|
| Number of Nodes | 32 | 8 | >64000 per network |
| Frequency Band | 3.1-10.6 GHz | 2.4 GHz | 868/915MHz, 2.4 GHz |
| Range | 100 m | 10 m | 20-200 m |
| Power use | High | High | low |
| Data protection | 32bit | 16 bit | 16 bit |

One of the latest technologies using Wi-Fi is the ESP 8266

As Shown in figure1.


Figure 1. ESP 8266 Wi-Fi network [12]

The ESP can operate as both in a station mode and in soft access point mode by which it can create its own wi-Fi network [12].

There are many systems that have been proposed and developed for monitoring and controlling power consumption [13]. In short, there's both the measurement of the energy as well as the control. The proposed system for Mitanda home for the aged people will deploy wireless sensors such as light sensor to implement the management of energy and control the electrical devices.

## IV. METHODOLOGY

The system was designed using Proteus Design Suite. A simulation was then carried out in Proteus to verify that the system works as expected. Figure 2 shows the block diagram of the proposed system. The system comprises the following components: Wi-Fi module ESP 8266, Power Supply, 3.3v, LCD (16 x 2), Sensors, driver, Wi-Fi router and the load which can be light, fan or TV.

The ESP 8266 is monitored by the controller through the Wi-Fi network. The controller is configured to switch the electrical devices on and off through power circuit depending on the consumption.

A threshold of the consumed energy is configured on the controller beyond which the controller switches off the load. The energy (E) equation is shown in equation (1):

$$E = P \times t = V \times I \times t \quad \text{(watt hours)} \qquad (1)$$

Here Voltage V is constant, time t is the time for the current I drawn by the load. A current sensor would be needed so that the total energy consumed is determined according to equation (1) and compared with the threshold configured on the controller. For example, if the desired energy consumption for the stove at Mitanda home is 5.2 KWH/day, the proposed system would monitor and maintain the supply until that value is reached. Then it switches off.

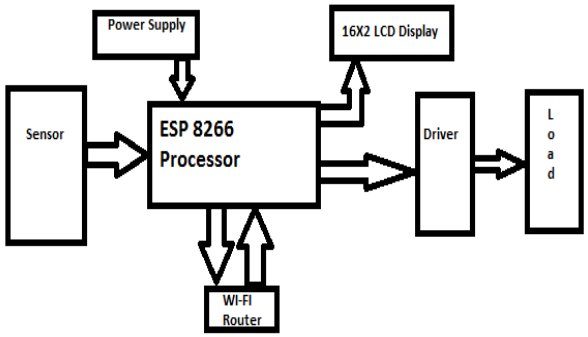



Figure 2. The block diagram of the proposed system

## V. PROPOSED SYSTEM

The proposed energy control system for Mitanda in Ndola will use Wi-Fi based energy measuring system in conjunction with a wireless sensor network system for remote monitoring and automatic control of the home devices. The advantage of using Wi-Fi is based on the fact that we are focusing on the implementation of Internet of Things using the ESP8266 module to connect to the internet. In other words the Wi-Fi is advantageous in the reasons below: Equipment can be placed almost anywhere, no unsightly cords running around [14].This will make it to be a feasible, low- power consumption, secure, efficient, flexible and scalable, cost-saving and finally supported by easy-to- deploy familiar interface. There are several technologies for application on laptops or mobile phones that can be utilized in improving the life style of the residents. In fact, the disabled people could find devices with touch screen control panel confusing and difficult to use. Therefore, they prefer a simpler remote control using laser-engraved backlit buttons, some switches, and equipped with a LCD screen to display necessary notifications. To each command button is associated a warning LED light that visualizes the situation status of the corresponding appliance. All operations are governed by a microcontroller where the EEPROM memory gives the opportunity to lock the remote by means of pass code stored in its memory [15].

The main component of the proposed system is the ESP8266 Wi-Fi Module which is a self-contained processor

with integrated TCP/IP protocol stack that can give any microcontroller access to a Wi-Fi network. The ESP8266 can perform the hosting an application as well integrating the Wi-Fi networking functions from other application processors. Each ESP8266 module comes pre-programmed with an AT command set firmware, making is simple to be connected to the Arduino device. The ESP8266 module is an extremely cost effective board with a huge, and ever growing, community [16]. Its high degree of on- chip integration allows for minimal external circuitry, including the front-end module, is designed to occupy minimal PCB area. All necessary information regarding how the ESP8266 can be fully implemented, even instructions on how to transform this module into an IoT (Internet of Things) solution appears in [17].

The installation of the sensors in the smart home will monitor the usage of appliances. The occupants' pattern of life at Mitanda in Ndola can be established by utilizing sensors to monitor usage of appliances and devices. The situation in a real world can be sketched and analyzed by using the data collected. Detection of the usage of electrical loads at Mitanda Home of the aged will be accomplished using transducers such as current and voltage sensors in which they will be the interface between the socket outlet and the load such as a stove or a fan. The circuit diagram for the proposed system is shown in Figure 3.

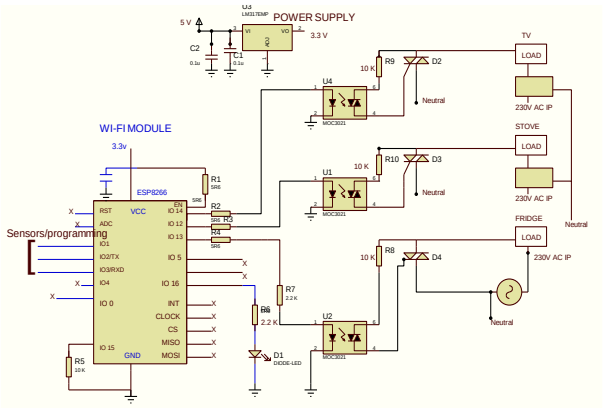



Figure 3. Circuit diagram of the proposed system.

## VI. RESULTS

In this work, home equipment such as a bulb or a heater was considered which can be monitored with the help of energy measurement and communication unit. This set of instruments is installed in the circuit of the bulb and the socket outlet for the measurement of energy and power consumption of any other home electrical devices. Energy measurement and communication unit transmits the measured values of the electrical parameters over a period of time to the home server with the help of Wi-Fi. Then home server incorporates the energy and power data from lights and from outlets through Wi-Fi access point (AP). Users can check the energy and power information as well as the energy consumption of home equipment and lights, with the help of home server as pointed out in [14]. The proposed system of addressing energy consumption problems will enable the care homes to have a sustainable means of energy consumption monitoring and controlling. Figure 4 shows the monthly energy consumption for 2017 in KWH with an average monthly consumption of 19 677.7 KWH.

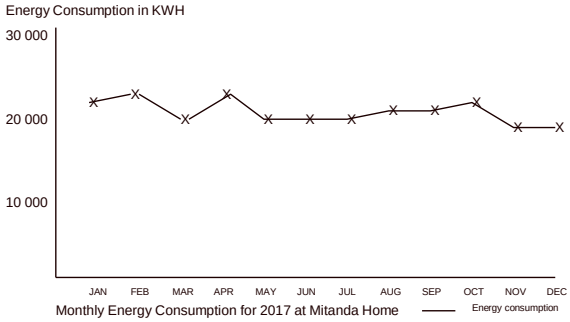



Figure 4. Monthly energy consumption for 2017

Figure 5 shows the simulation results of energy consumption with an average monthly consumption of 17 970 KWH.

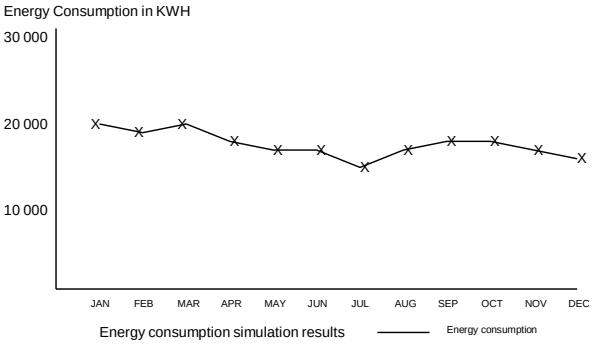



Figure 5. Simulation results for the consumed energy.

## VII. CONCLUSION

This work has identified areas of potential application of the Internet of Things (IoT) for the sustainability of old people's homes. This work has also shown other advantages that can be attained by the application of these technologies such as implementing the IoT applications in the medical monitoring and emergency response, agriculture, healthcare, energy management, and industrial automation. It has also shown that an efficient energy consumption home automation system for the care Homes of the aged people in Zambia based on Internet of Things (IoT) using WSNs was possible. The proposed system of energy management is intended to be used for monitoring and controlling the electrical energy consumption. The remote control of the devices can be achieved with the use of a computer or a smart phone. The smart phone will be used demonstrate the control of electrical devices. Future work will be needed for further application of sensor based control of devices. This work is undertaken for the purpose of reducing the cost of electricity in an old people's home and it is intended to influence decision makers in Zambia on the application of internet of things in automating of old people's homes and other application areas like health care homes for the physically challenged people.

As a result, this proposed system will be bring about sustainability in the management of electrical energy consumption in Zambia.

## REFERENCES


[1] Vijay S. Kale, Madan B. Matsagar, Avinash D. Sonawane, Chandrakant L. Ambekar, "Remote Temperature Monitoring System Using ARM, Arduino and ZigBee", (IJCTEE) volume 5, 2016 page 1.

[2] The Salvation Army, Zambia territory, "Mitanda home of the aged" (online) http://www.salvationarmy.org/zambia/F71F9C2ACAF93B5B802578030022483 available on 22.12.2017

[3] Nomusa Dlodlo, Josephat Kalezhi, "The Internet of Things in Agriculture for Sustainable Rural Development" International Conference on Emerging Trends in Networks and Computer Communications, May 2015, (available online on 16/01/2018).

[4] Hnin Nu Thaung, Zaw Myo Tun, Hla Myo Tun, "Automatic Energy Control and Monitoring System for Building", (IJSTR) volume 5, 2016

[5] Hoang Boi Nguyen and Tony Barnett, "Smart Homes for Older People", University of Tasmania, Australia, 2012

[6] United Nations, "world population ageing, 1950-2050" retrieved from https://population.un.org/ProfilesOfAgeing2017/index.html January 2018

[7] Anuroop Gaddam,"wireless sensor network based smart home for the elder care",massey university, New Zealand, 2011.

[8] Maramba Old People's Home.htm, 2009, retrieved from https://www.blogger.com/profile/09357164025343524658 available on 20.11.2017

[9] Avijit Mathur, Thomas Newe, "Comparison and overview of Wireless sensor network systems for Medical Applications "University of Limerick, Ireland. September 2014.

[10] [10] Bilal Ghazal and Khaled Al-Khatib, "Smart Home Automation System for Elderly, and Handicapped People using XBee". International Journal of Smart Home Vol. 9, No. 4 (2015) page 5

[11] Sandeep Pirbhulal, Heye Zhang, Md Eshrat E Alahi, "A Novel Secure IoT-Based Smart Home Automation System Using a Wireless Sensor Network", January 2017. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5298642/, December 2017

[12] https://arduino-esp8266 read the docs io / en / latest/esp8266 wifi/ readme.html

[13] Shenzhen Sailwider electronics co., ltd, "Wireless Home Energy Monitoring System" retrieved from http://www.sailwider.com/smartpower/index.htm, on January 22, 2018

[14] Devconhomesecurity, "Advantages of Utilizing a Wi-Fi Based Home Automation System" May 05, 2016, retrieved from http://devconhomesecurity.com/blog/advantages-utilizing-wifi-based-home-automation-system on 19th January 2018

[15] Archana N. Shewale, Jyoti P. Bari "Renewable Energy Based Home Automation System Using ZigBee", (IJCTEE)

[16] Sparkfun, "WiFi Module -ESP8266" retrieved from https://www.sparkfun.com/products/13678 , January 20, 2018

[17] Hivos people unlimited "Zambia:energy profile" retrieved from https://www.hivos.org/sites/default/files/zambia_profile.pdf

# *A Comparative Analysis of Web Searching and Information DiscoveryTechniques-Systematic Literature Review*

Darius Bwalya
School of Science, Engineering and Technology
Mulungushi University
Kabwe, Zambia
sys.engi.darious@gmail.com

Douglas Kunda
School of Science, Engineering and Technology
Mulungushi University
Kabwe, Zambia
dkunda@mu.ac.zm

***Abstract*-Never before has there been so much information explosion than in the 21st century. With this increase comes, a great challenge in accessing the actual desired content that is relevant information out of superfluous data links that are often harvested at every search, which often leads to more time being wasted in sifting unwanted links with ultimate delays. Just like how efficient is achieved when there is systematic and standard orderly arrangement of files along with efficient access method of this information, there has been even a greater need for standardised formatting of information and for technically devising efficient methods of accessing information in the cyber space. This has given rise to information formatting and search technology that are technically advanced namely Web search engines, Web crawling, Web indexing, Page and site ranking, Spam detection, Content ranking, Collaborative filtering, Social recommendation, Personalization, Social tagging. In this paper these web formatting and search technologies are analysed and compared in order to give insight when it comes to user assisting search methods to use in the design of specific purpose web sites and applications by software developers, e-marketers and advertisers and many other purposes too numerous to mention.**

***Keywords:*** *World Wide Web, Web search engines, Web crawling, Web indexing, Page and site ranking, Spam detection, Content ranking, Collaborative filtering, Social recommendation, Personalization, Social tagging*

## I. INTRODUCTION

Today it is not the question of whether one will find information online but relevant to the search. There are billions of hypertext documents and their multimedia content interconnected with hyperlinks that have formed a mesh of information called World Wide Web or simply www. This information is stored on computers called servers and accessed by computers called clients via client applications called web browsers such as Chrome. When a user types in the browser a search subject in a search engine, a web browsers sends a request for hypertext links for documents or web pages to the server that seem relevant to the search, but may not be relevant to the user.

The search engine is a software system that indexes web pages and allows internet users to query their contents for specific words and phrases[1]. The challenge is that a word or phrase can yield a thousand links as long as there is a much for the searched word or phrase especially to a novice online user. Though it requires good internet skills to have a quality search, advances have been made in the information searching and discovery. The objective is to help users in finding what they need without expecting or demand that they acquire specialized knowledge on advanced formatting techniques such as meta

## II. CATEGORIZATION OF WEB SEARCHING AND INFORMATION DISCOVERY TECHNIQUES

The following are the Information searching and discovery techniques that are now considered: Web search engines, Web crawling, Web indexing, Page and site ranking, Spam detection, Content ranking, Collaborative filtering, Social recommendation, Personalization and Social tagging. They are also classified in the table 1.0 below.

TABLE I. CLASSFICATION OF WEB SEARCHING AND INFORMATION DISCOVERY TECHNIQUES

| Study | Web Searching | Study | Information Discovery |
|---|---|---|---|
| [1] | Web Search Engines | [2] | Web Crawling |
| [3] | Content Ranking | [4] | Web Indexing |
| [5] | Collaborative filtering | [6] | Social Recommendation |
| [7] | Social Tagging | [8] | Personalization |
| [9] | Page and Site Ranking | [10] | Spam Detection |

## III. WEB SEARCH ENGINES

As already alluded to in the introduction the search engine is a software system that indexes web pages and allows internet

users to query their contents for specific words and phrases. They index tens to hundred million web pages and answer tens of million queries every day[1]. However Ricarte and Gomide(as cited [11] identified the common challenge of Information Retrieval(IR) as being considering only lexicographical aspects when it comes to matching web pages and queries, ignoring semantic aspects. This implies that queries are made on word for word matching process instead of concept-matching process which [11] proposed in their research. Another challenge was discovered to be a situation where a user may understand the foreign language but may not be conversant with formulating the query in the non-native speaking. It thus becomes a real challenge to cut across language barrier during information retrieval."Personalized IR systems do not only retrieve documents that are relevant to a query,but also relevant to a user's individual needs. Such systems greatly help users in satisfying their information needs by minimizing the information overload they experience." [8]

## IV. Web Crawler

This is a computer program that browses the World Wide Web in a methodical, automated manner or in an orderly fashion. Web crawlers are mainly used to create a copy of all the visited pages for later processing by a search engine that will index the downloaded pages to provide fast searches.[2]

There are obvious good reasons to this approach to gather the navigated websites. One most important application is in search engines such as google that then make use of this vast crawled data by indexing it and performing millions of simple to complex queries on behalf of internet users. Crawlers can also be used for automating maintenance tasks on a Web site, such as checking links or validating html code. Crawlers can be used to gather specific types of information from Web pages, such as harvesting e-mail addresses (usually for sending spam)[2] Apart from web crawling being valuable in information searching and discovery, the web crawling technology feature prominently in such applications as web data mining and extraction, social media analysis, digital preservation (i.e., ensuring continued access to information and all kinds of records, scientific and cultural heritage existing in digital formats), detection of web spam and fraudulent web sites, finding unauthorized use of copyrighted content (music, videos, texts, etc.), identification of illegal and harmful web activities (e.g., terrorist chat rooms), virtual tourism[12].

The major challenges to implementing web crawling are mainly to do with the sheer quantity of published web sites available on the WWW. They are now presented here:

- Downloading costs of already astronomical number of websites and content that is there and which is increasing exponentially by the day hinder most attempts by new crawling entrants such as individuals, Government agencies and others leaving this to commercially stable business entities mainly Google, Bing and Yahoo.
- Furthermore, not only should web crawling be done once, but if the search results must reflect updated content then it has to be done periodically which yet again is simply prohibitive to most attempts to web crawling.

## V. Web Indexing

This is a service in which harvested web site and content by web crawlers are classified and an index generated by search engines to make it easier for the data or information to be queried in order that relevant information is retrieved efficiently.

To achieve efficient indexing, most Information Retrieval(IR) Systems use index terms that match search queries from the user to document as matching search terms to full-text documents is computationally inefficient and does not guarantee the retrieval of the relevant information. Indexing involves assigning descriptive information about the page called metadata in a controlled process. Index terms represent the characteristics and relationships among information items. Indexers probes the content of a document conceptually and decide not only what the document is about, but also why it is likely to be of interest to a particular group of users. During indexing process index terms may be selected by the automated process from the initial set of terms. Final decisions about index terms, however, are made manually according to complex rules and standards. It undeniable that indexing is a highly formalized and knowledge intensive job. Though indexing is a sophisticated process combining significant human knowledge and expertise, index terms do not always reflect how readers think about a document.[4].

Readers may often confuse social tags and index terms. Though the two are related in being descriptive metadata [4] identify the difference stating that index terms come from highly controlled vocabularies whereas social tags mainly constitutes plaintext.

## VI. Social Tagging

In this form of information discovery, users annotate information resources like bibliographic references, multimedia or Websites with plain text or phrases collectively referred to as tags. Just like the index terms are used to describe information items, tags accomplishes the same thing.

The advantage of social tags over index terms is that users do need to have special skills or special training in order to use the tags and benefits are realised by users instantly. Users may use social tagging to for instance draw attention to a social issue e.g #corruption.

## VII. Page and Site Ranking

This often refers to the physical order or location of the search link on a Search Engine Results Page (SERP). According to the study conducted by [13] they established that when a website is highly ranked on a Search Engine Results Page, it has strong correlating effect to a user's probability that he or she will click on the web site link. This is illustrated in the fig 1.00 and 1.01 below.

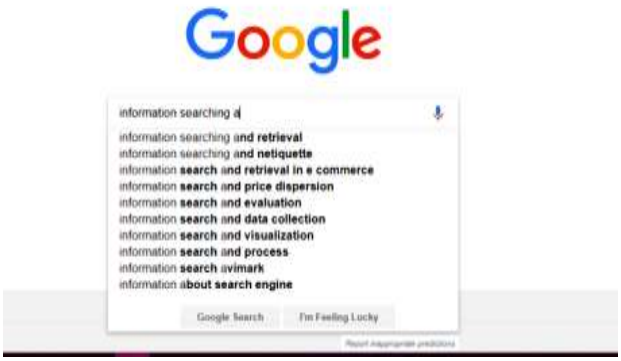



Fig 1.00 *Search prediction*

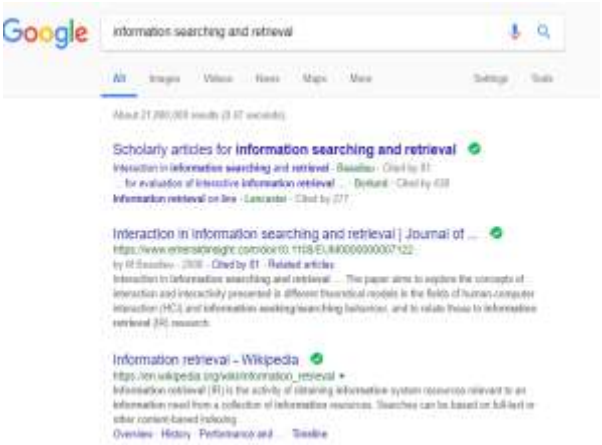



Fig 1.01 *Showing a Search Engine Results Page*

Fig 1.01 shows the highly ranked link and the study already mentioned discovered that the user in this case is likely to click the very first link "Scholarly articles for information retrieval". However there are other factors that may affect whether a link will be clicked or not such as domain own reputation which means despite it falling under other links on the SERP, a user is likely to click on it. This means that should it be high ranked it benefits even more. The descriptive metadata which is information about data may facilitates ranking of a websites. This is as a result of good use of the HTML tag <meta> which has values like "keywords" which allows a web designer to define keywords to the search engine like what the page is all about. See an example below in fig 1.02 below.

```
<meta name="keywords" content="HTML, CSS, XML, XHTML, JavaScript">
```

Fig 1.02 *definition of a meta element with name attribute value keywords likely to be searched for*

This may be an advantage thing to do on the commercial web site especially and besides the more popular the website the more successful its owners become.

## VIII. Collaborative filtering

This is a technique of analysing search behaviours and search practices of individuals which leads to discovering of other users of similar research interest allowing them to collaborate in the information search.[5] This technique plays a valuable role in increasing relevant links whereas lessening time spent to bad links among collaborators. This leads to forecasting interesting information to target users.

The research conducted by [5] assumed that the users with similar research speciality and interests show the tendency of having similar needs during their search for information. It follows therefore that establishing this trend will enable automatic connection to other users in the same discipline. [14] identified three categories of Collaborative filtering namely user-based, resource material-based and a hybrid. User-based algorithm is applied in anticipating what a current user is interested in by means of a neighbour's user rating. Item-based apply similarities to determine the nature.

## IX. Social Recommendation

With the advent of social networks more and more users get connected for various reasons among them interaction forming network of friends online, review of products or service, blogging, sharing of information, commenting on some news story trends [6]. This often affect users choices when it comes to purchase of services and products and online vendors and service providers tend to exploit the network group's preferences and behaviour which through social interaction in groups provides a clue as to what is popular [6].

How this common behaviour comes about is through what is now know us social recommendation wherein individual user in a social network may be influenced by the recommendation of the others for a certain product or service. These are usually trusted friends who may have used or experienced a particular product or service and share this on social media for example close trusted friends may give a rate to a certain product 5 stars out of 5 which obviously would influence a user to trust a product or service.

The challenge to social recommendation however was identified by Lai et al.(2018) as information overload because so many social networks in which users interact back and forth on one or more different topics. "Due to the discussions above numerous kinds of information are rapidly and constantly generated, so the problem of information overload is undoubtedly becoming a more serious in-depth issue"[6].

Further the coming of social networks allows users to publish information which may not be factual or which may not have been subjected to editing and review and simply opinionated for instance others may review a product based on mere heresay. [15] admitted that "Validating the data on the Internet is a challenging proposition and pitfalls by new users and experienced ones are far too often". [6] proposed in their research an effective recommendation method combining both

social networks and trust relationships. Fig 1.03 illustrates this their research proposal.

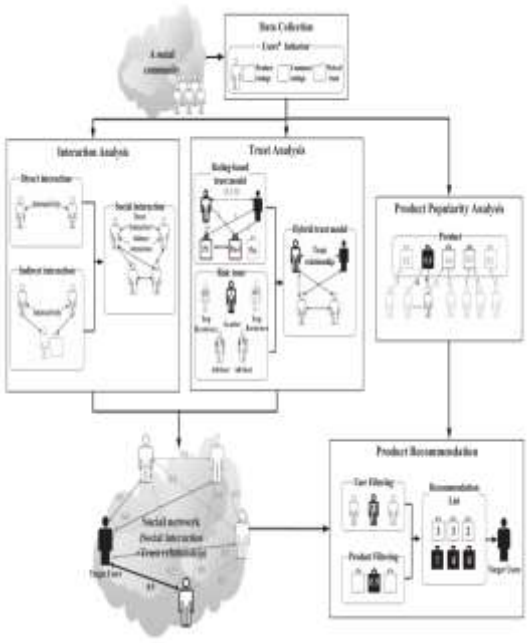



Fig. 1.03 *[6] proposed "Effective Recommendation" integrating Social Network and Trust relationship*

According to [6] their proposed work analyses users' behaviour and effectively integrates social interaction, trust relationship and product popularity to make product prediction for target users. Based on users' rating behaviour, they classified social interaction into direct and indirect interactions. Further the trust relationship is also explored and similarly classified into direct and indirect trust relationships. User interaction and trust relationships are integrated in the proposed method to identify the connections of a target user and other users. On the basis of the preference of highly connected users, popular products have high priority to be recommended to target users [6]. Thus, the proposed method can analyse the connections among users based on their explicit and implicit relationships.[6] It also considers the impact of popular products on user preference when making product recommendations to target users[6].

## X. Personalization

This is a method of collecting past customer buying behaviour that is, purchases made and browsing interests for example bookmarks. Based on this information tailored recommendation for product or service is provided to the target customer. An advantage of this technology is that it restrains information overload.

The issues that may arise with personalization are mainly to do with privacy concerns by customers[15] through their research discovered that although privacy issue may be a concern especially to customers that may have had privacy invasion in the past, their findings suggest that under certain circumstances, perceived personalization quality can outweigh the impact of privacy concerns. For example Howard and Kerin (2004 as cited in [16], hinted on how to achieve that in their research which revealed that when a direct marketing message was personalized (by adding a recipient's first name and a sender's initial), the response rate to personalised recommendation improved significantly. This compliments the social need that people have to want to be known by name. The challenge therefore to vendors and service providers is to improve the perceived quality of personalization service under offer so that customer fears for privacy safety are allayed and the research also indicated that product and service providers still have an opportunity to explore personalization to realize high sales.[15].

## XI. Spam Detection

Web spam is one of the main current problems of search engines because it strongly degrades the quality of the results leading many people to become frustrated by constantly finding spam sites when they query for legitimate content. [10].

This involves web pages that get undeservedly high rank rating through fraudulent means. As more and more techniques are developed to combat this web deceit, the spam also tend to evolve and become as sophisticated. [10] categorize web spam in three groups namely: link spam, Content spam, and cloaking.

- Link spam presents links that exits to deceive in terms of rank. The structure of a link spam exploits the link-based ranking algorithms such as HITS or PageRanks which ranks high a web site based on highly ranked websites linking to it [10].
- Content Spam involves deliberate altering of search engine's logical view (description) of a web page for example the use of unrelated metadata keywords that are more popular with query terms from various users but are not related to the page content [10].
- Web Cloaking involves delivering search to the user based on location or IP address. This may be misused by spammers to alter the search preferred by the user.

The web cloaking has its own benefits though, which [17] identified as being location based application in mobile devices that are able to offer customized service to the end users for example in the search and rescue operation where information based on the user location would so vital.

## XII. Content Ranking

This involves a ranking of the contents to the user query based on his search context and ranking the results according to his demands and priotization of what is relevant. [18] stressed the importance of trustworthiness when it comes to this type of ranking method arguing that it plays a significant role in prioritizing services depending on the requesters' demands. For example, if there are two services with similar functionality but one is more secure than the other, then possibly, the chosen service is usually the more secure service.

[18] defines trustworthiness in this context as a synthesized and a composite feature which includes sub-

features. For instance, security involves authorization, authentication, and integrity; and reliability includes failure and repair models. Selecting an airline is another example. When choosing an airline, a customer may consider some sub-features to make their decision: customer lounge, onboard service, and customer services quality. Probably one customer may rank the airlines with respect to the overall rating, yet another may prefer to rank them according to the sub-features. In that case, the services could be assessed according to their sub-features rather than the general feature.[18].

Application of content ranking would be in the social media context in which users may be posting content online on a regular basis. Overtime the user online content can be ranked according to importance. The challenge is how to measure importance.

The results of the research done by [3] showed that the social media (in which user are allowed to post content rank it e.g clicking likes) ranking scheme was clearly better than Random Selection Chronological Ordering(RSPICO), Random Selection non-Chronological Ordering (RSPIn-CO).

## XIII. Comparisons of Web Searchng and Information Discovery Techniques

Web searching as already noted implies all means by which users get what they need from the web using assistive tools like web search engine. On the other hand it has been made clear that information discovery involves machine learning about what is likely to be relevant to the information seeker and may involve discovering users of the same information needs and automatically link them or learning users search history and consequently discovering a browsing pattern. The table below compares the web search and information discovery techniques and studies that have been done on them in table I below

TABLE II. Categorizing Various studies Done on Web Searching and Discovery techniques

| | Study | Purpose | Methodology | Findings |
|---|---|---|---|---|
| **Collaborative Filtering** | | | | |
| | [5] Information Discovery | To propose the Collaborative Search System that attempts to achieve collaboration by implicitly identifying and reflecting search behaviour of collaborators in an academic network that is automatically and dynamically formed | Used the Digital Bibliography Library Project( DBLP). Using this enabled the formation of research communities implicitly and dynamically based on users research presence in the search environment and in publication scenario | The proposed system considerably improves the relevancy and reduces the time spent on bad links, thus improving recall and precision. |
| | [14] Information Discovery | Proposal of a method to discover the cloud services via semantic concepts and collaborative filtering | Experiments were conducted to examine and analyze the proposed method. To demonstrate the behavior of the proposed method, it was simulated in MATLAB 7.10. | The obtained results from experiments have illustrated that the presented approach improves the efficiency of service discovery in the cloud computing and obtains fewer execution times against the current approaches |
| **Content Ranking Studies** | | | | |
| | [3] **Category:** Information Discovery | To propose a novel content ranking component which ranks posted items based on a social computing method, driven by the power and influence of social network users. | Experimenting the proposed system on real social media. | Improved performance |
| **Page and Site Ranking Studies** | | | | |
| | [19] | To propose a new algorithm for ranking web pages based on web graph. The objective is determining the score of each web page based on paths which can be reached to that web page from other web pages as well as the out-degree (number of out links) of pages in the traverse paths. | In the first step, with formulation of ranking as an Reinforced Learning (RL) problem, a new connectivity-based ranking algorithm, called RL Rank, is proposed.<br><br>In the next step, we introduce a new hybrid approach using combination | Experimental results show using RL concepts leads significant improvements in raking algorithms |

| | | | | |
|---|---|---|---|---|
| | | | of BM25 as a content-based algorithm and RL Rank. | |
| | [9]<br><br>**Category:** Information Searching | To propose a ranking approach which considers visual similarities among web pages by using structure and vision-based features. | The conducted study is composed of two parts. In the first part, structural similarities are analyzed with the proposed concept of "layout components" along with visual inspection of DOM trees. In the second part, a computer-vision based method named histogram of oriented gradients (HOG) is employed to reveal local visual cues in terms of edge orientations.<br><br>Questionnaire study was done. | According to the findings of the comparative study, our approach outperforms some structure and vision-based studies in the literature. With this achievement, web pages could be employed as a query item to find other, similar web pages by taking into consideration that they are web pages, instead of images or anything else. |
| | [20] | This study aims to present a new web page recommendation system that can help users to reduce navigational time on the internet | The proposed design is based on the primacy effect of browsing behaviour, that users prefer top ranking items in search results. This approach is intuitive and requires no training data at all. | A user study showed that users are more satisfied with the proposed search methods than with general search engines using hot keywords. |
| **Personalization Studies** | | | | |
| | [21]<br><br>**Category:** Information Discovery | The study first proposes a method to determine user authority in a social tagging system, in which the quality authority and quantity authority of users are | A resource model is constructed by summing up the tags from each user and their corresponding weights. User models are then obtained based on the | The results show that the average GP relevance of the authoritative user based algorithm reaches 0.6115 much better |

| | | | | |
|---|---|---|---|---|
| | | calculated from a user co-occurrence network, which is derived from users' participation in the social tagging system. | resource models. An experiment was conducted on a dataset crawled from *Delicious.com*. | than two benchmark algorithms. |
| | [8] | The purpose of this paper was to present a comprehensive study of user profile representation techniques and investigate their use in personalized cross-language information retrieval (CLIR) semantic search systems, outlined within broad areas of classification | The user profiles consisted of weighted terms computed by using frequency-based methods such as tf-idf and BM25, as well as various latent semantic models trained on monolingual documents and cross-lingual comparable documents. The study paper also proposes an automatic evaluation method for comparing various user profile generation techniques and query expansion methods. | Experimental results suggested that latent semantic-weighted user profile representation techniques are superior to frequency-based methods, and are particularly suitable for users with a sufficient amount of historical data. The study also confirmed that user profiles represented by latent semantic models trained on a cross-lingual level gained better performance than the models trained on a monolingual level. |
| | [22] | This research examines consumers' percepti ons of personalized advertising (PA), a new and emerging trend in online advertising, within the context of social networking sites. | Privacy calculus theory was used and extended via its integration with constructs from other streams of literature. Both antecedents and outcomes regarding consumer privacy concerns toward PA | The results showed that a number of factors, such as invasiveness, privacy control, perceived usefulness, and consumer innovativeness influence |

| | | | | |
|---|---|---|---|---|
| | | | were empirically investigated. | consumer s' behaviour al intentions concernin g PA. |
| | [23]<br><br>**Category:** Information Searching | To provide technical review of semantic search methods used to support text-based search over formal Semantic Web knowledge bases. | This review focused on a single mode of user interaction and presents in detail several forms of algorithmic approaches for distilling information from knowledge bases to satisfy user queries. | The idea was to help the user explore the domain, to find out what is there, and to construct complex queries from possibly several atomic or incremental operations. |
| **Ranking Studies** | | | | |
| | [23]<br><br>**Category:**I nformation Searching | To provide technical review of semantic search methods used to support text-based search over formal Semantic Web knowledge bases. The focus was on ranking methods and auxiliary processes explored by existing | This review focused on a single mode of user interaction and presents in detail several forms of algorithmic approaches for distilling information from knowledge bases to satisfy user queries. | The idea was to help the user explore the domain, to find out what is there, and to construct complex queries from possibly several atomic or incremental operations. |
| | [3]<br><br>**Category:** Information Searching | The study investigated whether a search engine's ordering of algorithmic results has an important effect on website traffic. | Experiments by means of a unique dataset were done. | Additionall y to its relevance, the rank of a website strongly and significantl y affects the likelihood of a click. |
| | [7]<br><br>Category: Information Discovery | This article explores how hashtags are used to coordinate and accentuate the values construed in a corpus of Twitter posts (tweets) about depression. | The study employed a discursive system, communing affiliation, to interpret how particular values about depression are positioned as bondable in this ambient | Understand ing how people are forging alignments and negotiating meaning through social tagging practices. |
| | | | environment | |
| | [4] | The study compared CiteULike tags to Medical Subject Headings (MeSH) terms for 231,388 citations indexed in MEDLINE. | Three increasingly progressive levels of text processing, ranging from normalization to stemming, to reduce the impact of lexical differences. | Results showed that CiteULike tags and MeSH terms were quite distinct lexically, reflecting different viewpoints/ processes between social tagging and controlled indexing. |
| **Social Tagging** | | | | |
| | [7]<br><br>Category: Information Discovery | This article explores how hashtags are used to coordinate and accentuate the values construed in a corpus of Twitter posts (tweets) about depression. | The study employed a discursive system, communing affiliation, to interpret how particular values about depression are positioned as bondable in this ambient environment | Understand ing how people are forging alignments and negotiating meaning through social tagging practices. |
| | [4] | The study compared CiteULike tags to Medical Subject Headings (MeSH) terms for 231,388 citations indexed in MEDLINE. | Three increasingly progressive levels of text processing, ranging from normalization to stemming, to reduce the impact of lexical differences. | Results showed that CiteULike tags and MeSH terms were quite distinct lexically, reflecting different viewpoints/ processes between social tagging and controlled indexing. |
| **Social Recommendation** | | | | |
| | [6]<br><br>Category: Information Discovery | This paper proposes a novel social recommenda tion method on the basis of the integration of interactions, trust relationships and product popularity to predict user | Focused on analyzing user interactions to infer their latent interactions in accordance with the user ratings and corresponding reviews. | The experiment al results show that the proposed recommend ation method has a better recommend ation performanc e in comparison s to other methods. |

| | | | | |
|---|---|---|---|---|
| | | preferences, and recommend relevant products in social networks. | | Furthermore, the proposed method can not only reduce the time and effort users spend on querying information, but also positively relieve the problem of information overload. |
| | [17] Category: Information Discovery | The study proposed a privacy service that helps users to maintain their privacy policy in a flexible and incremental way. | Used a qualitative evaluation study whose results illustrate several challenges that should be handled in the design of such a service | It was discovered that the users should have much flexibility, but should not be overwhelmed with the configuration of their privacy preferences. |
| | [15] Category: Information Discovery | The paper explores the use of concepts in cognitive psychology to evaluate the spread of misinformation, disinformation and propaganda in online social networks. | The cognitive process involved in the decision to spread information involves answering four main questions viz consistency of message, coherency of message, credibility of source and general acceptability of message. We have used the cues of deception to analyse these questions to obtain solutions for preventing the spread of misinformation. | The validation of the proposed methodology has been done on the online social network 'Twitter'. |

## XIV. Methodology

A systematic literature study to find unique definitions, descriptions and issues of web searching and discovery techniques key technologies was done applying the two fold strategies:

The first and primary strategy was to utilize the following scholarly research resources emeraldinsight.com, researchgate.com, wileyonlinelibrary.com, sciencedirect.com, springerlink.com, Springeropen.com, link.springer.com, tandfonline.com,scholar.google.com,ieeexplore.ieee.org/Xplore/home.jsp and doaj.org.

In a second strategy, a search for gray literature on the Internet using the search engines Google and Bing was conducted. Google and Bing were used because these are the most widely used search engines. Advanced search option of google was utilized, selecting English as the preferred language, and turned the option for regional differences off. The key words stated on the outset formed a basis for selecting best literature for review. All searches in the gray literature were performed between August 2018 and October 2018.

## XV. Results

Scholarly resources research yielded 41 papers from a period of 1998-2018 which were studied intensively. Grey search areas yielded quite several thousand hits which were filtered systematically. For example a search on google search engine for "collaborative filtering" yielded 1,480,000 results, after narrowing down further with specific field of study with "collaborative filtering computer science", the results presented came to 828,000 hits, specifying results format such to 'pdf' for instance further narrowed down the search to only 552,000 hits. This systematic way was applied on bing search engine too.

## XVI. Challenges

Studies on web search engines revealed that often queries show results that are relevant to the query but not so much to the user's concept or idea.

Studies on Crawling revealed Downloading costs of already astronomical number of websites and content that is there.

Papers on indexing revealed that the index terms do not always reflect how readers think about a document.

Personalization concerns were to do with privacy concerns.

Most challenges highlighted by the scholars have been handled using various models detailed in the very research papers reviewed.

## XVII. Conclusion

This has been an analysis of various Web Information Searching and Discovery techniques that have become sophisticated over the years due to the abundant information explosion in this information age. It is clear that much work has been done to make querying of information not only practical but even more relevant while at the same time employing machine learning techniques by various research models.

This in depth analysis and comparison of the various techniques discussed here can provide an insight when it comes to the best design practices to do with user assisted information searching and discoveries and may be applied in different applications by local developers to include these technologies in their design in such areas as collaborative learning applications, emarketing, ecommerce and even social network users.

## REFERENCES


[1] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," *Comput. Netw. ISDN Syst.*, vol. 30, no. 1–7, pp. 107–117, Apr. 1998.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] S. AbdulNabi, "Effective Performance of Information Retrieval on Web by Using Web Crawling," *Int. J. Web Semantic Technol.*, vol. 3, no. 2, pp. 63–72, Apr. 2012.

[3] K. Ntalianis, A.-B. M. Salem, and I. E. Emary, "Social Media Content Ranking Based on Social Computing and User Influence," *Procedia Comput. Sci.*, vol. 65, pp. 148–157, 2015.

[4] D. H. Lee and T. Schleyer, "Social tagging is no substitute for controlled indexing: A comparison of Medical Subject Headings and CiteULike tags assigned to 231,388 papers," *J. Am. Soc. Inf. Sci. Technol.*, vol. 63, no. 9, pp. 1747–1757, Sep. 2012.

[5] S. Renugadevi, T. V. Geetha, R. L. Gayathiri, S. Prathyusha, and T. Kaviya, "Collaborative search using an implicitly formed academic network," *Aslib J. Inf. Manag.*, vol. 66, no. 5, pp. 537–552, Sep. 2014.

[6] C.-H. Lai, S.-J. Lee, and H.-L. Huang, "A social recommendation method based on the integration of social relationship and product popularity," *Int. J. Hum.-Comput. Stud.*, Apr. 2018.

[7] M. Zappavigna and J. R. Martin, "#Communing affiliation: Social tagging as a resource for aligning around values in social media," *Discourse Context Media*, vol. 22, pp. 4–12, Apr. 2018.

[8] D. Zhou, S. Lawless, X. Wu, W. Zhao, and J. Liu, "A study of user profile representation for personalized cross-language information retrieval," *Aslib J. Inf. Manag.*, vol. 68, no. 4, pp. 448–477, Jul. 2016.

[9] A. S. Bozkir and E. Akcapinar Sezer, "Layout-based computation of web page similarity ranks," *Int. J. Hum.-Comput. Stud.*, vol. 110, pp. 95–114, Feb. 2018.

[10] L. Araujo and J. Martinez-Romo, "Web Spam Detection: New Classification Features Based on Qualified Link Analysis and Language Models," *IEEE Trans. Inf. Forensics Secur.*, vol. 5, no. 3, pp. 581–590, Sep. 2010.

[11] P. J. Garcés, J. A. Olivas, and F. P. Romero, "Concept-matching IR systems versus word-matching information retrieval systems: Considering fuzzy interrelations for indexing Web pages," *J. Am. Soc. Inf. Sci. Technol.*, vol. 57, no. 4, pp. 564–576, Feb. 2006.

[12] D. Shestakov, "Current Challenges in Web Crawling," in *Web Engineering*, vol. 7977, F. Daniel, P. Dolog, and Q. Li, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 518–521.

[13] M. Glick, G. Richards, M. Sapozhnikov, and P. Seabright, "How Does Ranking Affect User Choice in Online Search?," *Rev. Ind. Organ.*, vol. 45, no. 2, pp. 99–119, Sep. 2014.

[14] N. J. Navimipour, B. Keshanchi, and F. S. Milani, "Resources discovery in the cloud environments using collaborative filtering and ontology relations," *Electron. Commer. Res. Appl.*, vol. 26, pp. 89–100, Nov. 2017.

[15] K. P. K. Kumar and G. Geethakumari, "Detecting misinformation in online social networks using cognitive psychology," *Hum.-Centric Comput. Inf. Sci.*, vol. 4, no. 1, Dec. 2014.

[16] C. Li, "When does web-based personalization really work? The distinction between actual personalization and perceived personalization," *Comput. Hum. Behav.*, vol. 54, pp. 25–33, Jan. 2016.

[17] V. Sacramento, M. Endler, and C. de Souza, "A privacy service for location-based collaboration among mobile users," *J. Braz. Comput. Soc.*, vol. 14, no. 4, pp. 41–57, Dec. 2008.

[18] A. Bawazir, W. Alhalabi, M. Mohamed, A. Sarirete, and A. Alsaig, "A formal approach for matching and ranking trustworthy context-dependent services," *Appl. Soft Comput.*, Aug. 2018.

[19] V. Derhami, E. Khodadadian, M. Ghasemzadeh, and A. M. Zareh Bidoki, "Applying reinforcement learning for web pages ranking algorithms," *Appl. Soft Comput.*, vol. 13, no. 4, pp. 1686–1692, Apr. 2013.

[20] L. Chen and C. Luh, "Web page prediction from metasearch results," *Internet Res.*, vol. 15, no. 4, pp. 421–446, Sep. 2005.

[21] J. Wei, F. Meng, and N. Arunkumar, "A personalized authoritative user-based recommendation for social tagging," *Future Gener. Comput. Syst.*, vol. 86, pp. 355–361, Sep. 2018.

[22] J. T. Gironda and P. K. Korgaonkar, "iSpy? Tailored versus Invasive Ads and Consumers' Perceptions of Personalized Advertising," *Electron. Commer. Res. Appl.*, vol. 29, pp. 64–77, May 2018.

[23] C. L. Koumenides and N. R. Shadbolt, "Ranking methods for entity-oriented semantic web search: Ranking Methods for Entity-Oriented Semantic Web Search," *J. Assoc. Inf. Sci. Technol.*, vol. 65, no. 6, pp. 1091–1106, Jun. 2014.

[24] Henry, L. A. "Information search strategies on the Internet: A critical component of new literacies." *Webology*, vol. 2, no. 1, Article 9. Available at:http://www.webology.org/2005/v2n1/a9.html.

# Collaboration on the Web: a Review on Web 2.0, Social Software and Wikis

Hellen Muchindu Syachaba
*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
msyachaba@gmail.com

Douglas Kunda
*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
dkunda@mu.edu.zm

***Abstract* - The Web has changed the manner in which we work these days, it has turned out to be progressively participatory, easier for users to contribute and work cooperatively. It offers new interaction possibilities that facilitate desktop-like interfaces; it is also the way in which new software is built by mashing up distributed components using service-oriented communication. The purpose of this article is to present literature review on some of the technologies and services that has enabled the web to be more interactive and collaborative for the users. In particular Wikis, Social Software and Web 2.0**

***Keywords: Wikis, Social Software, web 2.0.***

## I. INTRODUCTION

The World Wide Web is a system of interlinked hypertext documents accessed via the Internet. With a web browser, one can view web pages that contain text, images, videos, and other multimedia and navigate between them via hyperlinks [1]. The Web is also the largest transformable-information construct that was introduced by Tim Burners-Lee in 1989 at first. Much advancement has been made about the web and related technologies in the past two decades from Web 1.0 as a web of cognition to web 2.0 as a web of communication [2]. The web as an information space has had much progress since 1989 and it is moving toward using artificial intelligent techniques to be a massive web of highly intelligent interactions in a close future [2]. New services and software have transformed the Web from being a predominantly read-only medium to one where anyone can publish and share web contents. Web 2.0 tools promote different types of communication: one-to-one, one-to-many, or many-to-many, synchronous and asynchronous. In addition, th tools can be used to search, share and create different media: from text (Blogs and Wikis) to images in Flickr, audio, podcasting and video in YouTube [3]. Web 2.0 is the umbrella term used to encompass several developments within the web whilst social software is application/technology of the web 2.0 and the term wikis falls under the type of social software of the web 2.0. The wiki is a basic tool of web2.0

The rest of the paper is organised as follows: in the next section we discuss basic construct of Web 2.0 followed by discussion on Social Software in section III. Section IV discuss Wikis and Section V concludes the paper.

## II. WEB 2.0

There isn't full agreementon the definition of Web 2.0. For instance, Wikipedia defines Web 2.0 as a second generation of services obtainable on the World Wide Web that allow individuals to collaborate, and share data on-

line. Initially, Web 2.0 was first adopted by the youth audience [4].

The key characteristics of Web 2.0 include [6]:

*A.* It's an attitude, not a technology: an acknowledgement that Web 2.0 isn't primarily a couple of set of standards or applications, however a brand new mental attitude to how the web are often used.

*B.* A network effect: This describes applications that are more effective with increase in the numbers of users.

*C.* Openness: the event of additional liberal licensees (such as Creative Commons; open sources licenses for software) will enable integration of information and utilization of software while not encountering legal barriers.

*D.* Trust your Users: As opposed to making advanced access routines, an increasingly liberal methodology are often taken that make it easier for users to form and use services.

*E.* Network as a platform: the web is currently being used to offer access to web applications, and not simply informational resources. This enable users to use applications online without having to travel through the cumbersome exercise of putting the software on their local computers.

*F.* It's technically Open design, Open source software and Open customary with privileges to utilize and generate contents by anybody.

Web 2.0 tools promote people to meet virtually, share opinions and interests. It's additionally enabling content to be made and shared in real time, with end-users sometimes able to add content to applications themselves. This means that Web 2.0 technologies support open communications and deliver users the liberty to share their opinions and suggestions [7].

Figure 1 shows the comparison between Web 1.0 and Web 2.0

*Figure 1 A Comparison of Web 1.0 and Web 2.0* [2]

| **Web 1.0** | **Web 2.0** |
|---|---|
| Reading | Reading/Writing |
| Companies | Communities |
| Client-Server | Peer to Peer |
| HTML, Portals | XML, RSS |
| Taxonomy | Tags |
| Owning | Sharing |
| IPOs | Trade sales |
| Netscape | Google |
| Web forms | Web applications |
| Screen scraping | APIs |
| Dialup | Broadband |
| Hardware costs | Bandwidth costs |
| Lectures | Conversation |
| Advertising | Word of mouth |
| Services sold over the web | Web services |
| Information portals | Platforms |

## III. SOCIAL SOFTWARE

Social software (also referred to as "social networking sites" or "social media sites") is employed by individuals of all walks of life round the globe. They are outlined as web-based services that permit people to construct a public or semi-public profile inside a finite system, articulate a listing of different users with whom they share affiliation and looks [8]. Social software at its core relies on supporting people to interact socially and to attain their personal goals, alongside those who have similar interests. It works bottom-up: individuals sign up to a system and form communities through personal selection and actions. Social software isn't a new development because social computing, groupware and similar ideas emerged in the scientific literature from the 1980s. Social software has been the wide diffusion of internet usage that has dramatically enhanced its popularity [9]. Social software desires to arrange itself into teams and to collaborate by advancing personal interests contrasts with additional ancient approaches wherever individuals are placed into organizationally or functionally-defined groups [3]. The utilization of Blogs, Wikis, media-sharing services, and different social software, has been shown to make exciting new learning opportunities for individuals, and to support creation

of social networks and communities of practice among company workers [10].

Through Social software's, like Facebook, Twitter, Instagram and LinkedIn, people are given the chance to share their valuable thoughts, data with a wider audience with completely different demographics. They are not software as such but internet services. Most software tools of social networks are hosted remotely and may be accessed from anyplace with an internet connections. The key feature with any social software is that it is straightforward to use and is typically free at the purpose of use. Tagging is another key feature of most social software that helps users manage their resources and establish different users with similar collections or interests.

Social networking is the building of on-line communities. On-line social networking services give a spread of the way for members to act from emailing to instant electronic messaging to photograph tagging. The foremost well-liked sites offer how to interact with friends through multiple interaction ways [9]. Social networks can also be viewed as, pedagogical tools that stem from their affordances of information discovery and sharing [3]. A well-known social networking service like Facebook is a Web 2.0 service that is used in for online communication and interaction.[9] Some of the Examples of Social Software include:

*A.* RSS feeds

Really Simple Syndication (RSS) is a means of communicating data in a format that Feed Readers or News Reader softwares will understand. It has become a very important means of using social software effectively. Most blogs have RSS feeds to which users subscribe to get information of their choice [11].

*B.* Blogs

Blogs are a type of social software. Blogs are a transparent technique of making an online web site that is updated, often on a daily basis with "posts". Blogs often take the shape of an online journal and typically only have one main author. Blog software are often created and hosted on a server, however there are many blog suppliers who host the service without charge. Options that blogs embrace are: commenting facilities to all other online users to participate during a discussion; tagging to associate postings with a keyword or topic; and a calendar, thus postings will be retrieved by date. It's sometimes possible to subscribe to an RSS feed from a blog.

*C.* Social bookmarking and resource sharing

Social bookmarking tools allow users to store their bookmarks or web Favourites remotely on a web site. These bookmarks can be accessed from any PC connected to the Internet instead of storing the bookmark within a particular browser. Sites like del.icio.us (http://del.icio.us) allows users to store, prepare and share web resources. The project team set timely that this web site would be notably helpful for storing associated sharing any relevant internet sites for the needs of the project and created an account which has links to several valuable resources found.

*D.* Social networking sites

Social networking sites are another common form of social software. Users create a profile and be part of a network, which could be connected to wherever they live, what music they like, wherever they work or wherever they study. Facebook (www.facebook.com) is a common example of a social networking site that primarily have a social function permitting individuals to link with friends, speak on-line and share resources. Skilled social networking sites (such as LinkedIn www.linkedin.com/) and those targeted

on education (such as Elgg http://elgg.net/) also are turning into in style.

*E.* Media sharing

An example is YouTube that is a video sharing web site. Users will transfer video resources to websites and accessible to other users of the site. Tagging makes the resources recoverable by others interested in similar subjects. But alternative sites exist that permit sharing of various types of resources, an example is Slide Share that permits users to share PowerPoint presentations.

*F.* Virtual worlds

Second Life and alternative virtual worlds permit users to make a profile and move around a virtual world. Users will attend events, purchase and sell merchandise and there are currently variety of projects exploring its potential for teaching and learning. It is not clear whether virtual worlds are truly examples of Web 2.0 technologies, although they clearly serve a social networking function.

## IV. WIKIS

The term wiki comes from a Hawaiian term wiki-wiki that mean fast or quick. Wikis are websites that can be interactively altered by any range of individuals using straightforward on-line tools. An incremental version of the website is stored when an edit is saved, making it possible to 'rollback' the site to any previous version if consequent edits have to be compelled to be undone [11]. Wikis offer several authors with the chance of podcasting, deleting, and editing. Variety of wiki platforms are readily out there such a Wikipidea, MediaWiki, SocialText, PBWiki, Wetpaint.com [9]. The foremost known example of a wiki is Wikipedia that is an online reference work which may be altered by any registered user. Options embrace versioning and document history, therefore previous versions will be retrieved; discussion is additionally potential. Wikis can be used remotely or started on a local server [5]. It's platforms usually support two essential functions; open piece of writing and edit preservation. Open editing refers to the power for anyone to simply edit the content on a wiki. Edit preservation refers to the power of wikis to retain all edits to and versions of content contained on the wiki. Taken along, these two functionalities also enable users to "roll back" any changes to the wiki and restore content to a previous version. These two comparatively simple capabilities of wikis, separately and together, will produce a strong and clear collaborative atmosphere [12]. As wikis are free open source software, nobody authorizes the creation of wiki pages and everybody is mechanically licensed to put in writing, edit and publish.

Wiki options are:

*A.* Wiki markup language - format language of wiki is less complicated and more helpful than HTML. Internal links are created without requiring a lot of programs or accessories. Its center of attention is its content: though several wiki pages appear as if an easy HTML, the standard of wiki lies in the content, not the looks.

*B.* It provides version follow-up: Even the dates of the slightest changes on every page of a wiki is recorded for everyone to see. This suggests that any user will review the past on a page at any time.

*C.* Easy guide, supporting of multiple users

*D.* Built-in search feature and easy progress [2]

Wikis are used for supporting varied activities. Some areas of use of wikis are stated below:

*A.* Brainstorming: once an exact project or creation process is started, the participants will be invited to feature articles or opinions.

*B.* Group Projects: Wiki is a special cluster project and may function as a special computer network. Thus, all participants communicate, share resources, and write reports or books.

*C.* Meeting Support: Agendas for special conferences are additional on wikis and therefore the participants are often invited to the conferences.

*D.* Create Lists: it's the most effective approach for organizing the content in reaching the terms at any special space.

*E.* Collection of Links: It allows all users to transfer files, to create comments, etc.

*F.* Writing internet Contents: it's an ideal tool in shaping wiki main ideas and clearly expressing the content whereas writing a collective letter, position, statement and any legal subjects.

*G.* Making Group Portfolios: Any organization will use a wiki so as to be able to load the history and past projects of the organization. Such portfolio could be a strong tool of selling.

In addition to the preceding areas of use, wikis also are information sources which will be benefited in qualitative or quantitative researches. Recording the amendments created by the users, wikis give quantitative knowledge on who created the amendments and when and what kind of amendments were created. Furthermore, it provides qualitative knowledge by logging the discussions of noted users. Additionally to the current, wikis can even offer the qualitative and quantitative information needed for researching the cooperative learning theories, that may not be obtained at laboratory level before [12].

Wikis can be used in teaching method, for instance, as some professors have found varied ways in which to use wikis in their teaching, as a course of study and course management system updated throughout the semester, or as a virtual schoolroom, where students will create, comment on, and edit discussions. Some proprietary course management systems, as well as chalkboard, provide a wiki within them.

Wikis are touted as a collaboration tool that pulls on the input of the many people to craft one product. Wikis are thought-about to be effective tools in several fields as in learning and teaching, they'll facilitate cooperative learning, give cooperative writing, support project primarily based learning, promote creative thinking, encourage essential looking out, support inquiry primarily based and social creative person learning [4].

## V. CONCLUSION

The Web has progressed from being a predominantly read-only medium to one where anyone can publish and share web contents. The literature review highlighted how the different technologies of the Web have enabled it to be more collaborative and participatory. Web 2.0 enables people to collaborate, and share information online. Whilst Social software which are characterized as web-based services allow individuals to create a public or semi-public profile within a constrained system which enables individuals to view other users with whom they share a connection. Wikis on the other hand are websites that can be interactively edited by any number of people using simple online tools. However these are not the only technologies that are contributing to how collaborative the Web is today.

## REFERENCES

[1] C. Nupur, “World Wide Web and Its Journey from Web 1.0 to Web 4.0,” *(IJCSIT) International Journal of Computer Science and Information Technologies*, vol. 5 (6), pp. 8096–8100, 2014.

[2] S. Aghaei, “Evolution of the World Wide Web : From Web 1.0 to Web 4.0,” *International journal of Web & Semantic Technology*, vol. 3, no. 1, pp. 1–10, Jan. 2012.

[3] H. B. PATEL, “Is Social Software a New Mode of Second-Language Learning in the Information Age?” Journal of Education Culture and Society, 2013.

[4] J. Allyson Dooley, S. C. Jones, and D. Iverson, “Web 2.0: an assessment of social marketing principles,” *Journal of Social Marketing*, vol. 2, no. 3, pp. 207–221, Oct. 2012.

[5] J. Secker, “Social software and libraries: a literature review from the LASSIE project,” *Program*, vol. 42, no. 3, pp. 215–231, Jul. 2008.

[6] Umrav Naidu, Ghs, and Singh, “Web 2.0 Applications in library,” 2015.

[7] G. Sharma and L. Baoku, “Customer satisfaction in Web 2.0 and information technology development,” *Information Technology & People*, vol. 26, no. 4, pp. 347–367, Nov. 2013.

[8] D. Agostino, “The Effectiveness of Social Software for Public Engagement,” *International Journal of Engineering Business Management*, vol. 4, p. 31, Jan. 2012.

[9] A. Darwish and K. I. Lakhtaria, "The Impact of the New Web 2.0 Technologies in Communication, Development, and Revolutions of Societies," *Journal of Advances in Information Technology*, vol. 2, no. 4, Nov. 2011.

[10] S. I. Tamrin, A. A. Norman, and S. Hamid, "Information systems security practices in social software applications: A systematic literature review," *Aslib Journal of Information Management*, vol. 69, no. 2, pp. 131–157, Mar. 2017.

[11] T. Judd, G. Kennedy, and S. Cropper, "Using wikis for collaborative learning: Assessing collaboration through contribution," *Australasian Journal of Educational Technology*, vol. 26, no. 3, May 2010.

[12] S. Gokcearslan and S. Ozcan, "Place of Wikis in Learning and Teaching Process," *Procedia - Social and Behavioral Sciences*, vol. 28, pp. 481–485, 2011.

# Performance, Scalability and Quality of Service on Web: Challenges and Open issues.

Evaristo Chishimba
*School of Science, Engineering and Technology*
*Mulungushi University*
Kabwe, Zambia
evamulenga2@yahoo.com

Douglas Kunda
*School of Science, Engineering and Technology*
*Mulungushi University*
Kabwe, Zambia
dkunda@mu.edu.zm

***Abstract:***
The coming of web technology and accessing it through the Internet has led to the creation of a digital society, where almost everything is now connected and is accessible from anywhere. However, despite their widespread adoption, accessing the web sites pause a challenge to the users and the people managing the web sites. This is so because the development of the web sites demands a lot of time, it's costly and needs skilled man power. In this paper, we discuss challenges associated with performance, scalability and quality of service for web. We discuss the solutions which can be employed when developing the web sites. Furthermore, we present an outline of Challenges of website performance and issues which can be avoided when developing the web. In addition, we present some of the open issues in web technology.

***Key words:*** **Web technology, Internet, Performance, scalability and Quality of service.**

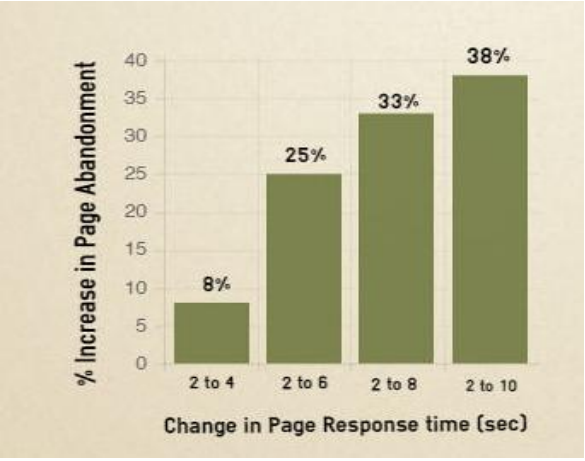
## I. INTRODUCTION

The performance of websites has always been critical [1]. A well-performing website improves the user experience a lot. This translates into your audience coming back and staying longer on your web site [1]. A better performing site is usually higher in the search results because the indexer is having fewer problems with it [1]. Also, mobile and smart devices have become so significant today that a website cannot afford to work at a snail's pace and without the proper optimizations [1].The author in [1] went on and said that most annoying issues which kill website are slow performance and these will result in customers abandoning your website in a blink of an eye as soon as they get the slightest feeling that things are going slow [1]. To emphasis the point the author in [1] gave an example that in 2007, Amazon.com reported that for every 100 millisecond increase in loading time of their website their sales declined by a whole percentage point [1]. Doing an analysis of what typically happens, strange loop Networks (now Rad ware), found that 57% of your visitors will abandon a site after waiting 3 seconds for a page to load [1]**.** Of these, 80% will not return and 40% will even spread the news and let everyone know how bad your website performs [1].

*Figure 1:*

**Every second counts.**

Source: akamai.com [17]



In [17] the delay for the web page to load in 2 seconds or less will translate in 8% of your users or customers abandoning your web site and this will mean losing your customer to other web sites which are performing faster than your web site. A further delay of about 10 seconds of your web page to load will mean that 38 % of the consumers expect a web page to load in 2 seconds or less. Otherwise, users or customers will abandon your web site and this will mean that nearly three quarters or half of your customers will abandon your web site.

According to the authors in [2] [5], Equally, Scalability is very important in Distributed systems in order to operate effectively and efficiently at many different scales, ranging from a small intranet to the Internet. A system is described as scalable if it will remain effective when there is a significant increase in the number of resources and the number of users. The Internet provides an illustration of a distributed system in which the number of computers and services has increased dramatically. As such scalability in the network performance is very vital where the accessing of the web sites is concerned.

Quality of service (QoS) is the description or measurement of the overall performance of a service, such as a telephone or computer network or a cloud computing service, particularly the performance seen by the users of the network, especially when accessing the web sites on the internet [3]. The author in [3] explained that QoS is sometimes used as a quality measure, with many alternative definitions, rather than referring to the ability to reserve resources. Quality of service sometimes refers to the level of quality of service, i.e. the guaranteed service quality. QoS is sometimes used in application layer services such as telephony and streaming video to describe a metric that reflects or predicts the subjectively experienced quality [3].

There are a lot of factors that could impact a website's performance. It could range from simple things that have an easy fix like for example having too small amount of server memory installed. Sometimes, it's out of your hands and the sheer amount of people visiting your website might be bringing your server to its knees [1].

But then, [1] there are also smaller and seemingly insignificant things like the size of your images. The design of the website and also the inclusion of a simple app on the website ruin the entire experience. So, according to the "end user", what's a good online experience then? The author looked at the following:

a. Navigation should be easy and clear.
b. The content should be simple to understand
c. It should be optimized for more than just your own browser.

According to the author this makes it obvious that no matter the size of your company, you should not overlook your website performance. As a result, it should deliver instant gratification. [1].

## II. WHAT ARE CHALLENGES FOUND IN WEB TECHNOLOGY

Author in [1] explained that, so many websites fail to score well on all of these requirements mentioned above. The author further explained that Webmasters responsible for maintaining highly optimized websites often times ignore crucial elements of the site's design, killing website load times. In the end, it usually is the simple mistakes that paralyze so many websites in the organizations.

The author in [1] also looked at some of the most common challenges found in web site development and these are the following:

*A. Speed optimization is completely neglected while building the website*

So many business owners do not feel the need to optimize their website during development. They only feel that need after the website has gone live and they are not getting any organic traffic. By then, the site is so complex that thorough optimization requires them to basically start over. Websites should be developed with performance optimization strategies in mind. Saves time and money [1].

*B. Cheap web hosting service*

Cheap web hosting is only cheap at first sight. With billions of websites online, the web hosting space is crowded. Mediocre web hosts sell cost effective services while sacrificing quality. They can do this by hosting thousands of websites on a single server. This slows web pages down so much that the loss in revenue far outweighs the cost reduction [1].

*C. Too many plugins and widgets*

Author elaborated that additional features that often come in the form of plugins and widgets put a burden on your website's performance. Adding a small widget can easily add a whole 2 seconds to the page load time. Also, plugins that stream a large amount of data to perform complex operations can have a huge impact. Ultimately, they can even reduce the functionality of the web page [1].

*D. Advertisements*

Advertisements slow things down. Your page has to load stuff from all over the internet to show it to your visitor, so don't overdo it [1].

*E. Designs with large hi-res images*

Beautiful design themes sometimes come at a price. By using large high-resolution images, everything looks great but some pages might slow down significantly. It may seem tempting to developers and site owners, and sometimes it makes sense. But more often than not, they end up only costing bandwidth and make you lose customers when your web sites are not performing well [1].

*F. Websites not optimized for mobile users*

Mobile is everywhere, and your website should be ready for that. Websites that are not optimized for mobile users usually suffer from issues like bloated graphics, non-playable videos and irrelevant cross-linking [1].

Although the mentioned issues seem pretty obvious, we've seen many websites facing exactly these problems and challenges. As a matter of time or money crucial factors like design, usability or website speed suffer [1].

## III. WAYS IN WHICH PERFORMANCE CHALLENGES CAN BE RESOLVED IN WEB DEVELOPMENT

*A. Poorly Written Code*

Poorly written code can lead to a host of web application issues including inefficient algorithms, memory leaks and application deadlocks. Old versions of software, or integrated legacy systems can also drag performance down. Make sure your teams are using all the tools at their disposal from automated tools like profilers to best programming practices like code reviews [4].

*B. Unoptimized Databases*

An optimized database allows for the highest levels of security and performance, while an unoptimized database can destroy a production application. Missing indexes slow down the performance of SQL queries which can drag down an entire site. Be sure to use scripts and file statistics to check for any inefficient queries [4].

### *C. Unmanaged Growth of Data*

Data systems tend to degrade over time. Developing a plan to manage and monitor data as it grows is indispensable to your web performance success. The first step is deciding who is accountable for data growth in your business. From there, your team will need to research and determine the appropriate storage for your data needs. Look at all your options, from databases to caches to more sophisticated layered storage solutions [4].

### *D. Traffic Spikes*

We generally think of increased traffic as a good thing. However, anyone who has experienced major traffic spikes after a marketing promotion or viral video knows what can happen when you are not properly prepared for them. Planning ahead is key, and set up an early warning system through simulated user monitoring systems like NeoSense [4].

### *E. Poor Load Distribution*

Poor load distribution can cause slow response times by incorrectly assigning new site visitors to bogged-down servers instead of others with cycles to spare. If too many people are on the same server, they're going to experience problems, even if the overall system is well under capacity. It is imperative to test with a product like NeoLoad as it will help you find any infrastructural weaknesses at hand [4].

### *F. Default Configurations*

Systems must be properly tuned. While default configurations make it easy to get new components up and running, they are not always appropriate for your web applications in a live production environment. Every setting should be checked, review thread counts, allocate memory and permissions. Confirm that all configuration parameters suit the demands placed on your web application [4].

### *G. DNS, Firewall, and Network Connectivity*

DNS queries make up the majority of web traffic. That's why a DNS issue can cause so much trouble, preventing visitors from accessing your site and resulting in errors. Likewise, network connectivity and firewall efficiency are crucial for access and productivity. By using DNS monitoring tools safeguards the problems at hand. Also, revise switches, check VLAN tags, and distribute tasks between servers. These are just a few ways to troubleshoot these types of performance issues [4].

### *H. Troublesome Third-Party Services*

If your users are experiencing problems, it's essential to determine if the problem is on your side or that of the third-party. If you decide to continue using the third-party service, make some design changes to protect your site from at least some of the effects of a third-party service issue [4].

### *I. Shared Resources and Virtual Machines*

Just about every web application today relies on virtual machines for everything from scalability to management to system recovery. However, sometimes the way these virtual systems are organized can result in problems where one is down and it will affect the others. After all, contention is bound to happen. You should monitor the systems closely so that if one VM is causing problems, you can easily deal with the side effects quickly [4].

### *J. The Domino Effect*

Make sure you realize that a failure in one location may affect other spots in ways you would not necessarily think of. Problems compound upon themselves, making it hard to determine what is really going on. You have to train your team to find root causes, back tracing through problems to find the real culprit [4].

*K. Integrating Web Application*

According to the author in [19] in present world, most business apps are more diversified and prefer to live outside the firewall. These businesses might use in-house BI tools, SaaS-based CRM system or hosting websites on the cloud. Although this may improve flexibility, still it causes a challenge for web developers to create a web application that can easily integrate with other applications. The author went further and explained that the solution to overcome this challenge is to employ the following measure that every integration point should involve understanding the coding, API, testing, and logging to troubleshoot any challenge.

*L. The talent challenge*

In [20] Web application development is becoming more complex and it is evolving faster than ever before. Developers now need an ever increasing and ever changing skill set of employees. But the problem is finding the wed developers with modern skills. The question is how do you find developers that understand security, integration, responsive design? How do you bridge the skills gap without hiring a dozen of new employees? We are in a talent challenge now [20].

## III SCALABILITY

In [5] the authors have described the system as scalable if it will remain effective when there is a significant increase in the number of resources and the number of users. The authors have explained that in order to achieve scalable in the system, the following factors should be taken into account and these are the following:

*A. Controlling the cost of physical resources:*

As the demand for a resource grows, it should be possible to extend the system, at reasonable cost, to meet it. For example, the frequency with which files are accessed in an intranet is likely to grow as the number of users and computers increases. It must be possible to add server computers to avoid the performance bottleneck that would arise if a single file server had to handle all file access requests. For example, if a single file server can support 20 users, then two such severs should be able to support 40 users [5].

*B. Controlling the performance loss:*

Consider the management of a set of data whose size is proportional to the number of users or resources in the system, for example the table with the correspondence between the domain names of computers and their Internet addresses held by the Domain Name System, which is used mainly to look up DNS names such as www.ltbc.edu.com. Algorithms that use hierarchic structures scale better than those that use linear structures. But even with hierarchic structures an increase in size will result in some loss in performance. For a system to be scalable, the maximum performance loss should be reduced. In general, algorithms should be decentralized to avoid having performance bottlenecks [5].

*C. Preventing software resources running out:*

An example of lack of scalability is shown by the numbers used as Internet addresses (computer addresses in the internet.) For this reason, the new version of the protocol is using I28 bit Internet addresses. However, the authors said to be fair to the early designers of the Internet there is no correct solution to this problem. It is difficult to predict the demand that will be put on a system years ahead. Moreover, over-compensating for future growth may be worse than adapting to a change when we are forced to. Large Internet addresses occupy extra space in messages and in computer storage [5].

D. *Some shared resources are accessed very frequently:*

In [5] the authors explained that many users may access the same Web page, causing a decline in performance that caching and replication may be used to improve the performance of resources that are very heavily used. Ideally, the system and

application software should not need to change when the scale of the system increases, but this is difficult to achieve [5]. The issue of scale is a dominant theme in the development of distributed systems [5]. A content distribution network (CDN) is globally distributed network of proxy servers deployed in multiple data centers [3] [5]. It simply means that instead of using a single web server for the website, it allows us to use a network of servers [3] [5]. As a result of this, we see benefits of using CDN on our networks and requests on the server will be routed to different servers so as to balance traffic and files are divided on different CDNs as a result there will be no queuing and wait for downloading different files like images, videos, text, on the network [3] [5].

## IV. QUALITY OF SERVICE

In [3] Quality of service is described as a measure of the overall performance of a service, such as a telephony or computer network or a cloud computing service, particularly the performance seen by the users of the network when accessing the web sites in the organization or a business setup.

Multimedia applications demand the timely delivery of streams of multimedia data to end-users. Audio and video streams are generated and consumed in real time, and the timely delivery of the individual elements (audio samples, video frames) is essential to the integrity of the application. In short, multimedia systems are real-time systems they must perform tasks and deliver results according to a schedule that is extremely determined. The degree to which this is achieved by the underlying system is known as the quality of service (Qos) enjoyed by an application [5].

Scalability is not about performance or making good use of computing power and bandwidth [3]. But it is about load balancing between the servers. That is why when the load increase there will be more traffic on the page, so additional servers must be added to balance it [3]. At the same time we should not just throw the entire load on a single server but we should design the software in such way that it can work on a cluster of servers [3]. Service-oriented architecture (SOA) can also help in improving the scalability when more and more servers are added. SOA gives us the flexibility to change easily [3]. Service oriented architecture is a design where application components provide services to other components through the communication protocol, basically over a network [3].

### *A. IP and Ethernet efforts*

Authors in [6] [7] have explained that unlike single-owner networks, the Internet is a series of exchange points interconnecting private networks. Hence the Internet's core is owned and managed by a number of different network service providers, not a single entity. Its behaviour is much unpredictable. Therefore, research has continued on QoS procedures that are deployable in large, diverse networks.

The authors in [6][7] have further explained that there are two principal approaches to QoS in modern packet-switched IP networks, a ***parameterized*** system based on an exchange of application requirements with the network, and a ***prioritized*** system is where each packet identifies a desired service level to the network**.** ***Integrated services*** ("IntServ") implements the parameterized approach. In this model, applications use the Resource Reservation Protocol (RSVP) to request and reserve resources through a network. Whereas, ***Differentiated services*** ("DiffServ") implements the prioritized model. DiffServ marks packets according to the type of service they desire. In response to these markings, routers and switches use various queueing strategies to tailor performance to expectations. Differentiated services code point (DSCP) markings use the first 6 bits in the ToS field (now renamed as the DS Byte) of the IP(v4) packet header [6][7].

At the Media Access Control (MAC) layer, VLAN IEEE 802.1Q and IEEE 802.1p can be used to distinguish between Ethernet frames and classify them. Queuing theory models have been developed on performance analysis and QoS for MAC layer protocols [6] [7].

One compelling example of the need for QoS on the Internet relates *to congestion collapse*. The Internet relies on

congestion avoidance protocols, as built into *Transmission Control Protocol* (TCP), to reduce traffic under conditions that would otherwise lead to "meltdown". QoS applications, such as *VoIP* and *IPTV*, require largely constant bitrates and low latency, therefore they cannot use TCP and cannot otherwise reduce their traffic rate to help prevent congestion. QoS contracts limit traffic that can be offered to the Internet and thereby enforce traffic shaping that can prevent it from becoming overloaded, and are hence an indispensable part of the Internet's ability to handle a mix of real-time and non-real-time traffic without meltdown[8].

### *B. End-to-end quality of service*

Authors in [9] [10][21] have described End-to-end quality of service that can require a method of coordinating resource allocation between one autonomous system and another. The Internet Engineering Task Force (IETF) defined the Resource Reservation Protocol (RSVP) for bandwidth reservation, as a proposed standard in 1997. RSVP is an end-to-end bandwidth reservation protocol. The traffic engineering version, RSVP-TE, is used in many networks to establish traffic-engineered Multiprotocol Label Switching (MPLS) label-switched paths. The IETF also defined Next Steps in Signalling (NSIS) with QoS signalling as a target. NSIS is a development and simplification of RSVP.

### *C. Circumvention*

Cryptography network protocols such as Secure Sockets Layer, I2P, and virtual private networks obscure the data transferred using them. As all electronic commerce on the Internet requires the use of such strong cryptography protocols, unilaterally downgrading the performance of encrypted traffic creates an unacceptable hazard for customers. Yet, encrypted traffic is otherwise unable to undergo deep packet inspection for QoS [11].

### *D. Doubts about quality of service over IP*

The Internet2 project found, in 2001, that the QoS protocols were probably not deployable inside its Abilene Network with equipment available at that time. Equipment available at the time relied on software to implement QoS. The group also predicted that "logistical, financial, and organizational barriers will block the way toward any bandwidth guarantees" by protocol modifications aimed at QoS. They believed that the economics would encourage network providers to deliberately erode the quality of best effort traffic as a way to push customers to higher priced QoS services. Instead they proposed over-provisioning of capacity as more cost-effective at the time [12].

## V. OPEN ISSUES IN WEB TECHNOLOGY

### *A. Community Networks*

In [13] the author discuss how the Community has established networks which are also referred to as "Community Networks" (CN) which have existed for many years and, provide a sustainable solution to address the connectivity gaps that exist in urban, remote, and rural areas around the world. In Africa, where these gaps are more evident, a recent survey was able to identify 37 community networks initiatives in 12 African countries, of which 25 are considered active. Community Networks are a key way to address this connectivity gap, says the Internet Society, a global non-profit dedicated to ensuring the open development, evolution and use of the Internet through the accessing of the web sites which are developed in developing countries like Zambia and others [13].

### *B. The Open Web*

The author in [14] has described the Open Web has an open and democratic Web, which is *interoperability*. The author has defined *Interoperability* that means the systems, software, sites, and applications which are developed and work with each other without conflict.

In [14] the author has explained that the root principles of the Web are to create and share content regardless of the following:

(a) Platform
(b) Operating System
(c) Data format
(d) Language/Location
(e) Ability or disability

In [14] the Web's creators, Tim Berners-Lee and others such as Robert Cailliau, saw the Web as a social environment in as much as a technical one. The first and perhaps most critical aspect to the Open Web is that it is *decentralized*. This means that it is not controlled by any single individual, company, or organization. Rather, it belongs to anyone who wants to use it [14].

Table 1: Technologies considered being Open Web [14]

| Open web technology | Version | Organization | Description |
|---|---|---|---|
| HTML5 | HTML5 | WHAT WG / W3C | <video> <audio> and |
| Media APIs | | WG / W3C | <canvas> |
| HTML5 APIs | HTML5 | WHAT-WG / W3C | Variety of application programming interfaces for a wide range of script access to application and document elements. |
| Micro formats | All | microformats.org | Extending existing mark-up to make it more human-readable and open. |
| WAI-ARIA | All | W3C | Web Accessibility Initiative / Accessibility for Rich Internet Applications |

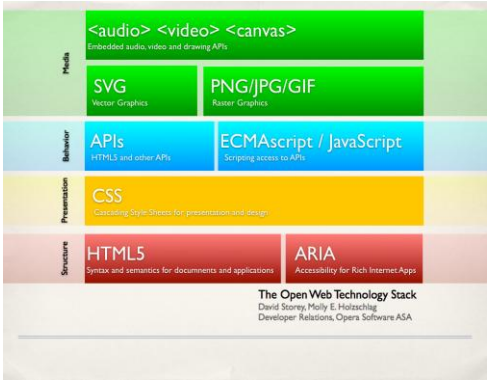
**OPEN WEB TECHNOLOGIES: FRONT END**

In [14] open web technologies are those that conform to the root and core principles of Web standards and Open Web. Front end development often has to play well with proprietary server technology. This is critical because technologies on the front end (Table 1) tend to be very adaptable, in large part due to their transparency.

While the majority of these technologies come from the W3C, the WHAT-WG and microformats.org are grass-roots organizations and work in the way compliant with the ideals of the Open Web. They are open to anyone who has the interest and desire to participate, a testament to the power of individuals to help build the very technologies with which we work [14].

*Figure 2*

***The Open Web Technology Stack:***




*A visualization of Open Web [14]*
**Source:** Opera software ASA

The authors in [14] said that there are four layers namely: *structure, presentation, behavior*, and *media.* As illustrated in figure 2 above. They went further and explained that many readers will be familiar with the three layer model, but along with HTML5 has come open APIs for embedded media.

On the first level, ***structure***, we see HTML5 along with ARIA. This provides the syntax, semantics, and accessibility features for the content and features of your websites and applications. Above structure is ***presentation,*** which is of course the design layer via CSS. Along with the content and style, there's the ***behavior*** layer [14].

It is in this layer we begin to see the influence of HTML5 in the Open Web, where HTML5 APIs (as well as other APIs that are related to but not HTML5, such as geolocation) and scripting sit at the same level. This is the real shift that HTML5 brings: incredibly rich scripting options we would never had in browsers before [14].

Finally, there is the ***media*** layer, where we also see the influence of HTML5 in the form of specific media APIs: video, audio, and canvas. SVG is *Scalable Vector Graphics* and can be used to create in-browser vector graphics and animation. Not new to the game, SVG is now becoming more desirable as the IE browser begins to mature [14].

The authors in [14] have explained that this stack is only one of many potential representations. Microformats, for example, could easily be added to the structural layer. In fact, their colleague at Opera Software, David Storey, who began this visualization and he has modified it from there. In the spirit of the Open Web, Paul Irish from Google saw it, asked for it, and now it's being modified and used to create a variety of visualizations.

*C. Universality*

In [15] the author explained that when he designed the Web protocols, he had already seen many networked information systems fail because they had made some assumptions about the users that would using a particular type of computer.

In [15] the goal was that anyone should be able to publish anything on the Web and so it had to be universal in that it was independent of all these technical constraints, as well as language, character sets, and culture.

Net Neutrality is essential to an open web in that, fair democracy is close to the principle of universality which means that no permission is needed from a central authority to post anything on the Web, since there is no central controlling node, and no single point of failure. This is critical to the Web's growth and critical to its future [15].

*D. Security*

In [16] the author explained that there must be security in the communication of data between application systems and other software to avoid improper access by authorized users. Trust or the lack of trust, is the main factor that is blocking the adoption of rapidly evolving Web technology paradigm such as software as service (SaaS) and data distribution services.

*E. Interface*

In [17] the web was a completely different place. By then Smartphones did not exist. Simpler and customer oriented web application are highly expected now. Sometimes it is that small UI elements that make the biggest impact. In the era of Smartphones, websites should be responsive enough on the smaller screens. If your web applications frustrate or confuse users, then it will be difficult to maintain your customer's loyalty for your website services.

*F. Web technologies are central to success*

According to author in [18] 85% of the people who stated that they were successful with their web and mobile applications also said that web technologies (HTML5, CSS and JavaScript) were central to their success. The quest for improving development efficiency was the number one reason people said they were moving away from native development in favor of web tech. *Quality*, *time*, and *cost* factors were also cited as key factors for moving to web technology.

## VI. CONCLUSION

In this paper, we have discussed the performance, Scalability and the Quality of Service on web. The authors have further discussed the challenges associated with performance of the network system and issues which can be avoided when developing the web. The authors have also described various protocols which have been developed to improve the performance and scalability in the network. However, more research in this field needs to be carried out in order to improve the quality of service. At the moment what mitigates is the Open Web, as it is built to be flexible, adaptive, and address the needs of today with the needs of the future.

As the networks and web technology advances there will be more security threats on our web sites and organization dealing on online businesses should ensure that they safe guard the data and information on the servers. The following are some of the security threats which should be prevented and these are: Cross-Site Scripting, Phishing, Cross-Site Request Forgery, Shell Injection, Session Hijacking, SQL Injection, and Buffer Overflow to mention a few. In order to also prevent these threats the websites should be carefully coded.

## VII. ACKNOWLEDGMENT

The authors would like to thank the people who helped in the writing of this paper for their support.

REFERENCE

[1] Thomas Peham “The most annoying issues that will kill your website performance”
https://usersnap.com/blog/annoying-issues-will-kill-website-performance/

**[2]** Neotys
The 10 Most Common Web App Performance Problems
https://www.neotys.com/blog/10-most-common-web-app-performance-problems/

**[3]** Quality of service – Performance and challenges of service – oriented Architecture.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5375822/

**[4]** *Menychtas Andreas (2009)*
https://menychtas.com/

**[5]** *George Coulouris, Jean Dollimore, Tim Kindberg (2003) “ Distributed Multimedia Systems” Distributed Systems , Concepts and Design, Third Edition, pp.614-622*

**[6]** Ben Erwin (December 16, 2008). "How To Manage QoS In Your Environment, Part 1 of 3". *Network Performance Daily video*. NetQoS. Retrieved October 15, 2011.

**[7]**Shi, Zhefu; Beard, Cory; Mitchell, Ken (2009). "Analytical Models for Understanding Misbehavior and MAC Friendliness in CSMA Networks". *Performance Evaluation*. **66** (9–10): 469. doi:10.1016/j.peva.2009.02.002.

**[8]** Ben Erwin (December 16, 2008). "How To Manage QoS In Your Environment, Part 1 of 3". *Network Performance Daily video*. NetQoS. Retrieved October 15, 2011.

**[9]** Bob Braden ed. L. Zhang, S. Berson, S. Herzog, S. Jamin (September 1997), *Resource ReSerVation Protocol (RSVP)*, IETF, RFC 2205

**[10]** “End –to- end quality of service Support over heterogeneous networks” *Project description*. European Community Research and Development Information Service. Retrieved October 12, 2011.

**[11]** Benjamin Teitelbaum, Stanislav Shalunov (May 3, 2002). "Why Premium IP Service Has Not Deployed (and Probably Never Will)". *Draft Informational Document*. Internet2 QoS Working Group. Archived from the original on September 12, 2010. Retrieved October 15, 2011.

**[12]** Andy Oram (June 11, 2010). "A Nice Way to Get Network Quality of Service?". *Platform Independent column*. O'Reilly. Archived from the original on September 12, 2010. Retrieved October 15, 2011.

**[13]** Third Summit on community Networks in Africa is held from 2 to 7 September 2018, in Wild Lubanzi Trail Lodge, Eastern Cape, South Africa.
http://www.internetsociety.org/event/summit-community-africa/2016

**[14]** Molly Holzschlag (August 01, 2011) “Understanding the open web stack” Test and Monitor. Post August 01, 2011.
https://smartbear.com/blog/test-and-monitor/understanding-the-open-web-stack/

**[15]** Tim Berners-Lee (January 2013) Open" conference as part of the "timbl down under" tour in January 2013. The Conference was on the theme of "Open" in general. October 2013

https://www.w3.org/DesignIssues/Open.html

**[16]** PK Sreenivasaiah ( 2010 ) ***open issues*** of database and ***Web*** application ... Published online 2010 Dec 3. Prepublished online 2010 Sep 28. doi: 10.3389/fphys.2010.00147

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3059952/

**[17]** Maruti Techlabs "challenges-in-web-application-development"

https://www.marutitech.com/5-challenges-in-web-application-development/

**[18]** Kristin Brennan (June 25, 2015) "Trends and challenges in web application development" The Rise of *Web Technology* 2015.

https://www.sencha.com/blog/trends-and-challenges-in-web-application-development

**[19]** TBI Infotech Posted on (January 6, 2016) **"**Commonly Encountered Challenges in Web Development Projects and their Solutions" 2016.

https://www.brihaspatitech.com/.../common-encountered-challenges-in-web-developm**...**

**[20]** Joe Stangarone (May 2015) "7 web development challenges you can not ignore"

https://www.mrc-productivity.com/blog/2015/05/7-web-development-challenges-you-cant-ignore/

**[21]** Vivian Mwenda Mwale and Mbuyu Sumbwanyambe, International Journal of Computer Science and Mobile Computing, Vol.5 Issue.4, April-2016, pg. 584-591

https://ijcsmc.com/docs/papers/April2016/V5I4201692.pdf

# Active learning environment: A Comparative Analysis of Web Services (SOAP, RESTful, WSDL, UDDI)

Ng'andu Wilson Mwiiya
*School of Science, Engineering and Technology*
*Mulungushi University*
Kabwe, Zambia
ngandumwiiya@gmail.com

Douglas Kunda
*School of Science, Engineering and Technology*
*Mulungushi University*
Kabwe, Zambia
dkunda@mu.edu.zm

***Abstract*—Attention to Web services has swiftly increased since their introduction. Exchanging information among applications in an acceptable way is the main objective of web services. This communication among applications has brought a need for uninterruptible and continuous web services which is centered on the SOAP and REST standard. Network traffic and processing delays are some of the limitations of SOAP communications. These limitations can be overcome by the use of the RESTful architecture because REST is lightweight. This paper will outline the motivation for the introduction of web services. The concept of an active learning environment is discussed followed by the use of web services in active learning environment. A brief background of the two significant types of web services is given and comparisons of these two frameworks based on their service discovery, interface, security and general performance. Each of these types of web services has its own value, advantages and disadvantages. Therefore, one has to understand the circumstances in which each of these designs ought to be used.**

**Keywords—*Web Service, SOAP, REST, RESTful, UDDI, WSDL, active learning environment***

## I. INTRODUCTION

Before the coming of web services, technologies like java RMI (Remote Method Invocation), DCOM (Distributed Component Object Model) and CORBA (Common Object Request Broker Architecture) were made use of in developing server and client applications. These technologies were anchored on the DCE (Distributed Computing Environment)/RPC (Remote Procedure call) procedural framework. Both the client and the server were highly coupled to each other. This technology had security and compatibility challenges. A standard means of dispensing services across the internet had to be introduced hence the emergence of web services.

The term Web service is used a lot these days, although sometimes with different meanings. Definitions of Web services range from general to specific and restrictive. The W3C (World Wide Web Consortium) defines a Web service as ''a software system designed to support interoperable machine-to-machine interaction over a network". It states clearly that the service should be identified by a Universal Resource Identifier (URI), and should have an interface that can be described in machine processable format (WSDL) and other systems interact with the Web service and the messages are sent using HTTP with eXtensible Markup Language (XML). Three things are stressed by the W3C definition and these are that Web services should be able to be defined, described, and discovered and clearly states that communication is conveyed by Internet protocols. In other words, Web services can be integrated into more complex distributed applications [1]. A web service is a software application designed to support interoperability in a standardized way between unlike applications running on a variety of platforms. This is done over a standard web protocol like HTTP. Different formats can be used by web services to provide data but the most common ones are XML and JSON. Another program that receives the data can easily recognize and parse the data in these text based formats. In this architecture, [2] adds to say clients need not have previous awareness of the web services before they actually use it. In this way, web services remain platform independent and are loosely coupled.

The two types of web services used to make web applications are discussed in this paper. These web services rely on the Simple Object Access Protocol (SOAP) and Representational State Transfer (REST) principle. Web services that rely on SOAP are called SOAP-based Web services and those that rely on REST are called RESTful Web services. To define SOAP in SOAP based web services, XML is used. The REST principle intended for distributed hypermedia systems is used for RESTful web services. REST is not a technology but a network architectural style. This is so for the reason that RESTful web services use HTTP protocol [2]. This paper aims at supporting active learning environment through the use of web services. A brief background of the two significant types of web services is discussed followed by a comparisons of these two frameworks based on their service discovery, interfaces, security and general performance.

## II. ACTIVE LEARNING ENVIRONMENT

The engagement of learners in the learning process through active learning environment is a challenging thing but is of importance in today's education system. It is agreed by researchers [3], that to have learners actively engaged means that the leaners must use higher order taxonomy of thought like analysis, synthesis and evaluation and it encourages exploration and development of skills [3]. Including learning strategies that support this kind of learning is of significant importance in this computer generation. The traditional teacher centered approach has challenges of not actively engaging learner in the learning process. Web services can address this challenge through employment of different application that can actively engage the learner. An active learning environment entails that learners can actively participate in instructional activities. A learning environment must be one which is exciting and active so that learners are all the time challenged and engaged in the process of learning. It is believed that the provision of an active learning environment can lead to increased student learning [4].

## III. USE OF WEB SERVICES IN ACTIVE LEARNING ENVIRONMENT

The learning style of the twenty-first century is being changed and formed by the introduction of digital tools and applications running web services. Students nowadays want a learning experience where they are actively involved and one which is social and has rich media support. Many of these social software tools allow autonomy and engagement of the learner [5]. By so doing learners can exchange ideas in social communities and knowledge is created as learners take the active roles in their learning [6]. In an active learning environment, learners do not take the passive role, but take an active part in their learning where the three learning domains – skills, knowledge and attitudes are aroused [4]. Web services through the use of social applications are able to provide a platform for conversation, reflexive dialogue and collaborative content generation. The use of web services can shift control from teacher centered approach to learner centered approach. This is done by engaging the learner in social networks that offer virtual learning [6]. Therefore the use of web services offer an active learning environment as they offer a dynamic, collaborative and a conducive environment for problem solving in the learning process [7].

Web services allows learners to take part in collaborative learning activities through applications like blogs, social networking sites, video sharing etc. An active learning environment [7] can be realized by applications that support reading and writing activities. Example of such applications would be Moodle, blogs and wikis. Learning can also be realized through applications that allow learners to interact among themselves and with the system to discover knew knowledge. Learners can solve problems in groups and provide feedback to each other.

Learning happens when learners are involved in groups [16]. Web services can be used to support such kind of learning through the use of social learning applications.

Applications like Moodle can be used to facilitate active learning. Documents and lecture contents can be offered on the Moodle platform. Learners can collaboratively do activities in groups. Collaborative work can also be done through online video conferencing. Learners can ask and answer question using forums and social networks. Learners can also use the chat facility to ask lesson facilitators questions [20].

## IV. SOAP

SOAP is a standard web service communication protocol. It was created by Dave Winer et al in collaboration with Microsoft in the year 1998. This protocol addresses needs of the enterprise market. SOAP Web services are based on the concept of Service-oriented architecture (SOA). SOA is as a result of distributed computing which enables building distributed systems that expose their functionality as services [8]. Web services that use the SOA are designed to enable asynchronous interactions between the client and the server. Service Oriented Architecture is a commonly used standard reference model for service oriented computing and is very useful in developing Web Services. According to [9], "Web services are loosely coupled software components delivered over Internet standard technologies." In other words, Web services are self-describing and modular business applications that expose the business logic as services over the Internet through programmable interface and where IP can be used for finding ways to subscribe and invoke those services


FIGURE 1: web service architecture

As shown in figure 1 above, adapted from [2], the SOAP-based web service architecture involves three parts: the service provider, the service registry, and the service requestor. A service provider is simply the provider of a web service. The service provider creates a SOAP-based web service with its service definition. The provider then publishes the service description in the service registry based on the Universal Description Discovery and Integration (UDDI) standard specifications. The registry is a place where developers can publish new services or locate existing ones.

### *A. Service Discovery*

The process of finding an appropriate service that meets the clients' requirements by locating a web service provider is called web service discovery. After the Web service has been published, a service requestor who can

be any consumer of a web service finds the service through the UDDI. This is done through querying the service registry. The UDDI registry makes a WSDL service description available to the service requestor with a URL (uniform resource locator) pointing to the service. The service requester then retrieves the service description, and uses it to bind the service implementation, and then invokes the service [10]. The service registry is the one that makes it possible for on-line discovery of services. And the exchange of information between the three web service entities rely on XML and use SOAP protocol.

To exchange SOAP messages between service provider and service requestor, transport protocols like Simple Mail Transfer Protocol (SMTP) or File Transfer Protocol (FTP) can be used although the most commonly used one is HTTP. The choice of using HTTP comes with the advantage that it makes it easier for SOAP model to pass through firewalls and proxies minus altering the SOAP protocol.

SOAP messages are made up of three parts; an envelope, a header and the message body. A SOAP message is identified as an XML document by the envelop element [11]. The header may be included or can be left out but it contains application-specific information like authentication data. And the actual SOAP message is found in the body element.

### *B. Interfaces*

Each web service must have a standard interface. The interface enables a web service to communicate with other web services for reasons of providing different functionalities. The Web Services Description Language (WSDL) is used to publish the service descriptions. WSDL provides information on how a web service can be used, including a description of the service methods and the binding information.

### *C. Performance*

An advantage about SOAP is that while using one format, a message can be sent through several middleware systems regardless of the protocol the middleware system is using. This means that the change of transport protocol along middleware while delivering SOAP messages from one point to the other cannot cause a problem. According to [10], SOAP can be appropriately used in a distributed computing environment and not in point-to-point integrations whereas REST can be appropriately used in a point-to-point integration but not in a distributed computing environment.

XML is the data format used by SOAP. XML enable the interoperability of data between applications running on different platforms. On the other hand, [12] points out that the rising use of XML raises concerns because XML is metadata-loaded. This makes processing to be inefficient as it can strain a network and storage space. While offering interoperability to loosely coupled systems, XML sacrifices the performance of SOAP. Another factor that adds to the degraded performance of SOAP is the considerable amount of time that is needed to extract the SOAP envelop from the SOAP packet [12].

### *D. Security*

Unlike REST which does not have a security protocol of its own, WS-Security is used as one of the standards in the SOAP protocol for encrypting messages so that the transfer of data is safe.

## V. REST

RESTful architecture is based on the Resource-Oriented Architecture (ROA) that provides rules and procedures for designing RESTful web services. The essential feature of ROA is the concept of resources. A resource is anything that has an identifier. Each resource has a representation and a URI as the identifier, and may be linked to other resources through hyperlinks [13]. When a client sends a request about a resource, what they receive is a resource representation (XML, JSON, Text, User-defined, etc.). The representation is any useful information about the current state of the resource. The communication protocol used by ROA is HTTP. Resources are accessed and manipulated using HTTP's operations like GET, PUT, POST, and DELETE [13]. REST architecture style is client server architecture where clients send requests to the server and in return the server processes the request and sends back a responses. REST was introduced by Roy Fielding in 2000 to be [14] an architectural style to be used by distributed hypermedia systems. In his academic dissertation Fielding states in summary that REST "provides a set of architectural constraints for designing networked applications that, when applied as a whole, emphasizes scalability of component interactions, generality of interfaces, independent deployment of components, and intermediary components to reduce interaction latency, enforce security, and encapsulate legacy systems." Therefore, REST is a way of developing Web services by following certain constraints between service providers and consumers. These architectural principles focus on system resources. RESTful Web Services are achieved by the use of Web standards (HTTP, XML and URI) and REST principles [13]. The following are the main design principles that are followed by the REST architecture:

There must be **stateless** Communication between clients and servers; with the exception of information that is used for authentication like sessions information, servers are not allowed to store any context information about a client between calls. The client includes all parameters, context, and data needed by the server to generate a response. Apart from simplifying the implementation and design of server-side components Statelessness improves Web service performance by removing the need by the server to synchronize session data with an external application. The server-side is less complicated to design, write, and distribute across load-balanced servers when it remain stateless. Additionally by shifting most of the responsibilities of maintaining state to the client, the stateless service's performance becomes better. According to [2] the server's responsibility in a RESTful Web service, is to generate responses and provide an interface that enables the client to retain the application state on its own.

**Addressability**; REST operations and data are seen as resources that can be identified by using a URI [15]. The

URI offers a way to address resources and to discover services. A resource in this sense is any information that can be named and is referenced, for example a document or a search result. A uniform and standard interface is used to access these resources.

And **Interface uniformity**; the HTTP protocol and its methods like GET, PUT, DELETE and POST form the standard and uniform interface for accessing REST resources. These methods, GET, PUT, DELETE and POST as pointed out by [2] are used to retrieve, update, delete and create resources as shown in table 1. This is considered to be one of the most unique features of the REST architectural style because of the specific set of restrictions it imposes on how components will have to interact [1]. This in return promotes visibility, which reduces the cost of integrating components. It also helps to decouple components. A uniform interface also has the advantage of promoting familiarity (the operations to be exposed are well known) and interoperability. As part of the uniform interface constraint, [16] HATEOAS (Hypermedia As The Engine Of Application State) does not require the client to know the URIs structure but just the entry URI. Other URIs that follow are dynamically discovered during the request, response communication. By so doing the client and the server become loosely coupled.

TABLE I: HTTP Methods and their corresponding CRUD actions

| **HTTP Method** | **CRUDE Action** |
|---|---|
| GET | Retrieve a resource |
| POST | Create a resource |
| PUT | Update a resource |
| DELETE | Delete a resource |

Other design principles that RESTful web services rely on include the following: There must be **Separation of concerns** between clients and servers; this implies that, clients should not be concerned about server specific operations like data storage, just as servers should not be concerned about client specific operations like the user interfaces. The client and the server therefore are loosely coupled.

Clients in using RESTful web service must be able to **cache** responses. Therefore enough cache-related information should be given by the server in each response so that the clients can decide for themselves when to cache the responses. This in return improves performance as it reduces the number of call a client can make to the server.

Because connections might occur through **several communication layers**, clients must be able to tell when they are directly connected to the application server or to an intermediary agent.

When followed, these design principles brings out the simple and lightweight characteristics in a REST application. And applications that follow these REST principles are called RESTful web services.

The Web Application Description Language (WADL) and WSDL2.0 are used to describe RESTful web services [17]. A WADL file sets in clear terms the requests that can be addressed to a service. The WADL file also includes the service's URI and the data the service expects and provides [18].

### *A. Service Discovery*

Service discovery is different in REST than in SOAP. In SOAP a service can simply be discovered by accessing the metadata in the repository. With REST, however, the self-descriptive message constraint requires that each representation contains all the metadata needed. Therefore all a client has to do to discover a RESTful web service is to follow the correct hyperlink.

### *B. Performance*

The performance of the RESTful web service can be determined by its execution speed and its efficiency**.**

- *Execution Speed*

Execution time is the time a web service would take to return a response back to the client. Comparing the response time between REST and SOAP, [10] found that REST had a shorter response time with a better data throughput than its counterpart the SOAP protocol. It is also made clear that in cases of slow data transfer speed or where the network has a large load, SOAP would not be the best choice. In [10] an experiment on response time was conducted by simulating a client-server scenario. The results of the experiment showed that SOAP-XML processing time is 4 to 5 times higher as compared to the REST approach.

- *Efficiency*

Because of some of the principles imposed by the REST protocol like caching and load balancing, the RESTful Web service can be used by large numbers of clients at the same time. It should also be taken into consideration that REST uses JSON which is a light data format or plain texts for the communication of messages which results in reduced server load [10].

Experiments have also been carried out in [19] to evaluate RESTful web service for mobile devices and the findings for the performance evaluation show that it is more advantageous to use RESTful Web services over SOAP based web services because of small message size and a quicker response time. The results of this performance comparison between REST and SOAP clearly showed that there is high performance in RESTful than in SOAP. And for this reason RESTful may be used in many implementations because of its higher flexibility and lesser overheads.

### *C. Security*

The SOAP protocol uses WS-security as one of the standards to ensure authentication of messages that are being transferred. REST does not have its own security protocol. It uses the security mechanisms built in HTTP or HTTPS. If the communication channel used is point-to-point, the Transport Layer Security (TLS) can be used. The use of TSL has challenges when mobile devices are involved in the communication [10].

## VI. Challenges And Issues

Web service offer a fast and flexible way of sharing information among its users and businesses. Web services take advantage of the ubiquity nature of the internet to connect applications, systems and resources so that new business relationships and processes can be supported. [20] Web services can provide access to corporate information that could only be accessed through the use of special software. Along with the many advantages of using web services comes security risks. The security issues associated with web service in a distributed environment include exposure of sensitive and private information to people who are not the intended users of the information.

Further, the adoption of web services in education has not been widespread mainly due to technical and social reasons. Studies conducted show that one of the factors that hinder adoption of learning media is security and privacy in social networked learning. This becomes a major concern in that issues of ownership and control are not easy to handle because content is shared freely [21].

## VII. Comparisons

Table 2 adapted from [22] summarizes the differences of the SOAP based web services and the RESTful web service.

TABLE II: Comparison of SOAP and REST based web services

| SOAP-based WS | REST-based WS | References/ citations |
|---|---|---|
| Focuses on exposing pieces of application logic (not data) as services | Focuses on accessing named resources through a single consistent interface | [10] [2] [15] |
| SOAP-based calls cannot be cached | REST calls can be cached | [2] [15] |
| SOAP cannot make use of REST since SOAP is a protocol and REST is an architectural pattern | REST can make use of SOAP as the underlying protocol for web services, because in the end it is just an architectural pattern | [4] [15] [2] |
| In SOAP, Client-Server interaction is Tightly | In REST, Client-Server interaction is loosely coupled. | [10] [2] |
| The Web Services Description Language (WSDL) is used to publish service descriptions | Web Application Description Language (WADL) and WSDL2.0 are used to describe RESTful web services | [10] [2] [15] [4] |
| SOAP has heavy payload as compared to REST | REST is definitely lightweight as it is meant for lightweight data transfer over a most commonly known interface, the URI | [10] [2] [15] [4] |
| It requires binary attachment parsing | It supports all data types directly | [2] [15] |
| SOAP web services only support XML data format | REST web services support different data formats like XML,JSON, plain text etc. | [10] [14] [4] |
| It consumes more bandwidth and resources because a SOAP response could require more than 10 times as many bytes as compared to REST | It consumes less bandwidth and resources because it's response is lightweight. | [9] [15] [4] |
| Designed to handle distributed computing environments | Assumes a point-to-point communication model- not for distributed computing environment where message may go through one or more intermediaries | [13] [2] [15] |
| Harder to develop, requires tools | Much simpler to develop web services than SOAP | [13] [2] |
| SOAP uses WS-Security standard specification for security | REST uses HTTP (or HTTPS) and the security mechanisms that are built-in to the protocol | [10][13] [2] [15] |
| Is the prevailing standard for web services, and hence has better support from other standards (WSDL, WS) and tooling from vendors | Lack of standards support for security, policy, reliable messaging, etc., so services that have more sophisticated requirements are harder to develop | [10] [2] [15] |

## VIII. Conclusions

Web services are becoming widely used over the internet. Because of this widespread use, the performance of web services is an important factor. Two significant web services were discussed in this paper. The SOAP based web services and RESTful web services. From the literature gathered in the paper, it can be concluded to say RESTful web service is a better alternative than the SOAP based web services in terms of performance. Unlike RESTful web services, SOAP based web services have large network traffic and the size of the message is also large. With regards to the choice of web service to use, the RESTful web services are suitable in implementations that need greater scalability, compatibility and performance. Therefore, if a point-to-point implementation or large-scale availability is needed in a project, REST is the right choice. On the other hand SOAP is suitable in implementations that require security and reliability. Additionally, a lot of business systems require asynchronous data processing requests, one of the SOAP advantages. SOAP is more suitable in large information system integrations like in banking.

We have not considered the security of data in transit or as it leaves or travels on the network. Security issues like confidentiality, integrity, privacy, authentication, and authorization all need to be looked at to give users confidence that their data is safe especially when the platform of web service implementation is the cloud. Many approaches have been put in place for SOAP-based web service security like WS-Security. One challenge is that the WS-Security has not been widely adopted. On the other hand RESTful web services support HTTPS as the default security mechanism which has its limitations. Therefore there must be intensive research with regards to the security of web services so future directions can be put in place to ensure web service security so as to

support secure services[23].Many services are secure and less vulnerable, one towards this is including light weight cryptographic algorithms.

## REFERENCES


[1] Q. Z. Sheng, X. Qiao, A. V. Vasilakos, C. Szabo, S. Bourne, and X. Xu, "Web services composition: A decade's overview," *Information Sciences*, vol. 280, pp. 218–238, Oct. 2014.

[2] S. Mumbaikar and P. Padiya, "Web services based on soap and rest principles," *International Journal of Scientific and Research Publications*, vol. 3, no. 5, pp. 1–4, 2013.

[3] J. Williams and S. J. Chinn, "Using Web 2.0 to support the active learning experience," *Journal of Information Systems Education*, vol. 20, no. 2, p. 165, 2009.

[4] S. A. A. de Freitas, W. C. Silva, and G. Marsicano, "Using an Active Learning Environment to Increase Students' Engagement," in *Software Engineering Education and Training (CSEET), 2016 IEEE 29th International Conference on*, 2016, pp. 232–236.

[5] N. Selwyn, "Web 2.0 applications as alternative environments for informal learning-a critical review," in *Paper for CERI-KERIS International Expert Meeting on ICT and Educational Performance*, 2007, vol. 16, p. 17.

[6] C. McLoughlin and M. J. Lee, "Personalised and self regulated learning in the Web 2.0 era: International exemplars of innovative pedagogy using social software," *Australasian Journal of Educational Technology*, vol. 26, no. 1, 2010.

[7] C. E. M. Luis, J. M. Gutiérrez, and A. M. G. Marrero, "Using mobile devices and internet technologies in problem-based learning: Design of a suitable active and collaborative learning environment in engineering education," in *Frontiers in Education Conference (FIE), 2014 IEEE*, 2014, pp. 1–6.

[8] M. S. Das, A. Govardhan, and D. V. lakshmi, "QoS of Web Services Architecture," in *Proceedings of the The International Conference on Engineering & MIS 2015 - ICEMIS '15*, Istanbul, Turkey, 2015, pp. 1–8.

[9] K. Wagh and R. Thool, "A Comparative Study of SOAP Vs REST Web Services Provisioning Techniques for Mobile Host," *Journal of Information Engineering and Applications*, vol. 2, no. 5, pp. 12–16, 2012.

[10] K. Gottschalk, S. Graham, H. Kreger, and J. Snell, "Introduction to Web services architecture," *IBM Systems Journal; Armonk*, vol. 41, no. 2, pp. 170–177, 2002.

[11] M. Shehab, K. Bhattacharya, and A. Ghafoor, "Web services discovery in secure collaboration environments," *ACM Transactions on Internet Technology*, vol. 8, no. 1, pp. 5-es, Nov. 2007.

[12] G. M. Tere and B. T. Jadhav, "How to improve XML web services performance?," in *Proceedings of the International Conference and Workshop on Emerging Trends in Technology*, 2010, pp. 257–260.

[13] F. Belqasmi, R. Glitho, and C. Fu, "RESTful web services for service provisioning in next-generation networks: a survey," *IEEE Communications Magazine*, vol. 49, no. 12, pp. 66–73, Dec. 2011.

[14] "Fielding Dissertation: CHAPTER 5: Representational State Transfer (REST)." [Online]. Available: https://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm. [Accessed: 30-Aug-2018].

[15] J. Tihomirovs and J. Grabis, "Comparison of SOAP and REST Based Web Services Using Software Evaluation Metrics," *Information Technology and Management Science*, vol. 19, no. 1, Jan. 2016.

[16] C. Pautasso, A. Ivanchikj, and S. Schreier, "A pattern language for RESTful conversations," in *Proceedings of the 21st European Conference on Pattern Languages of Programs - EuroPlop '16*, Kaufbeuren, Germany, 2016, pp. 1–22.

[17] "Migration of SOAP-based services to RESTful services - IEEE Conference Publication." [Online]. Available: https://ieeexplore-ieee-org.proxy.library.dmu.ac.uk/document/6081828/. [Accessed: 26-Aug-2018].

[18] F. Belqasmi, J. Singh, S. Y. Bani Melhem, and R. H. Glitho, "SOAP-Based vs. RESTful Web Services: A Case Study for Multimedia Conferencing," *IEEE Internet Computing*, vol. 16, no. 4, pp. 54–63, Jul. 2012.

[19] H. Hamad, M. Saad, and R. Abed, "Performance Evaluation of RESTful Web Services for Mobile Devices.," *Int. Arab J. e-Technol.*, vol. 1, no. 3, pp. 72–78, 2010.

[20] O. Kuyoro Shade, I. Frank, O. Awodele, and O. Okolie Samuel, "Security issues in web services," *IJCSNS International Journal of Computer Science and Network Security*, vol. 12, no. 1, 2012.

[21] C. Okello-Obura and F. Ssekitto, "Web 2.0 technologies application in teaching and learning by Makerere University academic staff," 2015.

[22] F. AlShahwan and K. Moessner, "Providing SOAP Web Services and RESTful Web Services from Mobile Hosts," in *2010 Fifth International Conference on Internet and Web Applications and Services*, 2010, pp. 174–179.

[23] "RESTful service composition at a glance: A survey," *Journal of Network and Computer Applications*, vol. 60, pp. 32–53, Jan. 2016.


.

# Web Engineering: Challenges and Open Issues

Selina Kadakwiza
*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
selina.halubanza@gmail.com

Douglas Kunda
*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
dkunda@mu.edu.zm

**Abstract**

Web-based systems haven't been developed following laid down software development approaches in the recent past. Managing quality control as well as software quality assurance has been a challenging task. A concern has therefore arisen regarding not only the security of web-based systems but also integrity issues. A systematic approach of software development is hence advocated for through software engineering. Web engineering promotes "the use of sound scientific, engineering and management principles, and disciplined and systematic approaches to development, deployment and maintenance of Web-based systems." This paper hence looks at key web engineering issues and challenges through a review of various similar literature.

**Keywords:** Web engineering, Web-based system development, Web Page Construction, Web application development.

## Introduction

The internet growth in the recent past is not only transforming the way information is shared online but has also had an influence on software development. The World Wide Web, for instance, had a profound impact on various sectors of the economy such as banking, entertainment and education, among others. Much of the software that was traditionally developed, also known as legacy systems, had to be redeveloped to enable then to be used on the web [1]. This software transformation requires that a software development approach is taken into consideration to develop web-based software that could be used on the internet. The ad hoc development trend for web-based systems has compromised the standard and quality of internet software. Web based software lacks a systematic way of software development which ultimately leads to poor software that attracts a lot of patches for them to cope with internet requests demands. As the demand for sophisticated web-based software increases, there is now a need to have a systematic way of developing internet software with a focus on the development process as well as its security [1].

According to [1], the lack of a systematic approach when developing web systems could ultimately lead to system failure. There could be other challenges such as deployment and software maintenance issues which if not handled properly could not only spill to other web software but instead lead to lack of trust or confidence in web systems [2]. Furthermore, the software crisis in web system development could have more catastrophic effects as compared to the challenges that software developers have been facing for years on end [2].

To instil confidence in web systems as well as avoiding web system crisis, better software development tools should be hence advocated for. This should be at all software development stage, deployment and maintenance stages. [2] further reiterates that "such approaches and techniques must take into account; (1) the unique features of the new medium, (2) the operational environments, and (3) scenarios and multiplicity of user profiles, as well as (4) the type (and skills and knowledge) of the people building Web-based systems."

The method that has been used to analyse the web engineering challenges and issues is based on the review of literature from various sources. Further insights have been added during the discussion phase to highlight the issues at hand.

### What is web engineering?

Web engineering is more than just a collection of software development practices that have been proven over the years but is instead a rather fast-growing field which is a more forward-looking development practice [3]. It draws from various disciplines and is not limited to computer science or software engineering among others. It is identified as the next stage in the computer software development evolution. Web Engineering, according to [3]," is the use of sound scientific, engineering and management principles, disciplined and systematic approaches with the aim of successfully developing, deploying and maintaining of high-quality Web-based systems and applications. **"**

Engineering on the other hand is a discipline that uses scientific principles derived from scientific processes to solve problems [4]. [4] further reiterates that through

increase in data and knowledge, the scientific process can hence be modified accordingly based on new evidence.

### What Knowledge Constitutes Web Engineering

According to [3], the mere aspect of learning web systems development gives us an insight into what knowledge constitutes web engineering. The web engineering knowledge is classified in many ways and it encompasses technologies, standards, methodologies and protocols as well as planning and management. New knowledge once acquired can be channelled to the relevant areas and among such areas is that of learning how to create a simple web page that can later transform into a complex web system, a process called, web construction [3]. According to [3], web construction requires the understanding of the basic technologies that are used when creating web pages such as Html. Specialised tools are also available that imbed the basic knowledge such as Netbeans, Http wrapper and Dreamweaver [5].

### Web Engineering: The Need and Principles

Web Engineering, according to [2], brings with it a well-defined process of web systems development that ensures that web systems developed are not only of good quality but also help to eliminate or minimise software errors while maintaining system quality and maintainability.

### Web Engineering and Software Engineering

Although there is some notable similarity between software engineering and web engineering such as programming in both, a clear distinction is evident[3].

Though Web engineering involves some programming and software development and adopts some of the principles of the software engineering, Web-based system development is different from software development, and Web engineering is different from software engineering. The following outlines some of the difference as stipulated by [2];

1. Most Web-based systems, at least as of now, are document-oriented containing static or dynamic Web pages.

2. Web-based systems will continue to be focussed on look and feel, favouring visual creativity and incorporation of multimedia (in varying degrees) in presentation and interface. More emphasis will be placed on visual creativity and presentation about to the front-end interface with which a user interacts.

3. Most Web-based systems will continue to be content-driven – often Web-based systems development include development of the content presented.

4. Multiplicity of user profiles – Most Web-based systems need to cater to users with diverse skills and capability, complicating human-computer interaction, user interface and information presentation.

5. The nature and characteristics of the medium of Web is not well understood as the software medium.

6. The Web exemplifies a greater bond between art and science than generally encountered in software development.

7. Most Web-based systems need to be developed within a short time, making it difficult to apply the same level of formal planning and testing as used in software development.

8. Also Web is different from software as related to the delivery medium.

Further, the type of individuals who build/develop Web-based systems are vastly varied in their background, skills, knowledge and system understanding, and as well as their perception of Web and quality Web-based system.

### Web Engineering: A Multidisciplinary Field

Web engineering is also considered as a multidisciplinary field and it involves “a mixture between print publishing and software development, between marketing and computing, between internal communications and external relations, and between art and technology” [2]. ” The ability to replicate and reproduce scientific results has become an increasingly important topic for many academic disciplines” [6]. It also incorporates fields such as human computer interaction and social sciences as inputs, among other sources [2].

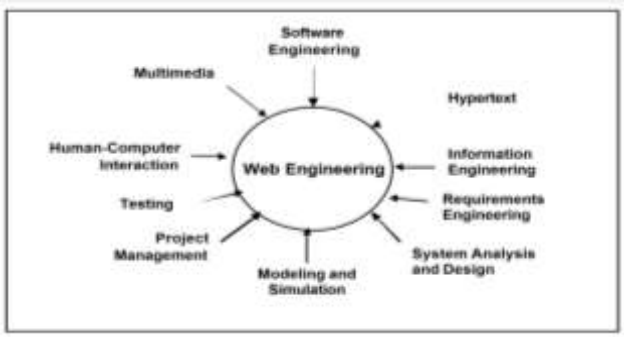



*Fig 1. Web Engineering – A multidisciplinary field [2]*

### Web Application Development and Methodologies

There are two important factors to be considered when developing a website and these are mainly the information or data that changes or is not static but changes with time as well as the website structure [1]. It is difficult to predict the structure of a website too because it also changes during and after the development process because of meeting the current website demands. A distinction is hence visible when we compare this type of software development to that of the tradition way of software development. Web application process should be measurable and trackable [1]. Furthermore, the web software development process’ ability

to break down the process into small or manageable parts helps the web software developers to successfully complete their projects on time.

**Complexity of Web Application**

Web applications' complexity depends on the scale of the web project. Some projects tend to be very complex and hence needs a lot of the developers' time. Such complex web applications mostly involve more than one developer and what is cardinal in this case is the interaction and coordination among web application developers [3]. Having in place good and easy to manage processes not only minimises the risks associated with accomplishing the web software design process but also caters for continual feedback management as well as development of the application within a specified time frame [3]. It being a complex web project needs the help of other disciplines that contributes to web engineering.

According to [3], "We need a sound process for building Web systems that :

- Assists us in capturing the changing requirements and managing the complexity of the development process,
- Assists in the integration of the know-how from various disciplines,
- Facilitates the communication among various members involved in the development process the development team, stakeholders and end-users , and
- Supports the continuous evolution and maintenance and management of the content."

**Web Engineering: Open Issues**

When a web application is being developed, there are some issues that need to be taken into consideration, and they can change in due course, perhaps by the web present trends or the accessibility of new technologies [7].

According to [8], "The need for context adaptability (such as content service meant for a special audience), plus the demand to essentially design this context utilizing rich semantics-founded models are among the key concerns cropping up within the last few years".

The application designer needs to address the audience in a high-level and frequently domain explicit approach such as" premium users". The present frameworks do infrequently provide settings to enhance and widen this context model accessible. Also, usually, "the web application per se does by now exist while its adaptation logic and use ought to be pulled out with such extra concerns [8]".

Therefore, it is authoritative to pose a question concerning the likelihood of extending the existing web structure to offer the needed compatible support [8]. "Another directional exploration within the field of the adaptability and user model-based interface is the Adaptive Hypermedia ". Adaptability is considered orthogonal to three perceptions: content, navigation configuration, as well as presentation as evident in Fig .2 below.

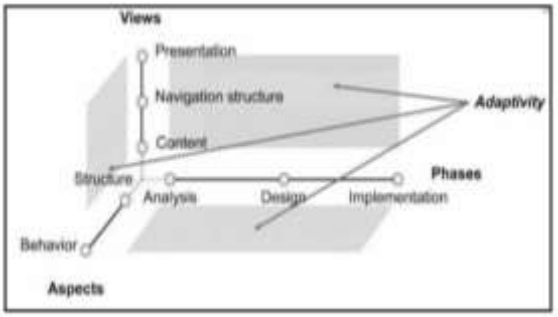



*Fig .2. Modelling Aspect in Web Engineering [8]*

Web engineering adaptability is orthogonal to three views generally: content, navigation structure, and presentation. Web engineering methods have weaknesses regarding adaptability with support features of modern web applications [8].

Web engineering looks at how to build applications and services for the web but with little focus on the fact that increasingly the distinction between internal and external should not be decided when designing and building an application. "Design decisions based on the environment and requirements usually are not static; they are an evolving set of properties shaped by the application, its users, its usage, and the larger context of the Web [9] ".

There have been massive changes in the field of web applications. "The phase has shifted from static to dynamic, and fixed layout has now taken the form of responsive layout, due to distribution of processing capabilities from server side to client side, mainly because of rich set of scripts for user interface and making request to server. This leads to reduction in network traffic". This is on the assumption of trustiness on client, eventually creating a web application more vulnerable [10].

The progression of the Internet has brought a significant change in the development of software, resulting in an increasing presence of Information Systems in Web environments and, subsequently, an increase in security vulnerabilities and threats. "In this context, secure application development has become a crucial component for information systems in the market". Discovering the main vulnerabilities in the coding of software for Web environments and the need for awareness and guidance for developers of information systems[11].

**Challenges**

Web Applications development comes with its unique challenges, some of which are not even found in traditional software environment. These include the need for real time,

changeability, complexity and provision of information that is personalised.[1] further emphasises that it is still a challenge to estimate the actual time required to develop a web application with exactitude. Failure to have a well-defined process of web application development makes it even more expensive and difficult to manage a web application project [2].

### Web based system testing and Cross-browser compatibility

Once developed, web systems need to be checked whether they can do what they were designed for as well as for security and usability. Testing web systems poses a challenge such as working well in one browser and not doing fine in another. According to [2], "the unpredictability of the Internet and Web medium makes testing Web based systems difficulty." There is a notable low performance for websites that have been poorly developed but needs gradual expansions [12].

"The inconsistencies between different browsers, versions and platforms are not only a major time-sink for web developers but also seem to be the reason why most developers avoid enriching the user experience with advanced features that are only possible with JavaScript, CSS2, or Flash[13]".

### Ensuring security

According to [12], web applications are prone to attacks and building a completely secure web system is practically impossible. Furthermore, web developers have also been known for their luck of trust in web applications. Despite the multinational's reliance on web applications in the recent past, the web applications are still vulnerable to cyber-attacks [14].

### Integrating different technologies

Integrating different web technologies in web design has been challenging to developers especially the ones that need to move from static to very advanced web application development.

### Debugging

Every software development has to some extent deal with software bugs. The fact that web system development incorporates various web technologies expose the web system development even to more software bugs than the traditional web development [12].

"Web developers face an extra challenge due to the number of technologies involved and the fact that a web application consists of a part that runs on the server and another on the client [12]".

| **categorize** | **Challenges** |
|---|---|
| Web 1.0 | It had limitations, it was the read-only web, static, and mono-directional, allowing organizations to transmit information to individuals, providing a limited user interaction and only permitting search and to read the information [15]. |
| Web 2.0 | The web 2.0 applications are openly accessible and dynamically generated. Despite the features of web 2.0 being more interesting, it causes higher security risks. For example, User/Hacker may upload content which can run code or carry malware to perform some malicious tasks.<br><br>Every now and then hackers may upload software like free anti-virus to social sites like Facebook (People these days are addicted to Facebook, or other social networking sites and users are blindly clicking each link and every application hence hackers taking advantage).<br><br>Certain websites contain software that should be used for virus removal, but they instead load a Trojan horse. 'Hackers may upload harmful code that could include key loggers that capture victims' keystrokes- including victims' credit card information, password and send them back to hacker" [16]. |
| Web 3.0(the Semantic Web) | The evolution of the web from web 1.0 to web 3.0 has various issues such as scalability, security and performance present in web1.0 and web 2.0. They also propagate to web 3.0 and create a big challenging task for IT experts. Because of the huge collaboration of public and private data it has made web 2.0 and web 3.0 to be more interactive and popular among web users and as well as for hackers also.<br><br>Data privacy in web 3.0 is one of most security issues for the IT professionals. Producers and customers are creating new contents, techniques day by day and publish it to the world for everyone.<br><br>They make deals, share their data and ideas among each other. If someone gives them full control over private data (like Online-games,), considering them as trusted and capable of controlling their data. They modify and publish private data to the world by mistake or intentionally. Illegal and manipulated forms of the same |

| | |
|---|---|
| | type of data can be available on the web, which can create multiplications of errors for everyone [16].<br>Inconsistency: These are logical contradictions which inevitably arise during the development of large ontologism and when ontologism from separate sources is combined [17]. |
| Web 4.0(symbiotic web) | One of the most critical developments of web 4.0 will be the migration of online functionality into the physical world. For example, imagine being able to Google our home to locate car keys or the remote control.<br>There is nothing safe on the internet. Web 4.0 will also be facing challenges on privacy, because personalized search is only possible if the user provides the respective search engine with his personal data. If we consider services like Google Earth that can provide users with information about other people's whereabouts and the tracking devices in some of the newer cell phones being released. This kind of information is easily available nowadays and no one can know where all that information goes, and who gets to see it[16]? |

***Summary Table 1****: Showing categorize and Challenges of Emerging Technologies*

The Web has become a platform for collecting, processing and exchanging large amounts of data through global networking and collective intelligence. Content is continuously updated, fed in and extracted [18].

The Internet has evolved from Web 1.0, with static web pages and limited interactivity, to Web 2.0, with dynamic content that relies on user engagement. This change increased production costs significantly, but the price charged for Internet content has generally remained the same: zero. Because of no transaction records, the "purchase" of this content and its value are not reflected when measuring growth and productivity [19].

Applications such as Twitter, Facebook or YouTube have taken the internet community by storm and have literally initiated a revolution in online communication. These social media applications are also often referred to by the general, somewhat vague but eloquent term – Web 2.0. The term implies a perceived second generation of World Wide Web, i.e. Web 1.0 → Web 2.0 [20].

Web 2.0 is the next big thing in the World Wide Web. It makes use of latest technologies and concepts to make the user experience more interactive, useful and interconnecting. It has brought yet another way to interconnect the world by means of collecting information and allowing it to be shared affectively [21].

Table 2: shows some popular examples of transformation of Web 1.0 based sites to Web 2.0 based sites:

| Web 1.0 | Web 2.0 |
|---|---|
| DoubleClick | Google AdSense |
| Ofoto | Flickr |
| Akamai | BitTorrent |
| mp3.com | Napster |
| Britannica Online | Wikipedia |
| personal websites | blogging |
| evite | upcoming.org and EVDB |
| domain name speculation | search engine optimization |
| page views | cost per click |
| screen scraping | web services |
| publishing | participation |
| content management systems | wikis |
| directories (taxonomy) | tagging ("folksonomy") |
| stickiness | syndication |

*WEB 1.0 and WEB 2.0 Transformation [21]*

Web 3.0 is a web where the concept of website or webpage disappears, where data is not owned but instead shared, where services show different views for the same web. Those services must be focused on content and personalization, and both will be reached by using vertical search. Web 3.0 is the next evolution of the internet. Some hypothesize that web 3.0 will combine the best bits of both web 1.0 and web 2.0 but will be a more user focused, personalized, intelligent, controlled or semantic web experience [22].

Web 3.0 refers to a supposed third generation of Internet based services that collectively comprise what might be called "the intelligent web, for instance, those using semantic web, micro formats, natural language search, data mining, machine learning, cloud computing and artificial technologies which put stress on machine-facilitated understanding of information with a view to providing a more productive and intuitive user experience. It is no wonder that Nova Spivack defines Web 3.0 as the third decade of the Web. Conrad Wolfram stated, "Web 3.0 is where the computer is generating new information, rather humans." [22].

The web 4.0 is also known as the "Symbiotic Web". The idea being the symbiotic web is that once the metadata are organized by web 3.0, human and machines can interact with mind-controlled interfaces. The machines would be clever on reading the contents of the web and react in the form of executing and deciding what to execute first to load the websites fast with superior quality and performance and build more commanding interfaces [23].

Web 4.0 will be read write concurrency web and it will spearhead global transparency governance, distribution, participation, collaboration in key communities such as industry, political, social and other communities. Web operating system (OS) will be such as a middleware which will start functioning like an operating system. Web OS will be parallel to the human brain and implies a massive web of highly intelligent interaction [17].

## Conclusion

The article looked at web engineering and how it can play a critical role in web design. Also looked at are the challenges of designing, deploying and maintaining web systems. The article furthermore, looked at the categories of web developments starting with web 1.0 and ending with web 4.0. More people are now appreciating what it takes to design large and complex web systems. “Web engineering lies at the centre of the Web revolution, which is one of the most important technological revolutions to affect our society[1]”. Web engineering challenges were also highlighted.

The complexity and errors associated with web system design could be reduced if developers appreciate the importance of web engineering [2].

## References

[1] Ginige, A. and Murugesan, S., 2001. The essence of web engineering. IEEE Multimedia, 8(2), pp.22-25.

[2] Murugesan, S., Deshpande, Y., Hansen, S. and Ginige, A., 2001. Web engineering: A new discipline for development of web-based systems. In Web Engineering (pp. 3-13). Springer, Berlin, Heidelberg.

[3] Ginige, A., 2002, July. Web engineering: managing the complexity of web systems development. In Proceedings of the 14th international conference on Software engineering and knowledge engineering (pp. 721-729).ACM.

[4] Mendes, E., 2005, November. A systematic review of Web engineering research. In 2005 International Symposium on Empirical Software Engineering, 2005. (p. 10). IEEE.

[5] Fayzrakhmanov, R.R., Sallinger, E., Spencer, B., Furche, T. and Gottlob, G., 2018, April. Browserless web data extraction: challenges and opportunities. In Proceedings of the 2018 World Wide Web Conference on World Wide Web (pp. 1095-1104). International World Wide Web Conferences Steering Committee.

[6] Cito, J., Ferme, V. and Gall, H.C., 2016, June. Using docker containers to improve reproducibility in software and web engineering research. In International Conference on Web Engineering (pp. 609-612). Springer, Cham.)

[7] Dayang N. A. Jawawi, Karzan Wakil , ,2018 January .A New Adaptive Model for Web Engineering Methods to Develop Modern Web Applications.the 2018 International Conference on Software Engineering and Information Management - ICSIM2018 (pp.32-39)

[8] Karzan Wakil , Dayang N. A. Jawawi ,2018 January .A New Adaptive Model for Web Engineering Methods to Develop Modern Web Applications.the 2018 International Conference on Software Engineering and Information Management - ICSIM2018 (pp.32-39)

[9] Wilde, E. and Gaedke, M., 2008, September. Web Engineering Revisited. In *BCS Int. Acad. Conf.* (pp. 41-50).

[10] Kachhwaha, R. and Purohit, R., 2019. Relating Vulnerability and Security Service Points for Web Application Through Penetration Testing. In *Progress in Advanced Computing and Intelligent Engineering* (pp. 41-51). Springer, Singapore.

[11] Leite, G.S. and Albuquerque, A.B., 2018, September. An Approach for Reduce Vulnerabilities in Web Information Systems. In *Proceedings of the Computational Methods in Systems and Software* (pp. 86-99). Springer, Cham.

[12] Rode, J., Rosson, M.B. and Pérez-Quiñones, M.A., 2005. The challenges of web engineering and requirements for better tool support.

[13] Manhas, J., 2015. Comparative Study of cross browser compatibility as design issue in various websites. BVICA M's International Journal of Information Technology, 7(1), p.815.

[14] Gupta, S. and Gupta, B.B., 2017. Detection, avoidance, and attack pattern mechanisms in modern web application vulnerabilities: present and future challenges. International Journal of Cloud Applications and Computing (IJCAC), 7(3), pp.1-43.

[15] Guarda, T., Leon, M., Augusto, M.F., Haz, L., de la Cruz, M., Orozco, W. and Alvarez, J., 2017, June. Internet of Things challenges. In Proceedings of the 12th Iberian Conference on Information Systems and Technologies, Lisbon, Portugal (pp. 14-17).

[16] Nath, K. and Iswary, R., 2015. What comes after Web 3.0? Web 4.0 and the Future. In *Proceedings of the International Conference and Communication System (I3CS’15), Shillong, India* (pp. 337-341).

[17] Khanzode, K. and Sarode, R., 2016. Evolution of the World Wide Web: From Web 1.0 to Web 6.0

[18] Kollmann, T., 2018. Grundlagen des Web 1.0, Web 2.0, Web 3.0 und Web 4.0. *Handbuch Digitale Wirtschaft*, pp.1-23

[19] Nakamura, L., Samuels, J. and Soloveichik, R., 2018. *“Free” Internet Content: Web 1.0, Web 2.0, and the Sources of Economic Growth* (No. 18-17).

[20] Sykora, M., 2017. Web 1.0 to Web 2.0: an observational study and empirical evidence for the historical (revolution) of the social web. *International Journal of Web Engineering and Technology*, *12*(1), pp.70-94.

[21] Lee, C., 2017. *A survey of the World Wide Web evolution with respect to security issues* (No. e2793v1). PeerJ Preprints.

[22] Khiste, G.P. and Surwade, Y.P., 2018. Publication Productivity of "Web 3.0" By Using Science Direct During

2008-2017. *International Journal for Science and Advance Research in Technology*, *4*(3), pp.1632-1634.

[23] Solanki, M.R. and Dongaonkar, A., 2016. A Journey of Human Comfort: Web 1.0 to Web 4.0. *International Journal of Research and Scientific Innovation (IJRSI)*, *3*(9), pp.75-78.

# Active Learning Environment: Web Metrics, Monitoring and Analysis, Open Issues

**Joshua Cheelo Mubila**
School of Science, Engineering and Technology
Mulungushi University
Kabwe, Zambia
mubilaj@gmail.com

**Douglas Kunda**
School of Science, Engineering and Technology
Mulungushi University
Kabwe, Zambia
dkunda@mu.edu.zm

***Abstract* - Technology as made the world to become a global village in the area of ecommerce, people in different localities are enabling to buy thing online as if they are right in the shopping store and supplier are enable to sell the product as if they are right in their shop with their customer. This is made possible through web browsers which create a platform for both the supplier and the buyer, it enables the buyer to have access to the supplier's website. The suppliers or web operators may want to know the number of people visiting their website and what the visitors are doing on their website, what pages they visit, how long it takes the visitor on their website before leaving.**

**Web operators may also want to know what makes visitor leave their website very fast, what makes them visit one page and leave, and how are the visitors getting to their website. The web operator must not ignore the needs of their visitors, taking care of the visitors needs will increase traffic coming to their website hence increasing sales. Therefore, this paper will look at Web Metrics, Monitoring and Analysis.**

***Keywords: web metrics; web monitoring; web analysis; visitors; page view; traffic source; referring words; bounce rate; active monitoring; server monitoring; application monitoring; email subscribes; conversion rate.***

## I. INTRODUCTION

In this fast-growing world of technology, the use of internet enables people to have access to different kind of information right at their fingertip. People are able to go online and search for the product they may wish to buy and be able to purchase the desired goods. This is made possible through the use of websites. The world is becoming increasingly aware that the Internet is evolving rapidly and constantly growing as more and more users get online. A presence in the web sphere is necessary for all organizations and businesses. The Internet provides numerous multimedia features enabling and changing the way organizations communicate with their customers, suppliers, competitors and employees [1].

All businesses require web presence and the nature of the business directly determine the metrics that will be used for tracking. All short visits to your website are trying to tell you things you are doing wrong, this is determined by slow downloading pages, every form on your website is wrongly filled out, every offer on your site rejected by visitors, very difficult to navigate and unstructured content. You have to listen to the feedback that you receive from the people visiting your website and their view about your website, this will create brand awareness, increase traffic to your website, organization becoming competitive on the market, well designed website, convincing offers, community support, providing quality web service and health website that bring success. Failing to pay attention to this feedback will cut you off your market and you will lose your royal customers for your brand [2].

There are thousands of different variables you could use to track visits to your website [3], this is what makes digital marketing tools so incredible when compared to traditional marketing tools such as newspapers, radio or television. With digital marketing you can track how many times a message posted on your website is viewed, who had access to it, did they do anything to it, what they did, how many times they did, and you can track were they are from. what kind of a device they used to access your website, whether it was through phone, tablet or desktop computer and what web browser they are searching on, you can also track their age, interest, and previous buying behavior, you can monitor and track for days, weeks, month or years in order to understand them better. The more you know about your customer the better you are able to serve them.

These objectives which include traffic and sales can only be achieved if you know how you are performing. To measure your performance in ecommerce you will have to employee web metrics. According to [4], "Metrics help organizations generate more effective Web sites and provide measures that managers understand and that academics can replicate and analyze. To provide practical value, metrics should identify frequency of measurement, frequency of review, source of data, rationale for introducing the measure, who will act on the data, and the purpose of the measure." Therefore, in this paper will discuss web metric, monitoring analysis in digital marketing.

## II. OPEN ISSUES AND CHALLENGES

web metrics were initially developed as a way to understand the success of e-commerce web sites by tracking customers and purchases. Three common Web Metrics, Monitoring and Analysis tools and approaches are: (1), combinations of

user panels and browser logging tools to track sample WWW user populations (2), collecting network traffic data directly from ISP servers and (3) using site-specific server log parsers or page tagging technologies to measure traffic through a particular site. All these methods record and report a wide range of data [5].

Despite the apparent plug-and-play nature of many web metrics tools – a perception encouraged by tool vendors – in practice they can differ widely in cost, functionality, sophistication, and ease of use. As with other types of software and hardware, potential issues with web metrics tools include: limited functionality; limited documentation; limited vendor server capacity and bandwidth, and poor response times when viewing data; inability to download and archive data locally; poor technical support; and the user privacy and trust concerns that can arise when a piece of page code collects data on users and stores and/or transmits those data to third parties. Some sense of these issues can be gained from vendor web sites and representatives, reading web metrics blogs and discussion boards, and consulting colleagues. Note that while some tools offer free demo versions, these may have a limited capacity and functionality that may not repay the learning curve involved in implementing the tool in the first place [5].

Page views count the number of times web pages on a web site are accessed during a visit, including repeat viewings of the same page. Issues have emerged with defining and tracking page views when AJAX (Asynchronous JavaScript and XML) technologies are present, as these permits in-page content refreshing without actual page refreshing, and constantly refreshed AJAX content on the same page will therefore count as only one-page view. [5].

## III. ACTIVE LEARNING ENVIRONMENT

Active learning environment simply means getting involved with the information presented, really thinking about it (analyzing, synthesizing, evaluating) rather than just passively receiving it and memorizing it. Active learning usually results in the generation of something new, such as cause-effect relationship between two ideas, an inference, or an elaboration, and it always leads to deeper understanding [6].

Website owners want to learn how their sites are performing and learn new ideas and were their site need improvements. Website owners learn in an active learning environment by interacting with visitors to the websites. they give feedback to visitor's concerns about their website, and answer frequently asked questions about by the visitors. Website owners use the information obtained from the visitors to improve their websites. tool that is used by website owner to monitor the performance of their website is google analytics. Google Analytics provided information on where visitors came from, what pages they visited, how long they stayed on each page, how deep into the site they navigated, where their visits ended, and where they went from there. By analyzing the data from Google Analytics [7], website owners are able to make improvements to their website.

## IV. WEB METRICS

In order to take advantage of the website analysis, one must have knowledge in metrics. Web metrics are the procedures that show how consumers are making use of a website. organizations use metrics to perfect their website. Web metrics basically depicts unfamiliar areas for most web companies. Web metrics are the measures that reflect how customers are using a website. Companies use these metrics for further improvement of their website. Although most companies today are engaged in website optimization, the actual use of web metrics is a relatively unexplored area [8]. Organization develop web metrics in order to monitor the use of their websites. One of the main areas where web metrics are used within the company is the search function on its website. Since it is expensive to answer phone calls, the company actively promotes the site's search function. Metrics that are used to optimize the search function include the most common queries and the number of queries originating from a particular site. A large number of queries in a certain area may indicate that this part of the website does not satisfy users with the information they are looking for. The number of queries is also used to determine the areas of the site that justify the commitment of resources for improvement [8]. Proper web metrics must be applied in order to effectively measure the effects of the website on its visitors. selecting a good web analytics tool is of extreme importance for unique needs of the users.

*Below are web metrics that will enable operators monitor the performance of their website;*

- *Visitors:* this metrics shows how many people are visiting your website, these include new visitors and returning visitors. The web metric helps to trace traffic pattern coming to your website over a period of time [9].
- *Page View:* this is the total number of times; a particular page of your website is visited. This metrics is very important because, its show which pages of the website receive high number of visits. Using this information, you will be able to see which pages are doing fine and which ones needs improvements [10].
- *Traffic source:* tell you where traffic is coming from to get to your website, so that you can see what percentage is direct traffic (traffic that come direct to your website), or it is referral traffic (traffic coming from other source such as google, yahoo, Bing or other source that send traffic to your website) [11].
- *Referring Keywords:* what keywords are driving traffic to your website, this is great way to help see

how your website effects are going and see which keywords, you need to optimize for throughout content on various web pages of your site [11].

- *Bounce Rate:* Your bounce rate is the number of people who visit your site and then quickly leave, or bounce off. Here is where we want to see a low number, as this means that people are sticking around and having more of a look through your content. This gives you the percentage of people coming to your website seeing one page and then leave, bounce rate very cross pages of your website, but it can communicate if the information presented on your website, especially home page or running page is favorable and significant, if it is not go back and make changes to your website and see if your bounce rate will go down [11].
- *Number of Email Subscribers:* it does not matter what business you are doing. You can take advantage of using email marketing strategy. Encourage your visitors of your website to leave their email address with you by offering something of value in exchange such as eBooks, pdf files, videos or consulting sessions. Just like with website visitors, you want to see a steady increase of email subscribers over time.
- *Click Through Rate (CTR):* this refer to how many people took action and clicked one of your post or advertisements. CTR is a valuable measure metric to watch any time you are creating an ad for any of your social media or search engine platforms. High click through rate means that what your ad is saying is important, easy to understand and interesting enough for someone to click it and learn more, it is rewarding. According to Dr. Birgit Weischedel, click through rates (percentage of visitors who clicked on a link or banner) and conversion rates (percentage of visitors who completed a desired action). The overall usage of a site is typically determined with traffic measures, such as the number of unique visitors, user sessions, page requests or visits [8].
- *server logs:* A web server is a computer that delivers pages to the computer that requests information by entering a web address or URL (Uniform Resource Locator). That request and the interaction between server, site and visitor create data that are stored in a log file on the server. Web server logs are not limited to usage data but can also show other information about the website, such as access logs (IP address, date and time of access), agent logs (browser, browser version, operating system), error logs (record error events), and referrer logs (list pages that link to documents on the server). Combining these logs creates a picture of users' behavior on the website, the technology they are using, problems on the website as well as the structure and navigation of the website [8].
- *Conversion Rate:* conversion rate is any call of action a customer completes that fulfill the purpose of your website. A conversion could be signing up for a newsletter or buying your product. The conversion rate is traced as a percentage of your traffic. The higher it is the better [11].

Metrics indicate when a system has fulfilled its purpose or is acting in accordance with its normative behavior. Thus, web metrics are intended to measure the extent to which a web site contributes to the attainment of business objectives. In electronic commerce situations, site metrics now tell a company not only how often users cannot reach its site fast enough to make purchases, but also how well it is performing compared with its competitors [12].

To understand the customer behavior from the data, web analytics tool presents a summary report, a group of basic metrics that are available immediately after logging into the tool. Google Analytics shows the following chart:

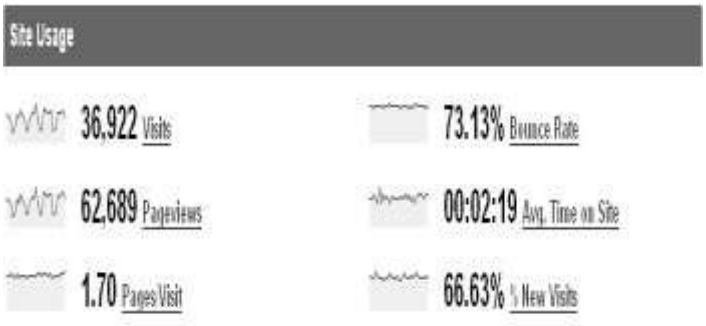



FIG. 1 GOOGLE ANALYTICS BASIC METRICS

The about chart was adopted from [13] shows the following metrics:

- *Visits:* the number of sessions on your website and number of times someone interacted with your 'site.
- *Bounce Rate:* the percentage of single page view visits (this metric can also have different definitions, such as a visit that last less than 5 seconds).
- *Page Views number:* the number of pages that were requested in all visits.
- *Pages/Visit:* how many pages were seen, on average, in each visit.
- *Average Time on Site:* how long people stayed on the 'site.
- *New Visits:* how many sessions were from people who visited your site for the first time.

The preceding numbers will vary from industry to industry, and for this reason there is no absolute benchmark to which a website owner can compare. The best way to proceed is to

trend it over time, as much data as possible, to understand if the website is improving or not [13].

Table 1 was adopted from [12], is a prioritized list of web metrics used by organizations. It means that many organizations use web metrics to check site traffic or to check popular content. The data show that few organizations use web metrics to improve site navigation or monitor visitors' usage behavior. It will be necessary to recognize that innovative use of web metrics can provide organizations with substantial benefits [12]. In addition, the metric would help learn about how visitors reach to and use a web site.

TABLE I
LIST OF WEB SITE METRICS

| | **Web Metric** | **Count** | **Percent of Responses** | **Percent of Cases** |
|---|---|---|---|---|
| 1 | Visits | 27 | 22.5 | 65.9 |
| 2 | Pageviews | 24 | 20 | 58.5 |
| 3 | Best pages | 19 | 15.8 | 46.3 |
| 4 | Page durations | 7 | 6.8 | 17 |
| 5 | Navigation paths | 7 | 6.8 | 17 |
| 6 | Entry/exit IPs | 5 | 4.2 | 12.2 |
| 7 | Visitors' traffic | 3 | 2.5 | 7.3 |
| 8 | Links to another web sites | 3 | 2.5 | 7.3 |
| 9 | Unique visitors | 3 | 2.5 | 7.3 |
| 10 | Conversion rates | 3 | 2.5 | 7.3 |
| 11 | Customer loyalty | 2 | 1.7 | 4.9 |
| 12 | Per-visitor revenue | 2 | 1.7 | 4.9 |
| 13 | CTR (Click-Thru-Rates) | 2 | 1.7 | 4.9 |
| 14 | Hits | 2 | 1.7 | 4.9 |
| 15 | Top search strings | 2 | 1.7 | 4.9 |
| 16 | Number of log-ins | 2 | 1.7 | 4.9 |
| 17 | User environment (browser, OS, etc.) | 1 | 0.8 | 2.4 |
| 18 | File uploads/downloads | 1 | 0.8 | 2.4 |
| 19 | Referrals | 1 | 0.8 | 2.4 |
| 20 | Number of accesses (e.g. accesses by hour) | 1 | 0.8 | 2.4 |
| 21 | Number of simultaneous accesses | 1 | 0.8 | 2.4 |
| 22 | Visitors by weekdays | 1 | 0.8 | 2.4 |
| 23 | Conversion rates | 1 | 0.8 | 2.4 |
| 24 | Total responses | 120 | 100 | 292.7 |

## V. WEB MONITORING

Website monitoring is used to monitor the performance and use of an organization's website to help understand problems and ultimately improve the effectiveness of the site by improving "metrics" such as conversion rates and user experience, or preventing or detecting undesirable behavior such as fraud, this is according to [14].

By "monitoring" we mean the automated process of testing, tracking and reporting on the availability and condition of the systems, services, and networks that make up a Web presence. The main purpose of web monitoring is to ensure that site visitors are able to access their online applications, websites, and perform actions such as searching, online shopping, checking an account balance, or simply researching and reading. The aim is to avoid and minimize downtime and keep the server and online applications running. [15].

Web monitoring is a method of examining and asserting that user able to communicate with a website or web application as required. Website monitoring is often used by organizations to ensure that website functionality, uptime and performance is as expected. Website monitoring is used by organization to consistently monitor how the website is performing, such as load times, server response times, page element performance that is often examined and used to further improve website performance. Monitoring is important to ensure that a website is available to users, downtime is minimized, and performance can be improved. Website monitoring also provide a standard by which a website can be measured against the performance of a competitors to help determine how well a site is performing.

A traffic monitor provides statistics of traffic using an activity input for receiving data related to activity on a server system. Events being monitored are binned by topic or term, where the terms are associated with categories. The categories can be a hierarchy of categories and Subcategories, with terms being in one or more categories. The categorized events include page views and search requests and the results might be normalized over a field of events and result output for outputting results of the normalizer as the statistical analyses of traffic [16].

Web monitoring will monitor your website on a regular basis, for your pages, your action, your visitors, and if there is a problem it will alert you. Below are different types of web monitoring.

- *Active Monitoring:* Active monitoring allows the web operator to diagnose problems, such as web application problems and network or server connectivity problems, and to ascertain if a website is going through downtime or it not performing as expected, before these problems' effects a large number of site visitors. Active monitoring can identify performance and connectivity problem from numerous Internet points. Active monitoring solutions reports if a website or application is slow or undergoing downtime within minutes of the occurrence. Active monitoring allows determine the cause of the problem, such as slow applications, Internet network problems, and groundwork issues. Active monitoring solutions does not depend on real Web traffic, applications can be tested continually, providing up-to-date status on website availability and performance. The advantage of web monitoring is that performance issues can be identified before they can happen and affect your website or web application. Web

monitoring allows simulation of user transaction scripted business processes, such as the use of shopping carts, secure log-ins, forms,

- *Passive Monitoring:* Passive monitoring is called real-user monitoring. Passive monitoring monitors how real visitors are actually visiting your website, which webpages are they visiting and how fast are the webpages loading. It also monitors what your visitors are doing, are they filling forms or they successfully been able to actually view your website and go through the website correctly. Passive monitoring is largely implemented in enterprises where the IT department use management tools to monitor the performance of the application on a network. Passive monitoring is implemented inside a local area network will monitor traffic flow as it enters and leave the website. it also monitors actual user interactions as they happen.
- *System Monitors:* A system monitors browser operation on a web page to identify objects referenced by the web page. For objects referenced by the web page, the system tracks a performance metric associated with each object. The system performs post-processing on the performance metrics of objects referenced by the web page to create a performance record for that web page, and transfers the performance record to a remote server for analysis of performance of web page operation in the browse. [17]
- *Server Monitoring:* This is monitoring the actual server, the actual device that is running your website. By looking at the actual specification of your website such as central processing unit, random access memory and disk usage.
- *Web Server Monitoring:* Web server monitoring specification allow performance management by monitoring the response time of the application and updates the status based on a given starting point and also provides complete management reports. Web monitoring can prevent problems by revealing patterns of resource usage and performance that might otherwise go undetected. For example, full disks typically cause a host of problems for applications and operating systems. Simply knowing that a disk is nearing capacity may save Web administrator hours of time fixing the problems caused by a full disk [15]. The Web Server performance metrics of interest are its CPU utilization and the HTTP bytes/sec data. Both can be found in the Web Server performance metric tables usually just after the WIPS, WIPSb and WIPSo throughput graphs in the FDR. Given the HTTP bytes/sec data, one can compute the Mbytes/sec or Mbits/sec network traffic per server and per CPU. You can determine what type and how many NICs have been used in the Web Server by examining the price data on the second page of the Executive Summary. Given the Web Server CPU utilization and network throughput rate, one can determine if the Web Server has any extra headroom or is it configured near its maximum capabilities. In addition, by multiplying the value by the number of Web Servers and computing similar data on the Image Servers and Web Caches, you can obtain an idea of the overall network traffic supported by the front-end servers and the switch [18].
- *Application Monitoring:* Application management also known as application performance monitoring guarantees that standard software application or web applications perform as expected. This technique estimates the performance of the applications and checks if everything is running as expected. Web monitoring applications, it is usually desirable to capture as much information as possible, with as little delay as possible. Dynamic Web pages undergo updates over time, and each updated version of the page potentially contains new information of value to the application [19]. If your company website is connected to external services either third part services or other system through application or external extension, you may want to monitor the other parts of your websites not just the web server its self, if is functioning correctly, you may actually need to query these specific external parts in order to check their functionalities if they are operating correctly.
- *Transaction Monitoring:* Every web application has particular functions that are target oriented. Including things like a customer's ability to purchase something on your website. Transaction monitoring, is the control of significant business applications and services by checking the individual transactions that flow across the application structure. Transaction monitoring tools measure response time of each component they are connected with each component. This data allows the operation team to see where performance is slowed down.
- *Log Files:* The first method of metric gathering uses log files. Every Web server keeps a log of page requests that can include (but is not limited to) visitor IP address, date and time of the request, request page, referrer, and information on the visitor's Web browser and operating system. The same basic collected information can be displayed in a variety of ways [11]. The log file searches aim at different actions requested by the user. This method can always be used in conjunction with a page tagging system, like Google Analytics in order to gain even higher precision. This approach allows Operational data from the e-shop to be read into the database of the analyzer. The reports

generated display real product Id's and category names, making the reports easier to read and is more focused on the e-shop. Transferring the log files to a relational database allows the flexibility of SQL to come into play with data manipulations. New reports and measurements can be generated easily. The generated data is completely manageable. It is loaded in batches that can be named and tagged [20].

- *Page Tagging:* The second method for recording visitor activity is page tagging. Page tagging uses an invisible image to detect when a page has been successfully loaded and then uses JavaScript to send information about the page and the visitor back to a remote server. Page Tagging requires an extra web server, to whom the visitor's browser is automatically sent. This server collects the log data generated by this visit and stores it to a specific data base for each site, based on an account number [20].

## VI. WEBSITE ANALYSIS

Website analysis is basically an inspection checks before continuing to any assignment or task. This analysis determination does not only represent the errors and issues linked to websites but also it will give you indication of how the works must be done on the website. website analysis consists page rank, page issues, website domain status, performance on search engine. Website analysis is a significant characteristic for any business, it indicates the complete statistics and data associated to your online business. Website analysis offer deep understanding of your website performance and user associated with data.

The process of monitoring the nature and behavior of traffic, rather than its content, is known as traffic analysis. Traffic analysis usually works equally well on encrypted traffic and on unencrypted traffic. This is because common encryption methods, such as SSL, do not try to obfuscate the amount of data being transmitted. Because of this, traffic analysis can usually tell you not only who the receiver and sender of the data is, but also how much data was transferred. In certain situations, an attacker having knowledge of the amount of data transferred can have disastrous results [21].

### A. *Importance of Website Analysis*

Website analysis monitoring assists to appreciate the user viewpoint, traffic source and search engine performance. It will also give you information about the improvements that have to be worked on your running website.

- *Keywords optimization:* keyword search is the actual words or phrase that visitors type into search engines in order to find your website. you have to study and learn the needs of your visitors to your website that is both current and new visitors, what kind of words or phrases are they using to search topics on your website. Website analysis play a significant role in keyword optimization, it explains the complete interpretation of your website keywords which are performing well and which are worse. If your website contains wrong keywords, will not attract enough traffic from your targeted audience. Use analytics data to come up with objective keywords. Keyword searches report rate of all the keywords that were used successfully to get to your website from all search engines.
- *Links Authority:* Website analysis helps to know your link profile of your website. A good number of links has the ability to increase your website trustworthiness from your visitors. Website analysis enables you to see the complete link structure of your website like inbound and outbound links. This can also help you in creating great amount of leads.
    - *Content:* Website analysis provides operators with understanding related to their website content; it will enable them to discover important and irrelevant content on their website. content on the website should be easily found and easy to understand by your visitors. More and more companies analyze website usage data (Content) in order to understand customers' needs to increase traffic and ultimately increase their revenue. Different sites can have different goals like selling more products and attracting more users to generate more income through advertisements. Websites want to keep visitors longer (reducing bounce rate) to encourage users to return and to make every visit end with completion of targeted action (conversion) [22].
    - *Traffic:* It also helps website operator to investigate where traffic is coming from to their website, including user activities and acquisition so that they can easily pay attention on the features which are performing great for their website. So, if you want to cultivate your online business you must analyze your website, as it helps you in every way in directing customers and improving search engine rank. To bring value, web analytics must differentiate between a wide variety of traffic sources, marketing channels, and visitor types. A common question is: "where did visitors learn that information?" For example, parameters used in tracking direct traffic from email, social media, or mobile devices allow correlation of traffic sources with marketing campaign cost, which helps to evaluate return on investments [22].

When analyzing the visit length, the measurements are often broken down into chunks of time. The goal of measuring the data in this way is to keep the percentage of visitors who stay on the Website for less than five seconds as low as possible. If visitors stay on a Website for such a short amount of time it usually means they either arrived at the site by accident or the site did not have relevant

information. By combining this information with information from referrers and keyword analysis, one can tell which sites are referring well-targeted traffic and which sites are referring poor quality traffic [11].

TABLE II
WEBSITE ANALYSIS

| | Metric | Description | Category |
|---|---|---|---|
| 1 | Visitor Type | Who is accessing the Website (returning, unique, etc.) | Site Usage |
| 2 | Visit Length | The total amount of time a visitor spends on the Website | Site Usage |
| 3 | Demographics and System Statistics | The physical location and information of the system used to access the Website | Site Usage |
| 4 | Internal Search Information | Information on keywords and results pages viewed using a search engine embedded in the Website | Site Usage |
| 5 | Visitor Path | The route a visitor uses to navigate through the Website | Site Content Analysis |
| 6 | Top Pages | The pages that receive the most traffic | Site Content Analysis |
| 7 | Referring URL and Keyword Analysis | Which sites have directed traffic to the Website and which keywords visitors are using to find the Website | Referrers |
| 8 | Errors | Any errors that occurred while attempting to retrieve the page | Quality Assurance |

The Table II was adopted from [11].

- *Visitor Type:* There are two types of visitors: those who have been to the site before, and those who have not. This difference is defined in terms of repeat and new visitors. In order to track visitors in such a way, a system must be able to determine individual users who access a Website; each individual visitor is called a unique visitor, [11].
- *Visit Length:* Also referred to as Visit Duration or Average Time on Site (ATOS), visit length is the total amount of time a visitor spends on a site during one session, [11].
- *Demographics and System Statistics:* The demographic metric refers to the physical location of the system used to make a page request. This information can be useful for a Website that provides region-specific services, [11].
- *Internal Search:* If a Website includes a site-specific search utility, then it is also possible to measure internal search information. This can include not only keywords but also information about which results pages' visitors found useful, [11].
- *Visitor Path:* A visitor path is the route a visitor uses to navigate through a Website. Excluding visitors who leave the site as soon as they enter, each visitor creates a path of page views and actions while perusing the site. By studying these paths, one can identify any difficulties a user has viewing a specific area of the site or completing a certain action (such as making a transaction or completing a form), [11].
- *Top Pages:* Analysis mentions three types of top pages: top entry pages, top exit pages, and most popular pages. Top entry pages are important because the first page a visitor views makes the greatest impression about a Website. By knowing the top entry page, one can make sure that page has relevant information and provides adequate navigation to important parts of the site. Similarly, identifying popular exit pages makes it easier to pinpoint areas of confusion or missing content, [11].
- *Referrers and Keyword Analysis:* A referral page is the page a user visits immediately before entering to a Website, or rather, a site that has directed traffic to the Website. A search engine result page link, a blog entry mentioning the Website, and a personal bookmark are examples of referrers. This metric is important because it can be used to determine advertising effectiveness and search engine popularity, [11].
- *Errors:* Errors are the final metric. Tracking errors has the obvious benefit of being able to identify and fix any errors in the Website, but it is also useful to observe how visitors react to these errors. The fewer visitors who are confused by errors on a Website, the less likely visitors are to exit the site because of an error, [11].

## CONCLUSION

The study contributes significantly to the web visitors needs that have to be fulfilled in order for online business to remain competitive on the market and increase their sales. It is also discovered that not paying attention to the needs of your website visitors, will keep you out of business as your competitors would have attracted your current and new customers. To be competitive on the market, online business should employee web metrics, monitoring and analysis as suggested in this study for web quality.

Forthcoming studies in this area must consist of a valuation of Web Metrics, Monitoring and Analysis tools based on digital marketing data for Web Metrics, Monitoring and Analysis tools highlighted to track performance on social networks and tools for collecting feedback from visitors, and for conducting different testing.

## ACKNOWLEDGEMENT

The authors would like to thank professor kunda, who has been a real significance in the compilation of this paper. His practical advice has been an inestimable source of knowledge and support for me during this process. I would also like to thank all student in my class for sharing knowledge and advice rendered to me.

## REFERENCE


[1] I. Bekavac and D. G. Praničević, "Web analytics tools and web metrics tools: An overview and comparative analysis," *Croat. Oper. Res. Rev.*, vol. 6, no. 2, pp. 373–386, 2015.

[2] D. Narayandas, "Building loyalty in business markets," *Harv. Bus. Rev.*, vol. 83, no. 9, pp. 131–139, 2005.

[3] J. Burby, A. Brown, and W. S. Committee, "Web analytics definitions," *Wash. DC Web Anal. Assoc.*, 2007.

[4] J. W. Palmer, "Web site usability, design, and performance metrics," *Inf. Syst. Res.*, vol. 13, no. 2, pp. 151–167, 2002.

[5] M. Khoo, J. Pagano, A. L. Washington, M. Recker, B. Palmer, and R. A. Donahue, "Using web metrics to analyze digital libraries," in *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries*, 2008, pp. 375–384.

[6] A. King, "From sage on the stage to guide on the side," *Coll. Teach.*, vol. 41, no. 1, pp. 30–35, 1993.

[7] W. Fang, "Using Google Analytics for improving library website content and design: A case study," 2007.

[8] B. Weischedel and E. K. Huizingh, "Website optimization with web metrics: a case study," in *Proceedings of the 8th international conference on Electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet*, 2006, pp. 463–470.

[9] D. Waisberg and A. Kaushik, "Web Analytics 2.0: empowering customer centricity," *Orig. Search Engine Mark. J.*, vol. 2, no. 1, pp. 5–11, 2009.

[10] S. Bhat, M. Bevans, and S. Sengupta, "Measuring users' Web activity to evaluate and enhance advertising effectiveness," *J. Advert.*, vol. 31, no. 3, pp. 97–106, 2002.

[11] D. Booth and B. J. Jansen, "A review of methodologies for analyzing websites," in *Web technologies: Concepts, methodologies, tools, and applications*, IGI Global, 2010, pp. 145–166.

[12] I. B. Hong, "A survey of web site success metrics used by Internet-dependent organizations in Korea," *Internet Res.*, vol. 17, no. 3, pp. 272–290, 2007.

[13] D. Waisberg and A. Kaushik, "Web Analytics 2.0: empowering customer centricity," *Orig. Search Engine Mark. J.*, vol. 2, no. 1, pp. 5–11, 2009.

[14] L. M. Goldspink and M. J. Duckett, "Website monitoring and cookie setting," Nov-2014.

[15] S. Popa, "WEB Server monitoring," *Ann. Univ. Craiova-Econ. Sci. Ser.*, vol. 2, no. 36, pp. 710–715, 2008.

[16] J. Yoo, K.-T. Lim, S. B. Wong, and E. Yasnokvsky, "Web site activity monitoring system with tracking by categories and terms," US7146416B1, 05-Dec-2006.

[17] P. Anastas, W. R. Breen, Y. Cheng, A. Lieberman, and I. Mouline, "Methods and apparatus for real user monitoring," Jul-2010.

[18] W. D. Smith, *TPC-W: Benchmarking an ecommerce solution*. 2000.

[19] S. Pandey, K. Dhamdhere, and C. Olston, "WIC: A general-purpose algorithm for monitoring web information sources," in *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*, 2004, pp. 360–371.

[20] C. J. Aivalis and A. C. Boucouvalas, "Log file analysis of e-commerce systems in rich internet web 2.0 applications," in *2011 Panhellenic Conference on Informatics*, 2011, pp. 222–226.

[21] A. Hintz, "Fingerprinting websites using traffic analysis," in *International Workshop on Privacy Enhancing Technologies*, 2002, pp. 171–178.

[22] G. Zheng and S. Peltsverger, "Web Analytics Overview," in *Encyclopedia of Information Science and Technology, Third Edition*, IGI Global, 2015, pp. 7674–7683.

# E-government Implementation Models and Challenges: The case of Zambia

Akabana Kalaluka
*Department of Computer Science*
*School of Science Engineering and Technology*
*Mulungushi University*
*P.O Box 80415, Kabwe, Zambia*
akabana.kalaluka@kafubu.co.zm

Douglas Kunda
*Department of Computer Science*
*School of Science Engineering and Technology*
*Mulungushi University*
*P.O Box 80415, Kabwe, Zambia*
dkunda@mu.edu.zm

**Abstract - Since the development of computers, the world has seen many innovations as a result of their extensive use. Information and Communications Technologies (ICT) have made a significant impacted how governments, business and society at large process and interact with each other. For over two decades now, governments the world over have realized that ICT's can facilitate new approaches to service delivery, stakeholder engagement and information access. With the push for e-government in full force, challenges emerged which are common in nature for countries in the developing world.**

**Key words: E-government, Maturity model, Implementation, ICT**

## I. INTRODUCTION

Since the development of computers, the w
orld has seen many innovations as a result of their extensive use. Information and Communications Technologies (ICT) have made a significant impacted how governments, business and society at large process and interact with each other. For over two decades now, governments the world over have realized that ICT's can facilitate new approaches to service delivery, stakeholder engagement and information access. This realization lead to the push for e-government which is simply the use of information technology to deliver/facilitate government services to citizens, business and other stakeholders [1]. E-government is not only about technology conversion of processes from traditional forms of transacting, but calls for a paradigm shift in the way governments function and can serve their citizens better [2].

While e-government benefits are easy to comprehend, there are many hurdles that governments face during implementation. This has made many developing countries lag behind their developed counterparts. Africa in particular, is are far off behind compared to developed countries [3]. The United Nations E-Government Development Index (EGDI) which represents the state of e-government development of the United Nations member states. Along with an assessment of the website development patterns in a country, the EGDI incorporates the access characteristics, such as the infrastructure and educational levels, to reflect how a country is using information technologies to promote access and inclusion of its people. The EGDI is a composite measure of three important dimensions of e-government, namely: provision of online services, telecommunication connectivity and human capacity. The world leader in 2018 is Denmark with a EGDI of 0.9150 and the sub-regional leader for Africa being Mauritius at 0.6678. The world average EGDI is 0.5491 and in Africa at 0.3375 [4].

E-government has also given birth to a number of models by different researchers which mostly address the four major areas of e-government [5].

**Government-to-Citizens (G2C)** – This focuses on services that government offers to its citizens through ICT's. The includes having a two-way communication system were citizens can communicate with their governments. Services offered can be e-Health, e-Education, e-Procurement and e-Banking.

**Government-to-Business (G2B)** – This is where governments and businesses interact with each other in dissemination of policies, payment of taxes and exchange of information.

**Government-to-Government (G2G)** – This is also called e-Administration. It uses ICT's to decentralize government functions like local and central government.

This paper presents e-governments models and challenges that face developing countries.

## II. E-Government Definitions

E-government has been defined in a number of different ways with ICT's being a common denominator in all the definitions.

[1] defines it as the "use of technology to enhance the access to and delivery of government services to benefit citizens, business partners and employees".

[6] definition pointed to all the "information and communication technology platforms and applications in use in the public sector of the use of the internet for delivering government information and services to citizens".

The Zambia's National Information and Communication Technology Policy 2006 states e-government to be "*delivery by Government of products, services, policies and the engagement of stakeholders in civic and government matters through the use of Information and Communication Technologies in order to achieve Government to Consumers, Government to Business and Government to Government interaction and transactions"*[7].

[8] stated that e-government is the "*delivery of improved service to citizens, businesses, and other members of society through the internet or other digital means"*.

[9] defined it as *"use of information and communication technologies to offer citizens and businesses the opportunity to interact and conduct business with government by using different electronic media such as telephone touch pad, fax, smart cards, self-service kiosks, e-mail / Internet, and EDI*".

On the world map, developing countries in Africa have lagged behind when compared with the world average e-government development index [4]. This is largely due to numerous factors which are both common to all countries implementing e-government and some unique to Africa and to be specific sub-Sahara.

## III. MATURITY MODELS

Many maturity models have been developed by different scholars to help in identifying progress made in achieving e-government objectives. These models have not brought consensus from all scholars as areas not tackled by others are brought to light by others with their own proposed solutions.

### *Layne and Lee's four-stage model*

Layne and Lee's model [10] based on technical, organizational and managerial feasibilities, suggested e-government to be an evolutionary phenomenal with initiatives to be accordingly derived and implemented. The four stages proposed by Layne and Lee's are explained below and Fig. 1 shows the various stages.

*Stage 1: Cataloguing:* An online presence is established creating a website with non-transactional information on the site. This stage offers the least amount of functionality for users. As this stage progresses, sites of other department are linked with services offered. Services at this stage are generally passive.

*Stage 2: Transaction:* At this stage, electronic services start to be developed using the internet. Citizens are able to fill in forms for their government requirements instead of filling in paper work. A two-way communication channel is opened allowing government and citizens interact.

*Stage 3: Vertical Integration:* This stage focusses on transformation of government services, rather than automating and digitizing existing processes. Government processes and systems at different levels (vertical) start to be integrated.

*Stage 4: Horizontal Integration:* This is when system integration is achieved across different functions enabling transactions to be run across
government institutions.

### *Andersen and Henriksen Public Sector Process Rebuilding (PPR) model*

Anderson and Henriksen [11],proposed a maturity model labelled the Public Sector Process Rebuilding (PPR) model which goes a step further than the Layne and Lee model to dictate that IT should be used strategically to encompass more areas than simply integration issues and supporting functions of governments. [11] proposal is to move away from seeing e-government as an operational and technical interfacing activity, but as a strategic use of IT to drive

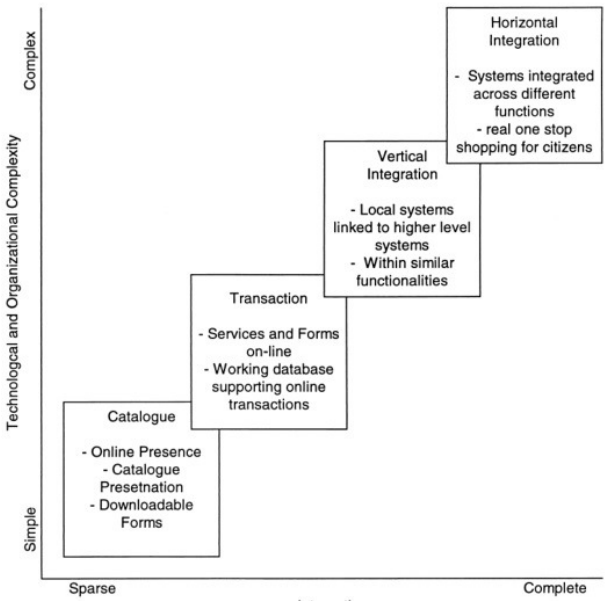



Fig. 1 Dimensions and stages of e-government development [7]

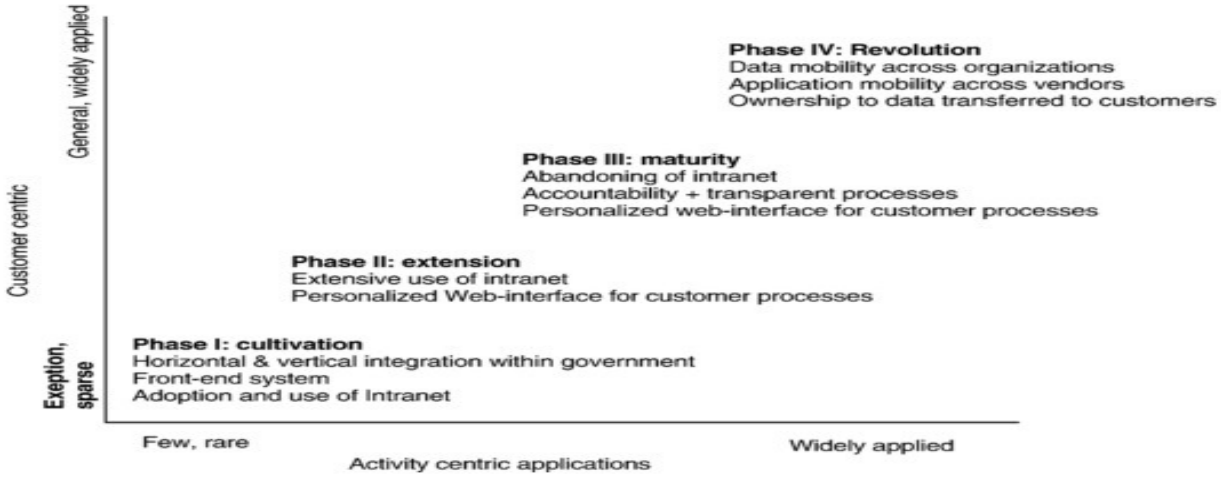



Fig. 2 The PPR maturity model: activity and customer centric stages [8]

governments functions. Figure 2 shows the stages.

*Phase 1: Cultivation:* This phase involves horizontal and vertical integration within government with limited use of front-end systems for customer services, and adoption and use of the intranet within the government.

*Phase 2: Extension:* In phase 2, there is extensive use of intranets, and adoption of user interface for citizen's processes.

*Stage 3: Maturity:* At this stage, the organization matures and abandons the use of the intranet, develops transparent processes, and offers personalized web interface for processing of customer requests.

*Stage 4: Revolutionary:* In the last stage, there is data mobility across organizations, application mobility across vendors, and ownership to data transferred to customers.

*Long, Y., & Siau, K. Synthesized e-government model*

Long, Y., & Siau, K. [12] synthesised a 5 stage model that is stated to be comprehensive and simple while capturing the main ideas of previous models. The five stages namely: web presence, interaction, transaction, transformation, and e-democracy. Figure 3 below shows how they interact. The model is such that, the first three stages deal with automation and digitization of government process which then follows the last two stages which aim to transforming government services, reorganizing the internal operational process, and reconceptualise the way citizens would participate in government decision-making.

*Stage 1: Web presence:* At this stage, simple web presence is established with basic information on websites such as the institutions vision and mission, office hours, contact information, and official documents.

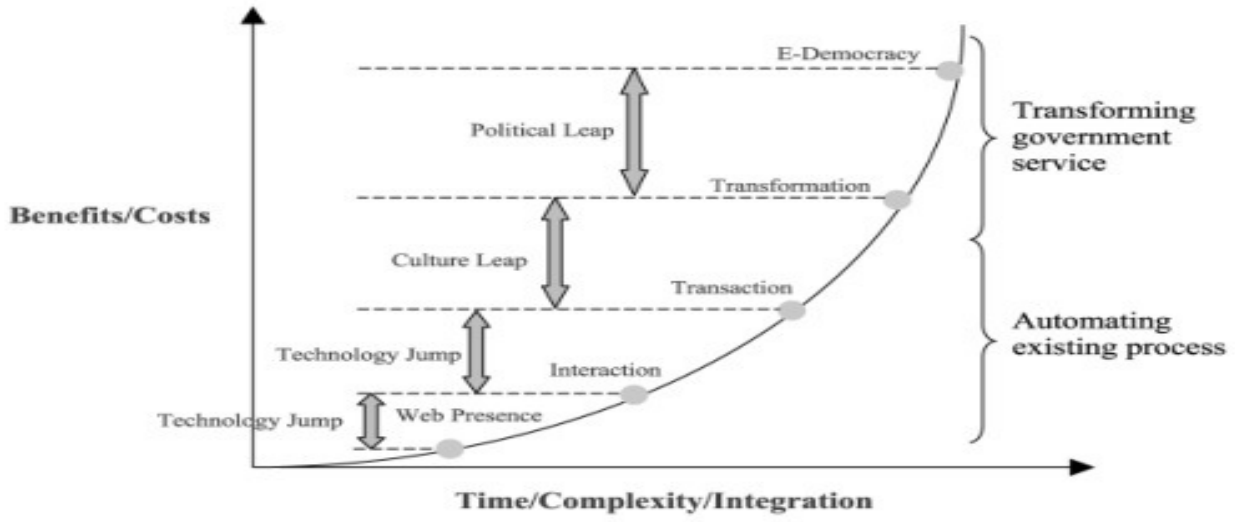



Fig. 3 Synthesizing e-government stage [9]

*Stage 2: interaction:* This phase offers simple two-way interaction between users, and the government with exposure to basic search engines, e-mail systems, as well as official form downloads.

*Stage 3: transaction:* During this stage, both individual citizens and businesses are able to fully transact online with services like license applications, tax filing, personal information updates, fulfilling tax forms, applying licenses and reporting financial data.

*Stage 4: transformation:* This stage involves the way government offers services. Transformation involves both horizontal and vertical integration of services and processes.

*Stage 5: e-democracy:* This is a long term goal offering facilities such as online voting, polling and surveys, governments attempt to improve political participation, citizen involvement, and politics transparencies.

*Gartner's four-stage model*

The Gartner Group [13] proposed a four-stage model for e-government which measures initiatives and establishes a road map to achieve desired levels of constituency service. The four-stages are list below including figure 4.

*Stage 1: Information:* The first stage enables web presence for government where simple information is provided. Communication is one way during this stage.

*Stage 2: Interaction:* The interaction phase offers some level of interaction between government and citizens which can be referred to as Government-to-Citizens (G2C), Government-to-Business (G2B). Email and interactive forms that enable interaction are
accessible allowing for two-way communication to take place.

*Stage 3: Transaction:* The transaction stage enables online services to be offered such as paying for license renewal and taxes by citizens and businesses.

*Stage 4: Transformation:* The final stage looks at the

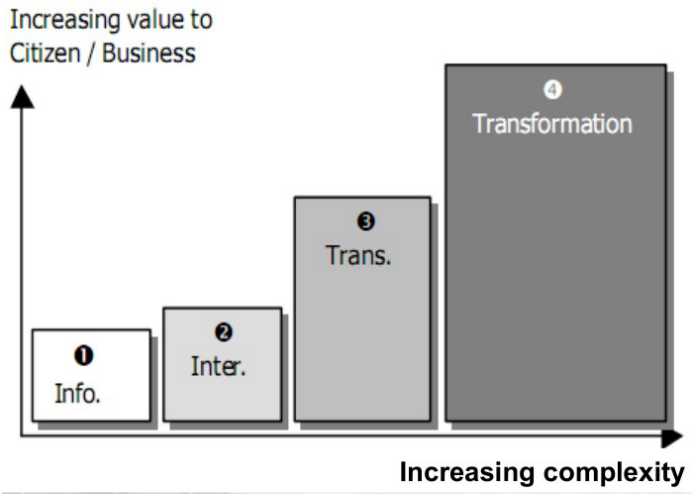



Fig. 4 Gartner's four-stage model [10]

overall picture of governance were functions and concepts are conceived and organised.

*UN's Five-Stage Model*

The UN five-stage model [14] offers a global perspective were all countries are included despite being developed or developing. The UN model identifies the five stages based on his form of benchmarking is a based primarily on analysing website content, special features, the quality and type of information offered and the capacity to conduct online transactions.

*Stage 1: Emerging:* The emerging stage is one were governments develop an online presence with websites giving basic information.

*Stage 2: Enhanced:* At this stage, governments are expected to have fully functional
websites which are regularly updated with dynamic content available.

*Stage 3: Interactive:* The interactive level is reached when citizens can have some form of interaction with their government through emails, forms and other interactive means that offer a two-way type of communication.

*Stage 4: Transactional*: At this stage, governments should be able to offer services that enable citizens and business be able to pay for services and good without going to a physical government institution.

*Stage 5: Seamless:* The last stage is where governments have integrated all systems and processes across different departments. E-services across administrative and functional boundaries should be fully integrated.

## IV. Challenges in Implementing E-Government

The challenges in implementing e-government in Africa and sub-Saharan Africa will be reviewed from available literature. Table 1 summaries the challenges.

*ICT Infrastructure*

ICT infrastructure in developing countries has played a role in negating how quickly e-government has been implemented. Limited access to ICT services is a great impediment and mobile telephony has great potential change the current status [15], [16]. The minimum threshold level of technological infrastructure in Africa which is one of the prerequisites to e-government implementation places most African countries below the minimum prerequisites [17], [18]. In some instances, despite having ICT infrastructure in place, the is lack of coordination and integration [19]. The poorly developed ICT infrastructure can be attributed to high costs in technology acquisition and
deployment (such as initial costs for setting up the ICT backbone infrastructure) and high
costs to access Internet-enabled ICT platforms.

*Financial*

The lack of financial support is a significant obstacle in implementing e-government [2], [22]. The high cost of

implementation and maintenance ICT infrastructure poses a challenge despite governments having a plan for effective and accessible e-government [2], [21]. The lack of allocated budget for e-government implementation results in constrained efforts due to considerable system requirements [23].

*Privacy and Security*

Privacy is a major concern the world over be it in developing or developed countries, and if not handled well can lead to poor adoption of e-government [10], [24], [26]. The lack of legal frameworks concerning cyber-security, digital signatures and personal data protection and confidentiality, privacy laws and access to information legislation in most African countries poses as a great challenge in adopting e-government [17], [25].

*Culture*

Technologies slowly and with great deliberation [37], [38]. However, cultural issue is not easily tangible, it must be given more planning so that technical change
is implemented successfully [30], [37]. In order to overcome cultural challenges, governments should consider the relationship of national culture and propensity to change and devise ways on how best to go around it [27], [29].

*Policy and Regulation Issues*

E-government has been said to be a nontechnical issue, but rather an organizational issue [38]. To implement e-government, principles and functions require a range of new rules, policies, laws and governmental changes to address electronic activities that spur. E-government entails signing contracts or digital agreements, which have to be protected and recognized by formalized law, which protect and secure activities and
processes which are not yet implanted in some developing countries [37]. Developing countries need to develop appropriate policy frameworks, supported by legislation for e-governance [17] and develop an e-government policy and strategy [32].

*Digital Divide*

Digital divide is the gap in opportunity between those that have access to the internet and computers, and those who do not [2]. The digital divide puts people with little to no access to computers and the internet at a disadvantage to access e-government services. [34] stated that a survy by the United nations in 2012 showed that there exists a wide gap in e-government development between developed and developing countries due to the digital divide. Economic factors such as economic status, and social factors such as educational attainment were identified as among the main predictors of the digital divide [33]–[36].

## IV. E-GOVERNMENT IMPLEMENTATION IN ZAMBIA

Zambia can be said to have embarked on developing an e-government strategy around 2003 when the development of the national ICT Master plan started which was subsequently approved by the Zambian Government in 2005 [39]. Since the development of the ICT Master Plan, Zambia has made great strides in realizing its implementation of an e-government. In 2017, the ICT Master plan was revised which now has been developed for the period 2017 – 2030 [40].

TABLE 1: E-government challenges

| | Category | Challenges | Literature |
|---|---|---|---|
| 1 | ICT infrastructure | Issues to do with infrastructure development, technological incompatibility, complexity, newness of technology are some challenges that can potentially affect e-government development | [15]–[19] |
| 2 | Financial | Issues to do with financial support, budgets to maintain and procure ICT equipment can lead to significant challenges for e-government implementation | [20]–[23] |
| 3 | Privacy and security | Privacy concerns, cyber-security, digital signatures and personal data protection and confidentiality issues have an effect on e-government adoption | [10], [17], [24]–[26] |
| 4 | Culture | Cultural acceptable norms and behavior can influence peoples acceptance and willingness to use e-government services | [20], [27]–[31] |
| 5 | Policy and Regulation Issues | Most developing countries do not have adequate policies and regulations which are required to protect and formalize laws to safeguard e-government implementation | [17], [20], [32] |
| 6 | Digital divide | Lack of access to ICT's and internet has an adverse effect on e-government because services cannot be accessed by people affected by the divide | [20], [33]–[36] |

Through the Zambia Information & Communications Technology Authority (ZICTA), a number of project have been embarked on which has demonstrated Zambia's commitment to move in tandem with the global village that we leave in [41]. The Zambian government further went ahead to set up a division under State house with a proposal to have an e-cabinet implemented to foster appreciation by policy makers in the benefits of e-government in 2015 [42] to give the much needed push in achieving its realization. In 2017, Zambia finished building its National Data Centre (ZNDC) worth $ 75 million that it is built to Tier III standard. The ZNDC is the first phase of the Smart Zambia project in conjunction with ZICTA valued at $ 440 million [43]. ZICTA has been working on a number of projects [41].

The National Numbering and Addressing project to ensure accurate and traceable telephone, mobile numbers, and residential and business addresses leading to identification of street addresses and postal addresses to enable provision of E-commerce, Social services and Socio-economic inclusion [41], [44].

With Smart Zambia project under phase 1, ZICTA established an ICT Talent Training Centre at the Zambia Information & Communications Technology College (ZICTC) in Ndola with the objective of providing quality ICT talent training delivered in a modern and well-equipped ICT laboratory environment. A learning management system is also used to offer training to the public and private institutions as well as individuals for certification and non-certifications courses [41].

ZICTA has been offering technical and financial assistance to the Ministry of Education (MoE) to acquire computing devices to assist in the administration of ICT based courses and examinations across the country targeting one hundred and ninety (190) centres [41].

A Computer Assembly Plant Project envisages to design, construct, install, commission and operate a computer assembly factory plant through a Private Partnership Programme (PPP). The plant is expected to produce a minimum of 250,000 (laptops, desk top computers and tablets) annually mainly for education sector with the capability of integrating local content. High end computers for retail and government ministries are also expected to be produced by 2020 [41].

Zambia plans to change its national registration cards issued with electronic national registration cards which will be able to seamlessly store integrated public citizen's information. When implemented, Zambia will the first country in eastern and southern Africa to introduce multipurpose electronic national registration cards that will be used for voting, accessing banking services and as a driving license [44]. In promoting universal access, Zambia liberalized the international gateway to spur tariff reductions down in as far as access to the internet and usage of ICTs is concerned [45]. Other notable projects done which gives a boost to e-government implementation are [45]:

- The Zambia Telecommunication Company (ZAMTEL) has erected, at a cost of USD48 million, a fibre optical network which spans a distance of 1,913 kilometres meant to provide faster data transmission and increased bandwidth in information interchange.
- The Copperbelt Energy Company (CEC) with approximately 700 kilometres of optical fibre network targeted to improve internal communication and increased bandwidth.
- The Zambia Electricity Supply Company (ZESCO) offers a fibre optical network dubbed "FiberCom" (broadband network).
- Broadband network for transmission of electrical energy, as a communication medium for data transfer and carrying internet traffic for video, voice, and so forth. This network covers a distance of approximately 1,700 kilometres from Sesheke to Lumwana through Livingstone, Lusaka and the Copperbelt. Upon completion, this network will cover approximately 3,000 kilometres.
- There are efforts by ZESCO and ZAMTEL to link their fibre optical networks to the Eastern Africa Submarine Cable System (EASSy) which is since operational.
- Introduction of government local and wide area network based on ring topology fibre networks and very small aperture terminal (VSAT) technology to encourage seamless inter-organ service integration, improve internal communication, and improve bandwidth.
- Establishing of affluent information systems and correspondingly e-government building blocks such as the integrated financial management information system (IFMIS) project, payroll management and establishment control project, the justice case management system intelligent human resource information systems, and payroll management and establishment control project.
- Other private entities such as Real Time Zambia has commissioned its independent end-to-end fibre project enabling corporate entities to link via fibre to the rest of the world through the SAT3 undersea cable

Policy wise, the ICT Master Plan covering 2017-2030 was updated in 2017 [40] and the Seventh National Development Plan (7NDP)[46] were produced. The 7NDP aims to enhanced ICT's using three key strategies which include:

1. Strengthen the legal framework of ICT
2. Improve ICT infrastructure for service delivery and
3. Provide electronic services

In 2018, the Information and Communications Technology Association of Zambia Bill was enacted which will provide for the registration of information and communications technology professionals and regulate their professional conduct in the interest of the information and communications technology sector [47].

## V. Conclusion

This paper aimed at looking at the different e-government models available and challenges faced by developing countries in implementing e-government. The models presented have a number of stages which are similar as they highlight areas that are easily identifiable in other models. As for the challenges, the ones looked at in this paper are not conclusive, but could be considered among the major ones. New research targeting specific countries can help to re-evaluate these challenges.

Zambia's efforts in implementing e-government were discussed to show what has been taking place since the realization to implement e-government. From literature, it can be deduced that almost all models point to the same stages of e-government development, and each country needs to choose a model to use in developing an e-government strategy with clear targets to achieve each stage.

## Reference


[1] R. Silcock, 'What is E-government', *Parliam. Aff.*, vol. 54, no. 1, pp. 88–101, Jan. 2001.

[2] M. Alshehri and S. Drew, 'Implementation of e-government: advantages and challenges', in *International Association for Scientific Knowledge (IASK)*, 2010.

[3] S. T. Ahmed, C. Sparkman, H. T. Lee, and D. Loguinov, 'Around the web in six weeks: Documenting a large-scale crawl', in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 1598–1606.

[4] UN EGDI, 'Data Center', 2018. [Online]. Available: https://publicadministration.un.org/egovkb/en-us/Data-Center. [Accessed: 17-Aug-2018].

[5] K. Siau and Y. Long, 'A stage model for e-government implementation', in *15th Information Resource Management Association International Conference (IRMA'04) New Orleans, LA*, 2004, pp. 886–887.

[6] D. Olowu, 'Bridging the digital divide in Africa: Making the governance discourse relevant', *Afr. Netw. Dev. Inf. ICTs Gov.*, pp. 81–91, 2004.

[7] The Zambian, 'Information and Communication Technology Policy', *The Zambian*, 06-Apr-2007. .

[8] V. Kumar, B. Mukerji, I. Butt, and A. Persaud, 'Factors for successful e-government adoption: A conceptual framework.', *Electron. J. E-Gov.*, vol. 5, no. 1, 2007.

[9] T. Almarabeh and A. AbuAli, 'A general framework for e-government: definition maturity challenges, opportunities, and success', *Eur. J. Sci. Res.*, vol. 39, no. 1, pp. 29–42, 2010.

[10] K. Layne and J. Lee, 'Developing fully functional E-government: A four stage model', *Gov. Inf. Q.*, vol. 18, no. 2, pp. 122–136, Jun. 2001.

[11] K. V. Andersen and H. Z. Henriksen, 'E-government maturity models: Extension of the Layne and Lee model', *Gov. Inf. Q.*, vol. 23, no. 2, pp. 236–248, Jan. 2006.

[12] Y. Long and K. Siau, 'Synthesizing e-government stage models – a meta-synthesis based on meta-ethnography approach', *Ind. Manag. Data Syst.*, vol. 105, no. 4, pp. 443–458, May 2005.

[13] C. Baum and A. Di Maio, 'Gartner's four phases of e-government model', *Gart. Group*, vol. 12, 2000.

[14] S. A. Ronaghan, 'Benchmarking e-government: a global perspective', *Assess. Prog. UN Memb. States U. N. Div. Public Econ. Public Adm. Am. Soc. Public Adm.*, 2002.

[15] A. B. Adeyemo, 'E-government implementation in Nigeria: An assessment of Nigerias global e-gov ranking', *J. Internet Inf. Syst.*, vol. 2, no. 1, pp. 11–19, Jan. 2011.

[16] K. J. Bwalya, 'Factors affecting adoption of e-government in Zambia', *EJISDC*, vol. 38, pp. 1–13, Jan. 2009.

[17] N. J. Hafkin, 'E-government in Africa: An overview of progress made and challenges ahead', in *Prepared for the UNDESA/UNPAN workshop on electronic/mobile government in Africa: Building Capacity in Knowledge Management through Partnership. United Nations Economic Commission for Africa*, 2009, pp. 17–19.

[18] R. Heeks, 'e-Government in Africa: Promise and practice', *Inf. Polity*, vol. 7, no. 2,3, pp. 97–114, Jan. 2002.

[19] M. Mzyece, 'A critical analysis of e-government in Zambia : Section II : Country perspectives on e-government emergence', *Afr. J. Inf. Commun.*, vol. 2012, no. 12, pp. 110–127, Jan. 2012.

[20] M. Alshehri and S. Drew, 'Challenges of e-government services adoption in Saudi Arabia from an e-ready citizen perspective', *World Acad. Sci. Eng. Technol.*, vol. 66, no. June, 2010.

[21] A. Carvin, J. Hill, and S. Smothers, 'E-government for all: Ensuring equitable access to online government services', in *The EDC center for media & community and the NYS forum*, 2004.

[22] M. J. Moon, 'The evolution of e-government among municipalities: rhetoric or reality?', *Public Adm. Rev.*, vol. 62, no. 4, pp. 424–433, 2002.

[23] N. Nkwe, 'E-government: challenges and opportunities in Botswana', *Int. J. Humanit. Soc. Sci.*, vol. 2, no. 17, pp. 39–48, 2012.

[24] K. J. Bwalya and M. Healy, 'Harnessing e-government adoption in the SADC region: a conceptual underpinning', *Electron. J. E-Gov.*, vol. 8, no. 1, p. 23, 2010.

[25] P. Ngulube, 'The nature and accessibility of e-government in sub-Saharan Africa', *Int. Rev. Inf. Ethics*, vol. 7, no. 9, pp. 1–13, 2007.

[26] N. P. Rana, M. D. Williams, and Y. K. Dwivedi, 'Analysing challenges, barriers and CSF of egov adoption', *Transform. Gov. People Process Policy*, vol. 7, no. 2, pp. 177–198, May 2013.

[27] O. E. M. Khalil, 'e-Government readiness: Does national culture matter?', *Gov. Inf. Q.*, vol. 28, no. 3, pp. 388–399, Jul. 2011.

[28] F. Li, 'Implementing E-Government strategy in Scotland: current situation and emerging issues', *J. Electron. Commer. Organ. JECO*, vol. 1, no. 2, pp. 44–65, 2003.

[29] S. M. Mutula, 'E-government divide: Implications for sub-Saharan Africa', *Inf. Ethics Afr. Cross-Cut. Themes*, pp. 59–69, 2013.

[30] J. Y. Weisinger and E. M. Trauth, 'The importance of situating culture in cross-cultural it management', *IEEE Trans. Eng. Manag.*, vol. 50, no. 1, pp. 26–30, Feb. 2003.

[31] F. Zhao, 'Impact of national culture on e-government development: a global study', *Internet Res.*, vol. 21, no. 3, pp. 362–380, Jan. 2011.

[32] M. Chango, 'Challenges to e-Government in Africa South of Sahara: A Critical View, and Provisional Notes for a Research Agenda', in *Proceedings of the 1st International Conference on Theory and Practice of Electronic Governance*, New York, NY, USA, 2007, pp. 384–393.

[33] K. Bagchi, 'Factors Contributing to Global Digital Divide: Some Empirical Results', *J. Glob. Inf. Technol. Manag.*, vol. 8, no. 3, pp. 47–65, Jul. 2005.

[34] A. Collier, H. Deng, and F. Zhao, 'A multidimensional and integrative approach to study global digital divide and e-government development', *Inf. Technol. People*, vol. 27, no. 1, pp. 38–62, Feb. 2014.

[35] P. DiMaggio and E. Hargittai, 'From the "Digital Divide" to `Digital Inequality': Studying Internet Use As Penetration Increases', p. 23.

[36] S. Kiiski and M. Pohjola, 'Cross-country diffusion of the Internet', *Inf. Econ. Policy*, vol. 14, no. 2, pp. 297–310, Jun. 2002.

[37] M. Alshehri and S. Drew, 'E-government fundamentals', in *IADIS International Conference ICT, Society and Human Beings*, 2010.

[38] F. Li, 'Implementing E-Government Strategy in Scotland: Current Situation and Emerging Issues', *J. Electron. Commer. Organ. JECO*, vol. 1, no. 2, pp. 44–65, Apr. 2003.
[39] V. Weerakkody, Y. Dwivedi, L. Brooks, M. Williams, and A. Mwange, 'E-government implementation in Zambia: Contributing factors', *EG*, vol. 4, pp. 484–508, Jan. 2007.
[40] C. Manzi, 'SMART ZAMBIA: NATIONAL ICT MASTERPLAN RE-ALIGNMENT COMPLETED', *Mwebantu.News*, 21-Jan-2018. .
[41] 'ZICTA - Zambia Information & Communications Technology Authority'. [Online]. Available: https://www.zicta.zm/#tab-a4. [Accessed: 10-Oct-2018].
[42] Lusaka Times, 'Zambia : President Lungu Launches E-Government Division', *LusakaTimes.com*, 22-Oct-2015. .
[43] M. Freedman, 'Zambia: National Data Center, an essential component of the Smart Zambia government project, achieved 98% - Extensia', 07-Feb-2018. .
[44] A. Nkunika, 'ADDRESSING THE CHALLENGE OF PATH DEPENDENCY IN THE CREATION OF SUSTAINABLE INFORMATION AND COMMUNICATION TECHNOLOGIES FOR DEVELOPMENT AN ANALYSIS OF ZAMBIA'S E-GOVERNMENT SYSTEM', 2017.
[45] T. Du Plessis, C. Rensleigh, and K. J. Bwalya, 'E-government implementation in Zambia – prospects', *Transform. Gov. People Process Policy*, vol. 8, no. 1, pp. 101–130, Mar. 2014.
[46] 'UNFPA Zambia | Zambia's Seventh National Development Plan (2017-2021) Implementation Plan'. [Online]. Available: https://zambia.unfpa.org/en/publications/zambias-seventh-national-development-plan-2017-2021-implementation-plan. [Accessed: 01-Oct-2018].
[47] 'Downloads – Information & Communication Technology Society of Zambia'. .

# Web Design Tools: Challenges and Open issues

Gondwe Raphael

*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*

gondweraphael@gmail.com

Douglas Kunda

*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*

dkunda@mu.edu.zm

**Abstract:**
**We are living in this world where technology has taken a center stage in all human endeavours. Accessing and sharing of information is just by the finger tips and this is facilitated by the web (Internet). The web in this scenario imply the way of exchanging information between computers on the Internet, tying them together into a vast collection of interactive multimedia resources. It is for this reason that web designers should carefully look at open issues and challenges of web design as regards to Hypertext, Hypermedia, Markup languages and XML related technologies. If only designers and developers of the web can embrace and address these issues and challenges pertaining to web design, then when can realise the potential merits that they may bring on board. Thus, the prime objective of this paper is to provide systematic theoretical discussion on web technologies in terms of open issues and challenges.**

***Keywords—Web design, Web Technologies, Hypertext, Hypermedia, Markup Language, XML, Security,***

## 1. INTRODUCTION

The web which is seldom referred to as World Wide Web (W3) is regarded as the way of exchanging information between computers on the Internet, tying them together into a vast collection of interactive multimedia resources. In other words, a web is basically a system of Internet servers that support specially formatted documents which are in a markup language called HTML (HyperText Markup Language) that supports links to other documents, as well as graphics, audio and video files. Internet is the tool that is considered to be cardinal for communication, learning, socialisation, shopping among others. For instance:

The usage of web technologies in e-learning are further enhanced with the web 2.0, which is a set of economic, social, and technology trends that facilitate a more socially connected Web where everyone is able to add to and edit the information space (Anderson, 2007). These include blogs, wikis, multimedia sharing services, content syndication, podcasting and content tagging services (Anderson, 2007).

It is from this background that web designers and developers should research more on who accesses the web, what type of web service they need and the approximate number of Internet users. Besides, security on the web can be compromised hence, developers should take keen interest on web technologies to use.

Web designers should take into consideration what users want in a website in order for it to perform well. What users want in a website is an important area to be considered (Palmer, J. W., 2002). The speed at which your website loads decides if it will be able to keep the users with it, or they will run away faster than your website's loading speed.

In addition, the web application runs on top of the Internet, hence the reason of web designers to take special attention on how user interfaces are created. Web designers should have it in mind that user interface defines the interaction that takes place between the end users and the web application as a result, it plays a major role of whether users will revisit the web site (Saputra, D. G., & Azizah, F. N., 2013). Reflecting more on web security, the web is prone to attacks as it is complex (consists of a large number of components and technologies) thus, web applications attacks can hardly be ignored. El-Hajj et al., (2016) reported that web applications lack sufficient built-in security assurance, as a result expose information to a third party. Thus, there is need for effective defense to the web platform in order to safe guard it.

This paper will discuss hypertext, hypermedia, markup languages and XML related technologies.

## 2. OPEN ISSUES AND CHALLENGES

### A. Hypertext, Hypermedia, Markup Language and XML Technologies

Despite the many exciting advances in web technology advancement, there is still more to be accomplished in terms of wed design and development. Baturay .M. & Birtane .M., (2013) writes that web designers have been striving to provide users with better web browsing to meet users' needs just like those on traditional website layouts. In computing, web technology can be regarded as the method by which computers communicate with each other through the use of

markup languages and multimedia packages.

The terms hypertext and hypermedia are sometimes used interchangeably but there is a difference. Hypertext is a powerful cross-referencing tool meant for user-driven access to an ocean wealth of interconnected information either static or dynamic in an electronic format. In other words, hypertext may refer to plain simple text that contains links to access other chunks of text within the same or different document. Hypermedia is an extension of hypertext that employs multiple forms of media such as text, graphics, audio or video sequences, still or moving graphics, etc. The structure of hypermedia is quite similar to that of a hypertext, except it's not constrained to be just text-based. It extends the capabilities of hypertext systems by creating clickable links within web pages to create a network of interconnected non-linear information which user can both access and interact with for a better multimedia experience.

The conclusion that can be made in terms of application is that both hypertext and hypermedia follow a similar structure comprising nodes that are interconnected by links except in hypermedia systems, the nodes can contain multiple forms of media such as text, images, audio, video and graphics. In terms of implementation, hypertext is used to represent multimedia content in electronic text format whereas hypermedia combines both hypertext and multimedia to provide access to a wealth of information usually in a non-linear sequence. The idea of hypermedia is to extend the functionality of multimedia elements to make content more interactive and better accessible than before.

When you are trying to comprehend what markup languages are, in other words, you are getting to know on how information is added to a document that enhances its meaning in certain ways, in that it identifies the parts and how they relate to each other. Markup means to structure it in a specific format hence, markup language because it is a language that allows users to organise, improve the appearance of, and link text with data on the internet.

XML (extensible Markup Language) is a markup language that allows users to define a set of tags which describe the structure of a document (Bray et al 2000, Miller et al. 1998, Sorenson and wood 1998, St Laurent 1997). XML provides a structured representation of data that can be implemented broadly and is easy to deploy.

What web designers and developers ought to know is that end users are exposed to a variety of devices that can aid them in accessing Internet and they should not restrict them to the traditional way of using desktops.

This sounds as a wakeup call to web designers and developers in coming up with web technologies that will embrace such gadgets. web designers and developers are supposed to use a responsive web design as it enables users to surf the web through a multi-device and not limited to desktops (Baturay .M., & Birtane .M., 2013). Besides, Baturay .M. & Birtane .M. (2013) notes that the key technical features of a responsive web design are:

- **Media queries and screen resolutions**: a web designer use HTML and CSS3 media queries to help the website decide content to be viewed depending on each devices' screen.
- **Fluid grid layouts**: responsive web design using fluid proportion based grids works on multiple devices. By doing so, it enables the content to resize as well as rearrange on the expansion or contraction of webpage grid depending on the width.
- **Flexible images and media**: The responsive web design by using dynamic resizing changes the layout and resizes the images hence three dynamic screens are designed at the same time, that is, desktop screen, mobile device screen like Ipad as well as smartphone screen like Iphone.

The highlighted web technologies have issues and challenges when it comes to web design and development. Following are the discussions:

Content search is one of the issues that demands a lot of attention. This is where you search over all the nodes and potentially over all the links individually and look for nodes and links that match a specific pattern. The content of those nodes and links should match a specific pattern.

The other issue that needs a lot of research on is the issue of augmenting the Node/Link Model. Basic hypermedia model lacked a composition mechanism, that is, a way of representing and dealing with groups of nodes and links as unique entities separate from their individual components. Hypermedia should support inclusion, that is, part-of, relations as a construct distinct from the standard kind of reference links (Frank G. Halasz 1990).

Discovered structures is yet another issue, this means links or composites that are created by the system on the basis of computing a similarity among nodes. The other issues include:

**Availability of Content**

Semantic Web content is web content annotated according to particular ontologies, which define the meaning of the words or concepts appearing in the content. Before the notion of the Semantic web existed, we were involved in an experiment ([Decker et al], [Benjamins et al Creating Semantic Web content is therefore a serious challenge for the Semantic Web. Since the infrastructure of the Semantic Web

is still being built (RDFS, OIL, DAML+OIL, etc.), currently, there is little Semantic Web content available. Apart from the infrastructure, researchers are currently building tools to support semantic annotation of web content. Such tools are important and critical to the success of the Semantic Web. However, they have two limiting characteristics: 1. Most of them annotate only static pages, and 2. Many of them focus on creating new content. This leads to the following not optimal situation:

- Dynamically generated content is not considered. 'Dynamic' content is generated from databases, and according to a study of [Lawrence & Giles 1999] referring to it as the 'Deep Web'- in March 2000, [Michael K. Bergman] its size is estimated to be 400 to 550 times larger than the commonly defined World Wide Web, which includes more than one billion static web pages.
- Existing content is running the risk to be excluded from the Semantic Web, even though XML content is gaining ground on content share.
- Extract the dynamic content from its source, annotate it (as if it were static) and store it. The problem here is the almost infinite amount of static pages that can be generated from a dynamic (database powered) site, and the almost continuous updates, creations and removals of generated static pages when data changes in the databases.
- Leave the content in the database and annotate the query that retrieves the concerned content. This option is less space-consuming, and consistency in the annotations is guaranteed with respect to the underlying sources of information, since the content is dynamically generated/annotated when retrieved from its sources.

**Ontology Availability, Development and Evolution**

An ontology is the working model of entities and interactions in some particular domain of knowledge or practices, such as electronic commerce or "the activity of planning." In artificial intelligence (AI ), an ontology is, according to Tom Gruber, an AI specialist at Stanford University, "the specification of conceptualisations, used to help programs and humans share knowledge." In this usage, an ontology is a set of concepts such as things, events, and relations - that are specified in some way (such as specific natural language) in order to create an agreed-upon vocabulary for exchanging information. In Semantic Web, ontologies are key reason they are the carriers of the meaning contained in the Semantic Web that is, they provide the vocabulary and semantics of the annotations.

There are three main issues to be solved regarding this challenge, the first two issues are related to traditional ontology development problems that haven't been solved completely until now, and the last one is much more related to the new context of the Semantic Web:

The first is the construction of kernel ontologies to be used by all the domains. Initiatives exist for the construction of some of these kernel ontologies in different domains: the IEEE Standard Upper Ontology Group8 aims to create a common unified top level ontology); initiatives in the e-commerce domain also exist, such as UNSPSC9, eclass10, RosettaNet11, NAICS12, etc.

The second is to provide methodological and technological support for most of the activities of the ontology development process [Fernandez-1997], including:

- Knowledge acquisition, conceptual modelling and ontology coding in Semantic Web languages (RDFS, OIL, DAML+OIL), and new languages that may arise in the coming years [Maedche, Staab – 2001].
- Ontology alignment and mapping, ontology integration, ontology translation tools, and ontology reengineering tools if existing ontologies are going to be reused [Fensel et al, 2001], [Noy, Musen 2000].
- Consistency checking tools for the ontologies to be reused [Gomez-Perez 1996].

Ontology representation is the most fundamental issue in ontology development. In addition to making ontologies understandable by computers and humans, an ontology representation language should also provide representation adequacy and inference efficiency. The Standardisation of ontology representation languages (e.g., RDF, RDFS [15], and OWL [16]) has taken big strides in the past few years. The above languages have mainly adopted a frame-based knowledge representation paradigm [17], some of which (e.g., OWL) incorporate description logics to enhance the expressiveness of reasoning systems (Lina Z, 2007).

Ontology acquisition refers to the creation of the content of ontologies such as concepts and relations. Given the strong dependence on domain knowledge, ontology modeling is traditionally carried out by knowledge engineers and/or domain experts. As domain knowledge evolves rapidly and workforces become increasingly distributed, subject-matter experts are not easily accessible and the experts' knowledge is likely to be incomplete, subjective, and even outdated. To keep up with the requirements of practice, people turn to other sources such as dictionaries, Web documents, and database schemas for the content of ontologies. As a result, ontology acquisition can significantly benefit from ontology learning. (Lina Z, 2007).

Ontology evaluation is another major issue as ontologies become available. Ontology evaluation can enhance the quality of ontologies, improve the inter-operability among systems, and further increase the wide adoption of ontologies. Ontologies can be evaluated from a variety of perspectives, ranging from content to technology, methodology, and application, using objective measures such as completeness, consistency, and correctness. Among others, ontology learning is regarded as an alternative method to evaluate the content of ontologies.

Ontology maintenance pertains to how to organize, search, and update existing ontologies. This issue looms large as more and more ontologies are accumulated. The constant evolving of the environment of ontologies makes it very important for ontologies to be evaluated and maintained to keep up with the change. SWOOGLE is a crawler-based indexing and retrieval system for the semantic Web, housing several millions of Web documents in RDF and OWL. The enormous number of semantic Web pages and ontologies makes manual ontology maintenance a daunting task. (Lina Z, 2007).

**Scalability of Semantic Web Content**

The first one is related to the storage and organisation of Semantic Web pages. The 'basic' Semantic Web consists of ontology-based annotated pages whose linking structure reflects the structure of the WWW, that is, pages connected to others by means of hyperlinks. This hyperlinked configuration does not fully exploit the underlying semantics of Semantic Web pages.

The use of **semantic indexes** to group Semantic Web content based on particular topics is a necessary step to make applications able to aggregate content in order to provide added value services. Semantic indexes will be generated dynamically using ontological information and annotated documents [Gomez-Perez 1996].

**Multilinguality**

Studies on language distribution over the WWW content show that even if English is the predominating language for documents, there exist an important resources written in other languages: English 68.4%, Japanese 5.9%, German 5.8%, Chinese 3.9%, French 3.0%, Spanish 2.4%, Russian 1.9%, Italian 1.6%, Portuguese 1.4%, Korean 1.3% ,Other 4.6% [Source: Vilaweb.com, as quoted by eMarketer]. The diversity of languages is much more acute for European WWW resources. Multilinguality plays an increasing role at the following levels: at the level of ontologies, of annotations and of user interface.

**Maintenance of Behaviour**

If one thinks about behaviour as something that has a fixed duration and does not have a lasting impact of any kind, it is straight-forward to keep track of what the body is doing. At any given time, one can simply sum up all the behaviours currently executing to see exactly what is going on. However, reality is not that simple. When a behaviour command is completed, the body is typically left in a new state, possibly even maintaining the behaviour until another behavior replaces it. One example is the gaze behaviour. If a character is asked to gaze at a certain in C. Pelachaud et al. (Eds.): "Intelligent Virtual Agents 2007", Lecture Notes in Artificial Intelligence 4722: 99-111, Springer-Verlag Berlin Heidelberg target with the command <gaze target='person1' stroke='g1: stroke'/>, it is clear that the gaze will fully rest on the target at exactly the same time another behaviour (g1) reaches its own moment of greatest effort. However, it is completely left undetermined what happens next. Does the character continue to look at person1? If person1 starts moving, should the character adjust its gaze accordingly? If the character is being requested to track person1 with its gaze, how should that be indicated and how long should it last?

**Constraints and Synchronisation**

The synchrony achieved through aligning behaviour sync-points demonstrates one type of timing constraint, but it is important to consider other types as well. These include:

- Physical characteristics constraints: Limitations on body movement speed.
- Rhythmic constraints: Requirement that different modalities stay in perfect synchrony.
- Global rule constraint: A general synchrony rule that is meant to hold true for all behaviors in a set of behaviors (such as stating that all gesture strokes should coincide or precede emphasized syllables).
- Fixed signal constraint: Synchronization to an external source with a fixed timing pattern (such as clapping to music).
- Perceptual signal constraint: Synchronization to an external signal with unknown timing (such as the actions of another person).

**Lack of quality control**

Reliance on primitives and hard-wired quality-control techniques is one of the serious limitations of existing approaches. Website developers cannot customise these approaches on their specific requirements because they are typically embedded in their host systems. Defining new approaches is also another challenge that website developers struggle with. Even though tools like TurKit that rely on current crowdsourcing systems let users define some quality control process, using these tools require programming

skills such as Java or C++. Baresi et al., 2006 said web developers' blame is on poor software quality on continuous change such as removing errors, meeting customers' expectations and improving the implementation. Below are some of the challenges of web design:

**Lack of skilled engineers in web software development**

There is an acute shortage of skilled IT engineers against a large number of IT jobs. This means that companies will often resort to hiring IT engineers with less skills and education than desired (Offutt, 2002).

**Lack of a common software standardisation**

Another challenge of web applications is on how the software in the different environments can be combined to function into a composite system. And also, how to build a coherent application out of a large collection of unrelated software modules (Vermesan et al., 2011). Below are some of the standardizations that can be addressed:

- Micro operating systems that are energy efficient.
- Distributed self-adaptive software for self-optimization, self-configuration, self-healing (such as autonomic).
- Self-management techniques in order to overcome increasing complexities.
- Password distribution mechanisms for increased privacy and security.
- Bio-inspired algorithms as well as game theory.

Vermesan et al., (2002) said standards have an important role to play both within an organization or entity and across organsizations because when these wish to share information or exchange information, standards allow them to do so in an efficient manner hence reducing ambiguity about the interpretation of the information being exchanged.

**Lack of reliability**

It is important to know that web software is critical to the commercial success of many businesses. Therefore, if the software does not work reliably, then the businesses will not succeed. The user base for web software is very large and so they expect the web applications to work as reliably as possible. So, if the web application does not work well, the users will not have to move further, instead they will point their browser to another URL. Hence if web software is not reliable, websites depending on customers will lose them and the business may loose very huge sum of money. Reliability of web software is crucial and companies can afford to spend resources to ensure high reliability (Offutt, 2002).

**Not meeting usability requirements**

Website users expect software to be very simple to learn how to use. Even though there is vast knowledge for how to develop usable software and websites, still many website do not meet the usability requirements as per expectation. Websites that are not usable will not be used by customers because they will quickly switch to more usable websites as soon as those are online (Offutt, 2002).

**Lack of security**

There are so many potential security problems in web software applications such as websites being cracked into and private information for customers being held for ransom. This can cause loss in revenue, legal consequences and lose customer credibility among others. Thus it is essential that web software applications handle customer data and all other electronic information as securely as possible. Software security is one of the growing research areas in computer science because web software developers are facing a huge shortfall in terms of both available knowledge and personnel who have the knowledge that is available (Hu et al., 2017; Offutt, 2002).

**Stakeholders often don't know beforehand what they expect from a system**

The existence of many stakeholders results in conflicting and intrinsically decentralised requirements. This is because some stakeholders usually do not know beforehand what they expect from a system. Thus it is an illusion to exhaustively gather requirements and preplan the process to avoid future changes: changes is not a nuisance to avoid but an intrinsic factor to address (Baresi et al., 2006).

## B. PERFORMANCE

Performance can be defined as the end-user responsiveness of system under various loading conditions. Most of the product related issues could be resolved if reliability tests are conducted on the workflow as those that need performance benchmarking. Capacity of a unit of production system can be determined by monitoring requests per second (RPS), requests per page and number of active users (van Eijl, 2002; Sia and Ho, 1997).

The Website performance in this regard can be determined by the speed in which web pages are downloaded as well as displayed on the users' web browser. Performance of a website is influenced by the characteristics of the design (Rodriques et al., 2017). If the web performance is good, it will respond efficiently to end-user interactions Therefore, this part of paper will focus on performance & development issues, performance and UI, performance & user navigation and performance & user engagement.

- **Performance and Development Issues**

Web technologies are used in web development in order for the website to perform better. According to web technologies perform well in terms of basic messaging rates. It is also important to determine what these technologies can offer for the other groupware requirements . Below are some deployment and development issues as regards performance of the web:

- *Graphics capabilities*. Browser-based graphics is still far behind stand-alone applications, but the development of tools such as the HTML5 Canvas and web ports of OpenGL mean that web-based graphics could soon approach the performance of plug-in technologies.
- *Access to devices and file systems.* The current security model of the web prevents web pages from accessing any devices or files outside a very narrow sandbox.
- *QoS control*. This is an area where web technologies are considerably less mature than stand-alone groupware approaches, providing essentially no support for application-level network control.

- **Performance and UI**

Usability is generally defined as a multidimensional property of a user interface described by five attributes - learnability, efficiency, memorability, errors, and satisfaction. A system with a higher learnability attribute is easier for the user to learn, navigate, and perform required operations. An efficient system is designed to support higher productivity levels, while a system with higher memorability attribute values is easier to remember. The error attribute property describes websites designed so that users make fewer mistakes while using the site.

Performance, as with usability, is also something web developers have to constantly keep in mind throughout the development cycle. Sukhpuneet et al., (2016) writes that usability is how well and easily a user can interact with a website without formal training. Every action taken during development has the potential to impact the performance of the finished site hugely. If during development, the web developer upload large images to test – that's going to come back to bite at a later stage. If not correctly manage the assets (css, javascript, images) from the beginning, then it's going to get harder and harder to optimise them as developers move forward. What is always wanted is a blazing fast site that looks amazing and performs exceptionally well for the user. That's the award winning trifecta of performance, usability and design that is strived for in every single project which comes through the studio, and that's usually what developers get if they have to manage to solve all of these problems.

Website users have grown to expect web software to be very simple to learn how to use. Although a lot of knowledge is available for how to develop usable software and web sites, many web sites still do not meet the usability requirements that most of the people expect. Web sites that are not usable will not be used: customers will quickly switch to more usable web sites as soon as they are put online. . Jankowski, J., and Decker, S., (2013) writes that websites should explain themselves, should not waste peoples' time because users simply scan web pages and if they find something that works, they simply stick to it and Sukhpuneet et al., (2016) adds that users leave one website for another if that site is not usable.

- **Performance and user navigation**

When it comes to websites, the navigation system acts like a road map to all the different areas and information contained within the website. Good website navigation makes it easy for the visitors to find what they want and for search engines to crawl which results in more conversions and greater search visibility. Website navigation (a.k.a., internal link architecture) are the links within your website that connect your pages. The primary purpose of website navigation is to help users easily find stuff on your site.

Search engines use the website navigation to discover and index new pages. Links help search engines to understand the content and context of the destination page, as well as the relationships between pages.

Website navigation is another part often neglected by developers. Navigation provide users with a strong sense of structure and place hence avoiding difficulty finding information (Webster, J., and Ahuja, J.S., 2016). Intuitive navigation creates a better user experience for the website visitor. Intuitive navigation is leading your audience to the information they are looking without a learning curve. And when the navigation is intuitive, users can find out information without any pain, creating a flawless experience preventing them from visiting the competitors. Webster .J

and Ahuja .J.S. (2016) defined computer-based navigation as "the decision and actions that contribute to a person's ability to find and examine data organized in the computer media". Therefore, web navigation system is one that aids users in the creation and interpretation of an internal mental model in order to find and examine data on a website. Navigation is one of the important characteristics that influence website performance (Rodrigues et al., 2017). Navigation system helps to avoid the user from being disoriented on the website.

Two types of navigation systems (global and simple) can increase web site performance. Global navigation system is one that provides the entire systems visual representation while simple navigation system just provides local links to other pages in the website. A simple navigation system suits in a small website where a link on every page is implemented and a navigation bar can link to the website's front page as well as other two sections of front pages. As for the global navigation system, it usually includes navigation bars with links to things like Catalog, About Us, Detailed Site Map, Table of Contents and also local Search Engine (Webster .J., and Ahuja .J.S., 2016).

A global navigation system is recommended over simple system because users can find information more quickly while simple system lack flexible path mechanism (Webster .J and Ahuja .J.S., 2016). Web designers believe that effective mechanism for navigation should be included in any usable hypertext system.

In other literature you may come across other terms that are used as categories of navigations and these include: structural, associative and utility.

***Structural***

Connects one page to another based on the hierarchy of the site; on any page you'd expect to be able to move to the page above it and pages below it.

***Associative***

Connects pages with similar topics and content, regardless of their location in the site; links tend to cross structural boundaries.

***Utility***

Connects pages and features that help people use the site itself; these may lie outside the main hierarchy of the site, and their only relationship to one another is their function

- **Performance and user engagement**

Being an important determinant in deciding the ranking of the website, User Engagement of a site requires all the attention it can get. Website user engagement is an important indicator determining the success and ranking of your site. The term user engagement may refer to any action taken by a user on a website such as posting a comment or a search query, filling in a contact form, signing up for a newsletter, etc. It also includes the time each user spends on a given website. The following should be addressed when designing websites: Reduce page load time, improve your internal linking structure, simplify navigation, choose your writing style and use a responsive design among others.

Web application designers should strive to come up with websites that are able to engage users. An engaging website will hold users attention hence they will be successful in the way they move across pages and through space of a website. Therefore, engagement relate positively to performance because the increased focus of attention lead to more effective web searches (Webster .J. & Ahuja .J.S., 2016).

Web designers should bear in mind that an engaging website have characteristics that will influence the user to use the website again. In other words, it should be visually appealing and elegant (include text, images and graphics) in order to bring back old users to the website as well as influence new users. Likelihood of return and satisfaction of user proves good performance of a web site (Palmer .J. W., 2002).

## C. SECURITY

One of the fastest growing research areas in computer science is that of software security and web software developers are facing a huge shortfall both in terms of available knowledge and personnel who have the knowledge that is available (Hu et al., 2017; Offutt, 2002).

Web applications are now more complex, dynamic and interactive providing users with bulky and sensitive information as well as a service. Web application security is the process of securing confidential data stored online from unauthorised access and modification. This is accomplished by enforcing stringent policy measures. Prokhoren et al., (2015) notes that, web applications are a valuable asset for sophisticated web-oriented attackers due to large amounts of private data processed and stored on them. Therefore, defense mechanisms are vital in web design in order to protect users and information of a website.

Web application security aims to address and fulfill the four conditions of security, also referred to as principles of security:

- Confidentiality: States that the sensitive data stored in the Web application should not be

exposed under any circumstances.

- Integrity: States that the data contained in the Web application is consistent and is not modified by an unauthorized user.
- Availability: States that the Web application should be accessible to the genuine user within a specified period of time depending on the request.
- Nonrepudiation: States that the genuine user cannot deny modifying the data contained in the Web application and that the Web application can prove its identity to the genuine user.

Web attackers and network attackers are some of the web threats found on the websites. Web attacker controls a server that responds to any HTTP request that is sent with malicious contents by an attacker while network attacker detect and intercept all traffic that is sent between two network nodes (Bugliesi et al., 2016). Attackers are able to inspect, forge and corrupt HTTP traffic that has been sent on the network. To avoid these web threats, SOP (same-origin policy) mechanism is offered by web browsers even though it is not enough to prevent many common attacks because of injection vulnerabilities. Injection vulnerabilities occur when validation user controllable data is missing or is insufficient (Deepa .G, and Thilagam P.S., 2016).

Many are web application security vulnerabilities that can compromise web security and the following are some of the threats discussed.

SQL injection (SQL) and Cross-site scripting (XSS) are rated as the top most threats by different security consortiums (Deepa .G. & Thilagam .P. 2016). Prokhorenko et al., (2015) also says SQL injection and XSS injection are the most common attacks. The application becomes vulnerable to SQL injection attack when the attacker access confidential information. And it becomes vulnerable to XSS attack when the attacker arranges for the server to produce a page that is able to execute a script constructed by an attacker. PHP scripting language represent first line of defense against SQL injection and XSS injection hence it becomes the first target for attackers. (Nguyen-Tuong et al., (2004).

SQL injection attacks are flaws enabling the attacker to compromise the database of the application due to lack of proper input sanitation and this, results in unwanted extraction and insertion of data into the database. As a result SQL injection attack succeeds if the attacker manages to predict the form of the generated SQL request because such knowledge enables an attacker to provide an input which in turn alters the structure of the initially-expected SQL request (Prokhorenko et al., (2015). In short, SQL injection attacks become possible if the attacker has HTML and SQL syntax. Huang et al., (2005) adds that malicious patterns resulting in execution of arbitrary SQL or system commands can be injected if data is improperly processed prior to SQL query construction. Tautology attacks, piggybacked queries, union queries, blind injection attacks, timing attacks, alternate encodings and attacks on stored procedures are some of the classifications of SQL injection vulnerability (Deepa .G. and Thilagam .P. S., 2016).

SQL injection attack can be addressed by secure programming, vulnerability detection & prevention and attack detection & prevention (Deepa .G. and Thilagam .P. S., 2016). In secure programming, the developer follows secure practices when developing an application. Vulnerability detection & prevention identify vulnerable injection points in which malformed data enter as well as propagate through the system. And finally, attack detection & prevention compare the structure of the query that is generated during normal and attack execution hence preventing the malformed query from being executed by the database (Deepa .G. and Thilagam .P. S., 2016).

XSS injection vulnerability enables the attacker to execute malicious scripts in the web browser of the client and occurs whenever a user supply input to the web application with no proper sanitization (Deepa .G. and Thilagam .P. S., 2016). It can be said that the browser executes malicious scripts whenever the user visits an exploited web page. When an XSS attack occurs, it leads to session hijacking, leakage of sensitive data, defacement of web content and cookie theft. It is therefore important to note that XSS attacks are in three types and these are Reflected, Stored and DOM XSS.

There are three forms of XSS, usually targeting users' browsers:

**Reflected XSS**: The application or API includes unvalidated and unescaped user input as part of HTML output. A successful attack can allow the attacker to execute arbitrary HTML and JavaScript in the victim's browser. Typically the user will need to interact with some malicious link that points to an attacker controlled page, such as malicious watering hole websites, advertisements, or similar.

**Stored XSS**: The application or API stores unsanitised user input that is viewed at a later time by another user or an administrator. Stored XSS is often considered a high or critical risk.

**DOM XSS:** JavaScript frameworks, single page applications, and APIs that dynamically include attacker controllable data to a page are vulnerable to DOM XSS. Ideally, the application would not send attacker controllable data to unsafe JavaScript APIs. Typical XSS attacks include session stealing, account takeover, MFA bypass, DOM node replacement or defacement (such as trojan login panels), attacks against the user's browser such as malicious software downloads, key logging, and other client side attacks

Protection against XSS attacks can be achieved by mixing multiple languages such as HTTML and JavaScript together as this an architecture of storing data and codes in the same memory (Prokhorenko et al., (2015). Deepa .G. and Thilagam P. S., (2016) also says XSS vulnerabilities are possible to eliminate when secure coding practices such as sanitization of untrusted input for removing harmful properties are adopted. This is possible because XSS attack prevention approach help to identify and prevent malicious scripts from execution by the user. Or better still, you can employ the following:

- Using frameworks that automatically escape XSS by design, such as the latest Ruby on Rails, React JS. Learn the limitations of each framework's XSS protection and appropriately handle the use cases which are not covered.
- Escaping untrusted HTTP request data based on the context in the HTML output (body, attribute, JavaScript, CSS, or URL) will resolve Reflected and Stored XSS vulnerabilities. The OWASP Cheat Sheet 'XSS Prevention' has details on the required data escaping techniques.
- Applying context sensitive encoding when modifying the browser document on the client side acts against DOM XSS. When this cannot be avoided, similar context sensitive escaping techniques can be applied to browser APIs as described in the OWASP Cheat Sheet 'DOM based XSS Prevention'.
- Enabling a Content Security Policy (CSP) is a defense in depth mitigating control against XSS. It is effective if no other vulnerabilities exist that would allow placing malicious code via local file includes (e.g. path traversal overwrites or r vulnerable libraries from permitted content delivery networks)

XML injection is a vulnerability that substitute malformed input in place of Xpath queries, OS commands and LDAP statements respectively as a way of distorting application behavior (Deepa .G. and Thilagam .P.S., 2016). Website attacker manipulates the value of HTTP header due to HTTP response splitting. Hence response stream is interpreted by the attacker as two responses instead of one.

Business Logic Vulnerabilities (BLVs) are easily exploited because they allow the attackers to manipulate the business logic of a web application. The type of attacks that exploit BLVs are legitimate application transactions that perform undesirable operation that is not even part of normal practice of business.

Parameter manipulation is a type of vulnerability that allows an attacker to compromise web application behaviour. Deepa .G. and Thilagam .P.S., (2016) says, “attacks are caused due to violation of the semantic restrictions of the user-input and that input could be provided through user interface or manipulated in the HTTP request and cookies”. If the business logic is implemented incorrectly or is absent, then attacks can occur.

For the web application to be secure from injection vulnerabilities there is need to take care of security issues of the application at each phase of the software development life cycle as well as providing a second layer of protection after deploying the application.

## 3. CONCLUSION

All in all, we have comprehensively evaluated the web technologies that include hypertext, Hypermedia, markup languages and XML as regards to web design and development. This has been heavily influenced by the introduction of the web-technology as a platform for many different types of systems. This means major changes in the approaches, institutional principles, methodologies, tools among others that we use for web-application development as this paper centered on web technologies in terms of open issues and challenges. Besides, it calls for continuous research zero in on web improvements.

## 4. ACKNOWLEDGMENT

This study was supported by Prof. Kunda, D, Department of Computer Science and ICT, School of Science, Engineering and Technology, Mulungushi University.

## 5. REFERENCES

Anderson, P. (2007) “What is Web 2.0? Ideas, technologies

and implications for education. JISC Technology and Standards Watch". http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf accessed 11 December, 2006.

Barakovic, S and Skoun-Kapov. 2017. Modelling the relationship between design/performance factors and perceptual features contributing to Quality of Experience for mobile web browsing. Zagreb, Croatia. Elsevier. Computer in Human Behaviour 74(2017)311-329.

Baturay, M and M, Birtane. 2013. Responsive web design: a new type of design for web-based instructional content. TASET. Ankara, Turkey. Doi:10.1016/;j.sbspro.2013.12.259.

Bugliesi, M. S, Calzavara and R, Focardi. 2016. Formal methods for web security. Venezia, Italy. Elsevier. ScienceDirect. 2352-2208.

Cavus, N. 2016. Development of an intelligent mobile application for teaching English pronunciation. Mersin, Turkey. Elsevier. Doi:10.1016/;procs.2016.09.413.

Deepa, G and P, Thilagam. 2016. Security web applications from injection and logic vulnerabilities: Approaches and challenges. Surathkal, India. Elsevier. ScienceDirect 0950-5849.

El-Hajj, W, G Ben Brahim, H Hajj and H Safa. 2016. Security-by-construction in web applications development via database annotations. Lebanon, Saudi Arabia. Elsevier. ScienceDirect. 0167-4048.

Hannes Vilhjamsson et al (2007) The Behavior Markup Language: Recent Developments and Challenges. C. Pelachaud et al. (Eds.): "Intelligent Virtual Agents 2007", Lecture Notes in Artificial Intelligence 4722: 99-111, Springer-Verlag Berlin Heidelberg

Huang, Y. C, Tsai. I. Lin, S, Huang, D.T. Lee and S, Kuo. 2015. A testing framework for web application security assessment. Taipei, Taiwan. Elsevier. Doi:10-1016/;.comnet.2005.01.003.

Jankowski, J and S, Decker. 2013. On the design of Dual-Mode User Interface for accessing 3D content on the World Wide Web. Galway, Ireland. Elsevire. Int. J. Human-Computer Studies 71(2013) 838-857.

Nguyen-Tuong, A. S. Guarnieri, D. Greene and D, Evans. 2004. Automatically Hardening Web Applications Using Precise Tainting. Virginia. Technical Report CS-2004-36.

Palmer, J. W. 2002. Website Usability, Design and Performance Metrics: Information Systems Research. Vol. 13, No. 2, Measuring e-commerce in Net-Enabled Organisation (part 1 of 2), pp.151-167.

Prokhorenk, V. K, Raymond Choo. H, Ashman. 2015. Web application protection techniques: A taxonomy. Austrilia. Elsevier. ScienceDirect 1084-8045.

Rodrigues, L. F. C, Costa, and A. Sliveira. 2017. How does the web game design influence the behavior of e-banking users? Lisboa, Portugal. Elsevier. Computers in Human Brhaviour 74(2017) 163-174.

Saputra, D G and F N Azizah. 2013. A Metadata Approach for Building Web Application User Interface. Bandung, Indonesia. Elsevier. Doi:10.1016/;portly.2013.12.274.

Tim Berners (2001) A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities: May 17, 2001 The Semantic Web

Webster J and J, Ahuja. 2016. Enhancing the Design of Web Navigation System: The Influence of User Disorientation on Engagement and Performance. MIS Quarterly, vol. 30, No.3, pp.661-678.

# Electronic Publishing on the Web: Challenges and Open issues

Mark MuKuba Mwale

*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
markmukubamwale@gmail.com

Douglas Kunda

*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
dkunda@mu.edu.zm

***Abstract:***
**The web as an internet based hypertext system is a good platform for disseminating and sharing information. Electronic books, magazines, newspapers and the development of digital libraries are all examples of electronic publications available on the web. Over the years, electronic publishing on the web has undergone a lot of innovations thereby becoming too complex. These include; editing electronic books, magazines, journals and other documents meant for public consumption on particular areas of concern or interest. In view of these developments, the paper discusses electronic publishing on the web with special interest in challenges and issues.**

***Key words: Publishing, Paper publishing, Electronic Publishing on web, Consumer, web***

## I. INTRODUCTION

Since the Second World War, the world has seen unprecedented technological evolution likened to the industrial revolution of the early 19th Century. Computer science is the engine behind these technological advancements which have given birth to multiple use of the computer. One of the most significant transformations or innovations brought by computer technology is the utilization and exploitation of information.

Because of improvements in the use and application of computer literacy, it is increasingly becoming absurd to cling to cultural lag bound by material culture contrary to modern times when all sectors of global economic growth and development are influenced by web technology[1]. Colleges and universities are also at the point of departure from archaic paper work to electronic information systems. The web is at the helm of all this and key to driving the global wheels of economic growth and development.

As observed by some scholars, the web and computer technology in particular will one day be as important as literacy and arithmetic. Due to this significant transformation, the world computer access percentage is increasing exponentially**.** For example, from 16% in the year 2005 to 47% in 2016 and an internet penetration rate of 54.4 % by 2017[23].

This simply explains why electronic publishing on the web is now the only plausible way to reach out to the masses. Thanks to innovations and improvements in computer technology, access to information through Electronic publishing on the web, e-Journals, e-books and e-newspapers as they are sometimes called is no longer a problem. As opposed to limiting information to paper edited books, newspapers and journals, the increase in internet penetration rate is a solution to functional illiteracy and expansion of scientific knowledge.

The web is a haven for transformed processes of electronic publishing thereby turning out to be the cradle for creating and developing of a new society of the computer age in which access to the right information at the right time is no longer a problem. Information from e-Journals or books is just a click of a computer's button away. The e-library is readily available on the web and can be accessed at any time provided one is in an area where there is internet. It is convenient, faster and effective.

In 1993, online publishing appeared like the doorway into a new world of instant global communication although today it appears that the hypertext links in early web publications served as holes in a sieve through which it has leaked. Therefore in the last two decades, e- publishing has involved greatly [2]. The paper is therefore poised to discuss electronic publishing on the web, challenges and issues.

## II. WHAT IS MEANT BY ELECTRONIC PUBLISHING ON THE WEB?

Electronic publishing on the web is also called digital publishing or e-publishing. To understand the meaning of the term 'electronic publishing on the web' it is imperative that we look at each of the key concepts in the statement, that is, electronic, publishing and web.

### A. *Electronic*

The word electronic is an adjective of electrical. It is a term that is used to refer to a device or system of operations triggered by electric current or a flow of electrons. All electronic devices use and control electrical current. An example of an electronic device that does all these operations is a computer. An electronic catalogue, database, data and records are examples of electronic information dealt by the computer network system.

Quite often we use the phone and other forms of computer to send electronic messages, upload songs, e-books, and newspapers among other things. In other words, electronic implies the movement, conversion and management of electronic information or data for consumption or use by an individual in the correct form, for example, in picture, video, word or voice message.

### B. *Publishing*

Publishing is the broadcasting or distribution of written information, audio and video content to the general public. The author needs a media or someone to sell his or her ideas to the general public. The media or individual marketing the publications is called a publisher. However, it is possible for a person to play both roles of being an author and a publisher. Publishing is a means to making information reach the intended audience or the general public.

The classical or orthodox (traditional) process of publishing involving paper is made up of seven main stages. These are acquisition, copy editing, production, printing, marketing and distributing [3].

### C. *Electronic publishing*

It incorporates all the above stages in one package taking advantage of the internet whose world penetration rate as already alluded to is at 54.4%, meaning, that the web acts as media or platform for all the processes involved in publishing to take place within a short period of time.
The latest web version is 2.0 which literally refers to the second generation of the Web, wherein interoperable, user-centered web applications and services promote social connectedness, media and information sharing, user-created content, and collaboration among individuals and organizations [4].

### D. *Electronic Publishing on the web*

Electronic Publishing on the web refer to the acquisition, editing, production, printing, marketing and distributing of journals books, magazines and other forms of data using the world internet or web as media [5].
Electronic publishing on the web is faster, ideal and reaches out to the masses because of the perceived increase in the number of internet users in the world from the year 2000 to 2018 as confirmed by the table below.

*Table 1*

**WORLD INTERNET USAGE AND POPULATION STATISTICS**
**JUNE 30, 2018 - Update**

| World Regions | Population ( 2018 Est.) | Population % of World | Internet Users 30 June 2018 | Penetration Rate (% Pop.) | Growth 2000-2018 | Internet Users % |
|---|---|---|---|---|---|---|
| Africa | 1,287,914,329 | 16.9 % | 464,923,169 | 36.1 % | 10,199 % | 11.0 % |
| Asia | 4,207,588,157 | 55.1 % | 2,062,197,366 | 49.0 % | 1,704 % | 49.0 % |
| Europe | 827,650,849 | 10.8 % | 705,064,923 | 85.2 % | 570 % | 16.8 % |
| Latin America / Caribbean | 652,047,996 | 8.5 % | 438,248,446 | 67.2 % | 2,325 % | 10.4 % |
| Middle East | 254,438,981 | 3.3 % | 164,037,259 | 64.5 % | 4,894 % | 3.9 % |
| North America | 363,844,662 | 4.8 % | 345,660,847 | 95.0 % | 219 % | 8.2 % |
| Oceania / Australia | 41,273,454 | 0.6 % | 28,439,277 | 68.9 % | 273 % | 0.7 % |
| WORLD TOTAL | 7,634,758,428 | 100.0 % | 4,208,571,287 | 55.1 % | 1,066 % | 100.0 % |

*Source: Internet World Stars- ww.internetworldstats.com/stats.htm*

## III. ETHICS GUIDING ELECTRONIC PUBLISHING ON THE WEB

The achievement of a singular, open, shared communications platform has huge constitutional implications for publishing and its role in society . Topics under review or discussion in electronic publications do not come from thin air. They come from a prescribed social, religious and economic context. It is the same context which turns out to be the audience and consumer of the publication. For this apparent reason, there is need for both the author and the publisher to adhere to scientific procedures while at the same time consider societal values on a particular subject of interest in electronic publishing on the web. This is called ethics.
As a branch of moral philosophy, ethics is concerned with right and wrong conduct[6].
Like any other scientific research and writing, electronic publishing on the web is expected to be guided by ethics. The autonomous and accessibility of online publishing does not give the author and publisher immunity from taking full

responsibility of heeding to ethical considerations, especially that electronic publishing on the web targets the entire world, meaning that any person with an electronic gadget such as a mobile phone, tablet or computer has easy access to the document. The only driving force is availability of internet. It is thus imperative for the author and publisher to understand that there is open access to all online publications. Open access provides unrestricted access via the web to publication outputs [5].

### **A.** *Censoring by the author and publisher*

The combination of open access online publishing with the demand for increased publication rates from academics has created the opportunity for predatory and counterfeit publishing to exist within the sector [5]. As a measure to this effect, ethics are critical in electronic publishing on the web because they compel the author and publisher to censoring their data before releasing them on the web, taking into account web's nature of open access, a free license for all to read and subject the publication to criticism.

**B.** *Posting on the web and plagiarism*

The author (s) is/are the most important player (s) in electronic publishing on the web. As soon as the author approves the final revisions, the manuscript is posted online, and the digital object identifier (DOI) information is provided with the online publication date for other articles to quote and cite . It is therefore ethical for the author (s) to ensure their work is unique and original. Plagiarism must be avoided at all cost bearing in mind that it is a serious academic crime that jeopardize the essence and values of scientific research.

Moreover, all scientific works are subjected to empiricism, skepticism and determinism. Based on these elements of scientific approaches, especially skepticism, no data is said to be absolute truth. It is subject to critique and analysis, thereby, making plagiarism unethical and something that should not be entertained. However, it should be borne in mind that authors and publishers are at greater risk of falling prey to temptations of plagiarism because of the availability of related or the same information under discussion on the web.

**C.** *Engagement of all parties involved in the publication*

It is ethical for the author (s) and publisher (s) not to take pecuniary advantage of other concerned parties in the publication. This means that for the document to be published on the web, all parties should be in agreement[7]. The significance of such an undertaking brings about collective ownership of mistakes, praises and success the publication cause on the web.

In as much as online publications are said to have an open access, their authority and authentication is a preserve of the author (s) and publisher (s).Non-affiliated researchers are able to access less than half of the peer-reviewed literature .

**D.** *Honesty and impartiality*

It is ethical to consider that the most important persons in the publishing chain are consumers. The consumers of electronic publishing on the web include all people in the world. Bearing this in mind, the quality of work must be thoroughly scrutinized and all grey areas with the potential of causing harm or injury on consumers owing to us living in a pluralistic world amid pervasion of human rights must be put into consideration and set as a priority. This tie the author (s) and publisher (s) to tenets depicting high levels of honesty. The understanding here is that all those involved in electronic publishing on the web should be factual, precise and without exaggerations, a phenomenon common to Eurocentric, Asiacentric and Afrocentric writers [8]. An African writer portraying the exaggerated revolutionary or radical condemnation of western technology considering it as cultural imperialism on the African soil, when in the real sense was supposed to be factual and acknowledge the social, economic and political developments it has brought to the continent is an example of an Afrocentric writer or scholar. The opposite may be a Eurocentric or Asiacentric writer. Such authors and publishers are against the ethical consideration of not causing harm to the audience via electronic publications on the web.

**E.** *Integrity*

In addition to honesty on electronic publishing on the web, integrity is ethical. Integrity refers to reliability and dependability. The expectation is that data should be precise with little or no errors at all, and subsequently give the same results adjudged on the ground. Reliability of data published electronically must therefore undergo peer review before it is brought to the public through the web [9]

**F.** *Confidentiality*

Open access is also very important, as it allows anyone who wishes to share an article to copy the link and e-mail it to someone else who, with one click, may view the full text of the article [10]. However, it seems absurd to consider the need for privacy and confidentiality in electronic publications on the web and somehow antagonizing the principle of open access. Just as the case is for none electronic books, journals and newspapers, electronic publishing is bound by privacy and confidentiality to guarantee security. This is a responsibility of the electronic editor.

**G.** *Justice*

Electronically published material is supposed to be guided by the ethical principle of justice. Justice literally refers to treating the equals equally and the unequal, unequally. It works on the principle of fairness. When referred to electronic publishing on the web, it is synonymous to access. Both the author and the publisher should know from the beginning that access to information is a right that must not be subjected to discrimination. For example, if the information is only electronically published, how about those who do not have a computer? The same can be said to the impaired that entirely depend on assistive technology.

If they are also a target population or universe for the information in question the ethical principle of justice demands that journal or information published electronically is also printed. Dual publishing thus works on the principle of not leaving anyone behind. It is inclusive and accessible to all including those limited to assistive technologies like the braille.

Other issues that need to be tackled regarding access or justice have to do with whether the e-publication is free or not, meant for a specific group of people or the general public. In instances when the publication is only accessible at a fee, the amount must be reasonable, meaning that the clients or would be buyers of the e-publication be it a journal or book should not be exploited.

## IV. REASONS WHY ELECTRONIC PUBLISHING ON THE WEB IS MORE VIABLE OVER PRINT MEDIA/PUBLICATIONS.

A. *Cost Effective*

Electronic publishing on the web is cost effective. Costs involved in the process of publishing are relatively low compared to print publishing. No need for establishing a publishing house with expensive machinery and many employees to do the different tasks. The huge costs associated to printing drafts for editing are reduced. A reliable and efficient internet is the only viable tool needed to have all the processes done. The raw data are posted to the electronic editor's e-mail box, corrected and sent back before the two can agree to post it to the web. The whole process is relatively cheaper. The author looks at the corrections made or highlighted by the editor and act accordingly.

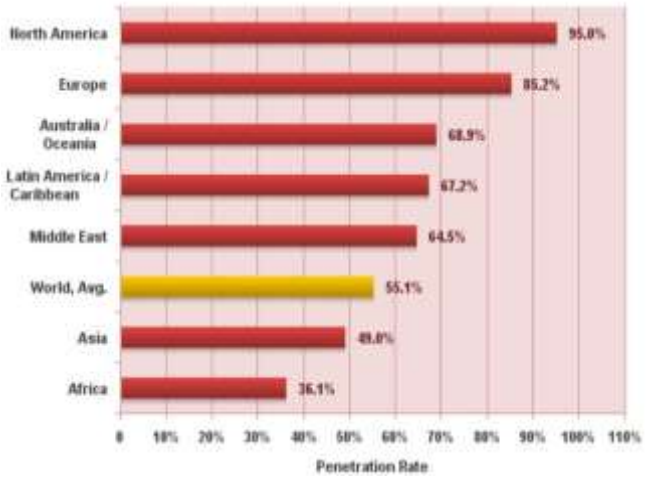
B. *Faster than printed media/publications*

Electronic publication on the web escalates the end to end process of the information or electronic material. Contrary to printed publication which takes a lot of time to reach the intended target due to a number of processes involved such as printing, binding and transportation, the internet and working computers in good condition are enough to have the entire processes carried out within a short period of time.

At the click of a button, the e-journal or book is posted to the web for the intended audience which in turn has access to it almost immediately. A delivery of the publication that would have taken days or months to reach the targeted people or population with numerous procedures as the case is for paper publications is carried out within a second. Moreover, the web penetration is so pervasive due to the expansionist policy computer and web technology, the globe has adopted in the recent past. In this regard, the web penetration at the moment is as per illustration in figure 1

*Fig 1*

**Internet World Penetration Rates by Geographic Region – June 30, 2018**




Source: Internet World Starts- www.internetworldstats.com/stats.htm

C. *Immediate feedback*

The author receives feedback on his publication almost immediately. He is able to assess the contribution the publication has done to the international body of knowledge. This gives room for self-evaluation, improvements and learning within a short period of time. The Electronic publication is open for critics on the international platform.

It is perfected by the specialists in the field or subject matter on the global platform, giving the author (s) an opportunity to learn from others. In other words, Electronic Publishing on the web is an open field for corrections and perfecting one's writing skills.

D. *Integration*

Readable electronic publications on web can easily be integrated with other forms of communication like audio recordings, videos and simulations.

E. *Easy access*

Searching for electronic publication is easier and faster. The reader just need to search for the author's name, title of the publication or the key concept to get to the material. Unlike paper publications and traditional libraries where one may take the whole day looking for the concept, book or journal on a specific topic/subject, the researcher only takes seconds to access the information in question.

Both the author and the reader have easy access to other electronic publications related to the subject or topic under research. This is good because other publications act as references.

F. *Encourages competition, innovativeness and discovery*

References and quotations and possible plagiarism are easily noticed. The work must therefore be original and free from plagiarism a thing that encourages high levels of competition, innovativeness and discovery. The author claims ownership of the information contained in the publication.

G. *Easy storage and durable*

All electronically published journals, books, newspapers and other forms of documents are achieved in the international data base. Durability is therefore guaranteed. It is available on the web and not subject to tear and wear as the case may be for printed publications. There is guarantee that they will be in their original form with all the pages intact for many years to come.

Furthermore, the ArchiMed, the electronic publication server, in which the document is archived, makes it readily available for research and reading in all other search engines on the web. Digital publishing provide significant opportunity for academic libraries that had not been heavily involved in publishing activities previously and has led to a new purpose for library publishing [11]. To this effect, there are no cases of the publication missing or not being there because it is on the international electronic library. The electronic library, which is home to e-publications, is not static or stationed at a particular place. Search engines on a computer and operating systems are the main gates into the e-library.

## V. CHALLENGES/ISSUES RELATED TO ELECTRONIC PUBLISHING ON THE WEB

### A. *Not easy to market*

Marketing of the publication entirely rests on the author. He has to do everything possible to ensure the publication is marketed on the web especially with the coming of search engine optimization (SEO). The challenge is that there are so many good writers who may condemn the publication such that it becomes very difficult to sell.

The method does not allow for storefront window and purchase. They cannot be put on the shelves like paper publications. Over reliance on the web therefore makes the marketing of e-publications complicated.

### B. *Privacy of the Author and readers not guaranteed*

The privacy of the author and readers remains much to be desired using this method of publishing [12]. It is at the mercy of the electronic editor and may lead to chaos if he lacks understanding of ethical principles governing e-publishing on the web. Copyright and other patents rights are threatened making the author vulnerable to plagiarism and other forms of academic crimes without consent.

### C. *Usually disadvantages the author*

The low cost of publishing is more of a phenomenon than reality. This perceived low cost of publishing favours the publisher while it is very high on the part of the author [13]. The author needs a lot of money to gather data and internet cost for a publication that may be offered almost at no cost to the consumers (readers).

In some instance, the e-book or journal is sold out at a relatively higher price much to the benefit of the publisher. As the adage goes 'the sweetest tree receives the maximum number of stones', a book that is so attractive will fetch a higher price. The challenge however is the extent to which the publication meets the criteria to attract a large following of readers and researchers for it to have a relatively good price, especially to new authors on the international market.

### D. *Vulnerablity to wanton criticism*

It is argued that a bad carpenter always condemns his tools. However, it is also true that even when you are to subject a player with shackled legs to being trained by the world's best coach like Sir Alex Furgerson, he would still remain poor. This applies to writers limited in scope and language braving themselves to posting their works on the web. They receive the worst form of abuse or wanton condemnation thereby being discouraged to make further attempts [14]. Unfortunately, the worst victims of these forms of abuse are beginners and writers from least developed countries, especially on scientific research involving simulations and theories of economic growth and development. This perception is to a greater extent the reason why there are very few African scholars taking the challenge to participate in electronic publications on the web and is said to be generating at less than 1 % of the World's research. Even then, most typical African Authors are of European descent.

### E. *Preference for printed publications or media*

The sales are not as greater as paper publications because of cultural-lag. Despite the international e-library being flooded with books, journals and other documents for academic purposes, scholars are attracted to paper publications [15]. On the contrary, Until the question of sustaining support can be answered, traditional publishers will continue to insist that the new models of sharing knowledge such as that of electronic publishing on the web are flaws.

Every discussion on publishing models will certainly lead to the question of cost bearing, "Who Pays?" No matter the medium, licensing, interactive vs static content, there are costs. There is no publishing which is free. Web publication needs resources to sustain an infrastructure of equipment, basic services and software to ensure that intellectual resources are available for the future.

## VI. CONCLUSION

The world is now a global village and has set for itself goals enshrined in the Sustainable Development Goals (SDGs). All these Seven SDGs can only be attained through effective implementation of e-governance systems, especially in the area of scientific research and innovations. It is in this respect that the globe is advocating for a paperless society in which e-publishing on the web takes centre stage. Unless much emphasis is made on electronic publishing on the web, attainment of the SDGs remains a pipe dream.

REFERENCE

[2] L. Fillmore, "Twenty Years Into IT: Online Publishing: Threat or Menace? Revisited," *J. Electron. Publ.*, vol. 18, no. 4, Dec. 2015.

[3] F. Dodds, "Changes in the role of the commissioning editor in academic book publishing," *Learn. Publ.*, vol. 28, no. 1, pp. 35–42, Jan. 2015.

[4] D. Wilson, X. Lin, P. Longstreet, and S. Sarker, *Web 2.0: A Definition, Literature Review, and Directions for Future Research.* 2011.

[5] K. McNaught, "The Changing Publication Practices in Academia: Inherent Uses and Issues in Open Access and Online Publishing and the Rise of Fraudulent Publications," *J. Electron. Publ.*, vol. 18, no. 3, Jun. 2015.

[6] M. Ananny, "Toward an Ethics of Algorithms: Convening, Observation, Probability, and Timeliness," *Sci. Technol. Hum. Values*, vol. 41, no. 1, pp. 93–117, Jan. 2016.

[7] M. Price, K. Handley, and J. Millar, "Feedback: focusing attention on engagement," *Stud. High. Educ.*, vol. 36, no. 8, pp. 879–896, Dec. 2011.

[8] R. D'Andrea and J. P. O'Dwyer, "Can editors save peer review from peer reviewers?," *PLOS ONE*, vol. 12, no. 10, p. e0186111, Oct. 2017.

[9] I. F. A. Shaikhli, A. M. Zeki, R. H. Makarim, and A.-S. K. Pathan, "Protection of Integrity and Ownership of PDF

Documents Using Invisible Signature," in *2012 UKSim 14th International Conference on Computer Modelling and Simulation*, Cambridge, United Kingdom, 2012, pp. 533–537.

[10] G. Coulter, "Launching (and Sustaining) a Scholarly Journal on the Internet: The International Journal of Baudrillard Studies," *J. Electron. Publ.*, vol. 13, no. 1, Mar. 2010.

[11] K. M. Conrad, "Public Libraries as Publishers: Critical Opportunity," *J. Electron. Publ.*, vol. 20, no. 1, May 2017.

[12] S. Kirrane, S. Villata, and M. d'Aquin, "Privacy, security and policies: A review of problems and solutions with semantic web technologies," *Semantic Web*, vol. 9, no. 2, pp. 153–161, Jan. 2018.

[13] F. Bowes, "An overview of content accessibility issues experienced by educational publishers: A current overview of content accessibility," *Learn. Publ.*, vol. 31, no. 1, pp. 35–38, Jan. 2018.

[14] D. Robinson, "A Justification of Command Responsibility," *Crim. Law Forum*, vol. 28, no. 4, pp. 633–668, Dec. 2017.

[15] M. Dragoni, S. Tonelli, and G. Moretti, "A Knowledge Management Architecture for Digital Cultural Heritage," *J. Comput. Cult. Herit.*, vol. 10, no. 3, pp. 1–18, Jul. 2017.

[16] L. Fillmore, "Internet Publishing: How We Must Think," *J. Electron. Publ.*, vol. 1, no. 1&2, Feb. 1995.

[17] L. Fillmore, "Internet Publishing in a Borderless Environment: Bookworms into Butterflies," *J. Electron. Publ.*, vol. 1, no. 1&2, Feb. 1995.

[18] A. Marchant *et al.*, "A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown," *PLOS ONE*, vol. 12, no. 8, p. e0181722, Aug. 2017.

[19] A. Luke, *Publishing e-books for dummies*. Hoboken, N.J.; Chichester: Wiley ; John Wiley [distributor, 2012.

[20] J. W. Maxwell, "Publishing Education in the 21st Century and the Role of the University," *J. Electron. Publ.*, vol. 17, no. 2, May 2014.

[21] J. W. Maxwell and K. Fraser, "Traversing The Book of Mpub: an Agile, Web-first Publishing Model," *J. Electron. Publ.*, vol. 13, no. 3, Dec. 2010.

[22] T. Lieb, "Q.A.: Basic Journal-ism: Tips for Electronic Publishers," *J. Electron. Publ.*, vol. 3, no. 1, Sep. 1997.

[23] D. Krairit, "The New Face of Internet User Typology: The Case of Thailand," *J. Theor. Appl. Electron. Commer. Res.*, vol. 13, no. 2, pp. 58–79, May 2018.

[24] S. Kim, E. Chung, and J. Y. Lee, "Latest trends in innovative global scholarly journal publication and distribution platforms," *Sci Ed*, vol. 5, no. 2, pp. 100–112, Aug. 2018.

[25] D. Greenstein, "Next-Generation University Publishing: A Perspective from California," *J. Electron. Publ.*, vol. 13, no. 2, Nov. 2010.

[26] L. McKnight *et al.*, "Information Security for Electronic Commerce on the Internet: The Need for a New Policy and New Research," *J. Electron. Publ.*, vol. 1, no. 1&2, Feb. 1995.

[27] M. Milutinovic, "Making a Library a Digital One," *J. Electron. Publ.*, vol. 18, no. 1, Jan. 2015.

[28] D. O'Donnell *et al.*, "Aligning Open Access Publication with Research and Teaching Missions of the Public University: The Case of The Lethbridge Journal Incubator (If 'if's and 'and's were pots and pans)," *J. Electron. Publ.*, vol. 18, no. 3, Jun. 2015.

[29] J. Owuor *et al.*, "Does assistive technology contribute to social inclusion for people with intellectual disability? A systematic review protocol," *BMJ Open*, vol. 8, no. 2, p. e017533, Feb. 2018.

[30] W. Rodgers and S. Negash, "The effects of web-based technologies on knowledge transfer," *Commun. ACM*, vol. 50, no. 7, pp. 117–122, Jul. 2007.

[31] D. G. Tracy, "The Users of Library Publishing Services: Readers and Access Beyond Open," *J. Electron. Publ.*, vol. 18, no. 3, Jun. 2015.

[32] B. White, "Total availability of journal articles to Internet users," *Libr. Rev.*, vol. 63, no. 4/5, pp. 295–304, Jul. 2014.

# Benefits and Challenges in the use of Cloud Computing in Colleges of Education in Zambia

Phyllis Siyomunji
*Department of Information and Communication Technology*
*David Livingstone C*ollage of Education
*Livingstone, Zambia*
psiyomunji@gmail.com

Douglas Kunda
*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
dkunda@mu.edu.zm

***Abstract*** **- Cloud computing (CC) is a computing technology that allows or enables access to a pool of shared configurable resources. It involves computer networks, servers, storage, applications and services and these can be quickly provided over the internet without much effort from management. There are several cloud service models that are used and these include Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Mostly these models offer increasing abstraction and they are often seen as layers in a stack. The paper discussed the benefits the education sector and its clientele stood to derive from the usage of cloud computing applications. Benefits such as automated assessments and examinations ,lower long-term costs, instant feedback to students ,creation of digital records of student growth and development, greater storage efficiency increased productivity and low operational variability where discussed. The paper discussed the various challenges that come with cloud computing in education. Some of these challenges include security, data privacy, as well as insufficient network. However, it was concluded and recommended that cloud computing technology as a tool can be used and it should be rapidly expanded in public Colleges of Education in Zambia.**

***Keywords:*** *Cloud Computing, Cloud Computing Models, Education, Information Communication Technology.*

## I. INTRODUCTION

According to [1], Information and Communication Technologies (ICTs) are a diverse set of technological tools and resource used to create, store, manage and disseminate information. Additionally [2] posits that the term ICTs as applied to education, are those technologies that include computers, the internet, broadcasting technologies (radio and television), and telephone that can facilitate not only delivery of instruction, but also learning process itself. These technologies have been identified as important tool for realizing a new paradigm of learner-centred education that better support learner' needs through differentiated and personalized instruction. [3] contends that the issue of 'computers in education started to become popular in educational policy- making in the early 1980s, when relatively cheap microcomputers became available for the consumer market. Kofi Anan, the former United Nations Secretary General, in the [4] pointed out that we must ensure that ICTs unlock the door of education system. One area that ICTs can be used is cloud computing (CC) in education.

## 2. CLOUD COMPUTING

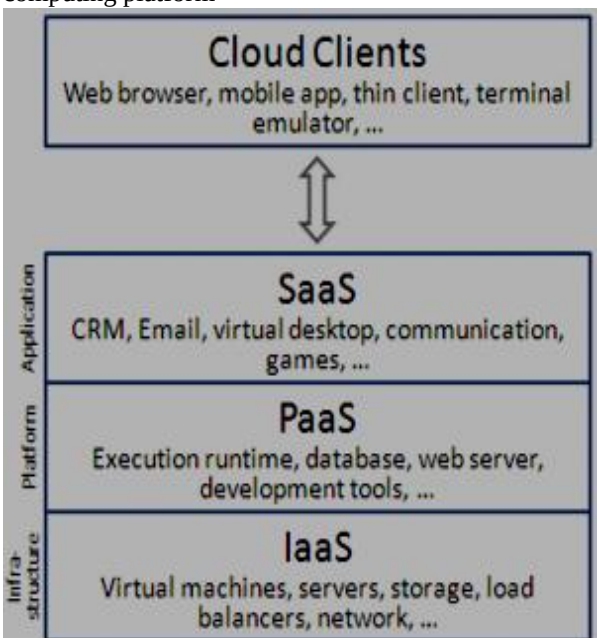
[5] points out that The National Institute of Standards and Technology (NIST) defines cloud computing as a model that enable ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal effort from management or service provider interaction . Additionally, [5] reports that the US National Institute of Standards and Technology's definition of cloud computing identifies "five essential characteristics":

Firstly is the on-demand self-service where users can obtain computing services, such as server time and network storage, as needed automatically without interacting with individual provider of services. Then there is also broad network access. This is where capabilities are available over networks and can be accessed via mobile phones, tablets, laptops, and workstations. Resource pooling is another characteristics of cloud computing. On this one, the services provider's resources are put together to serve multiple users using a multi-tenant model assigned according to user demand. Furthermore is rapid elasticity characteristic. This is where capabilities can be elastically provided and released mostly automatically outward and inward in line with user demands. To the user, the capabilities provided often appear unlimited and they seem to be in any quantity and any time. Lastly is the measured service characteristic. On this one, the cloud computing system on its own control and optimize the usage of resources. In this vein, the use of resources usage can be controlled, monitored and reported in a transparent manner to the service provider and the user of services. For cloud computing to work well, there are models that have been developed.

## 3. CLOUD COMPUTING MODELS

It should be noted that there are several cloud service models that are used. According to the NIST as reported by [5], only three standard models are mostly used. These include Software as a Service (SaaS, Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). In this vein, [5] submits that The SaaS model in cloud computing entails cloud providers offering the application to the end user on or through a cloud infrastructure. Consequently, the user does not manage or control the underlying cloud infrastructure. Subsequently, cloud applications using the SaaS model could be accessed through a program interface in the way that Dropbox works or directly through a client interface as is the case with web-based email like Gmail. On the other hand, the PaaS model of cloud computing gives an opportunity for consumer a platform

to deploy applications into the wider existing cloud infrastructure. In this regard, PaaS allows the end user a space to build their own application using shared languages, libraries and services that are either supplied or supported by the cloud platform provider. Lastly, the IaaS cloud-computing model provides the processing, storage, networks, and other fundamental computing resources that give opportunities for user to deploy and run the software of their choice. This may involve programs and applications and this could even be extended to running operating systems via the cloud infrastructure service. Figure 1 below is an example of cloud computing platform




Source: http://www.smart draw.com/network diagram

**Figure 1: Cloud computing service models arranged as layers in a stack**

Generally, [6], outlined benefits and challenges associated with the usage of cloud computing. The following table 1illustrates these benefits and challenges;

**Table 1: Benefits and challenges of cloud computing**

| BENEFITS | CHALLENGES |
|---|---|
| Access to applications from anywhere | Speed /lack of internet can affect work |
| Software free or pay-per-use | Standards adherence |
| Protection of environment by using green technologies | Risks related to data protection and security and accounts management |
| Increased access of students to new technologies | Lack of confidence |
| Increasing functional capabilities | Maturity solutions |
| Open to business environment and advanced research | Security and protection of sensitive data |
| 24 hour access to infrastructure and content | Dissemination politics ,intellectual property |
| Off line usage with further synchronization | Organisational support |
| Support for teaching and learning | Not all applications runs in the cloud |

Cloud Computing applications can be used can be used in any sector. One of the sectors is education. In this regard, educational institutions such as College of Education in Zambia can adopt cloud computing models in the management of their data to derive such benefits. In Zambia, there are ten (10) public Colleges of Education in total. Generally, most of these public Colleges of education in Zambia, have the following departments; Languages, mathematics, Natural Sciences, Early Childhood education, Social Sciences and Expressive Arts. The subjects' students are being taught are; Religious education, Guidance and counselling, Geography, Zambian Languages, Educational psychology, Business studies, Agriculture science, French, English, Civic education, Mathematics, Special education, History, Philosophy of Education, Theory and practice of Education, Sociology of Education, Physical Education, Art and Design, Communication Skills and Information & Communication Technologies. Mostly, the total number of Lecturers is estimated to be one hundred (100) with estimated the total number of students to be one thousand (1,000). In terms of computer facilities; colleges have the One hundred (100) computers and ten (10) laptops. Mostly, these computers are connected to the internet and they are being used for educational purposes. However, Colleges of Education in Zambia have not adopted cloud computing applications.

## 4. CLOUD COMPUTING AND EDUCATION

Cloud computing technology as a tool can be used in education institutions. It should be rapidly expanded so that it becomes a more integral part of the collegiate experience. This is important to do because cloud computing has the ability to collaboratively share, edit, process, and store huge amounts of data.

On the other hand [7], submits that the economies of scale is one of the main characteristics of cloud computing. This is so because cloud services could be provided at a lower cost when compared with other in-house mainframe computers, networks, and computer infrastructure that are normally provided by the educational institutions. Additionally,[8] submits that the application of cloud computing services in the educational sector provide more access to more resources resulting to increased quality of teaching and learning. This happens because cloud computing applications allow educators and students to have quick and easier connections to the core materials. Ultimately [9], argues that cloud computing can provide educational institutions benefits such as:

interactive teaching and learning; flexibility to create structured learning environments; ubiquitous availability of online applications; support for mobile learning; and scalability. In this regard [10], contends that the use of cloud computing is increasing becoming popular in the educational environment. Consequently [11], submits that as educational technology is being infused in education institutions, many educational institutions are turning to cloud-hosted learning management systems (LMSs) that connect student databases with learning content. This is happening because cloud computing is seen as viable option for numerical modeling and visualization. It also facilitates collaboration and data storage. Additionally, cloud computing is an affordable resource that enables fast processing of information, large data-storage capacity, and the sharing of resources. Ultimately, the benefits offered by cloud computing by far outweigh the challenges introduced by using this emerging technological resource. The usage of cloud computing applications in education is depicted in figure 2 below.

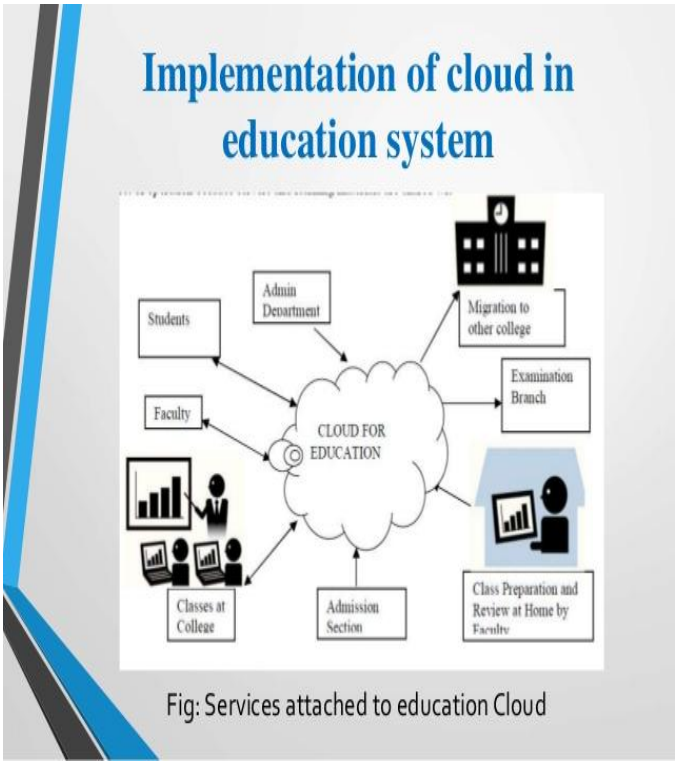



Fig. 2: Application of Cloud computing in education( Source : http://dentrodelasala.com.2018

## 5. BENEFITS OF CLOUD COMPUTING IN EDUCATION

The main benefit of Cloud computing in education is financial. [7] submits that using cloud computing, organizations do not have to make large investment in their own IT infrastructure like hardware and software. In this regard, Public Colleges of Education in Zambia will be able to reduce costs. This is possible because cloud-based applications can be run on Internet browsers. They are also compatible with mobile devices as well. This means that Public Colleges of Education administrators and students in Zambia do not necessarily need to own expensive computers. Furthermore, [12], contends that access to latest technologies and having maximum resources, while, at the same time minimizing costs are considered the main benefits of cloud computing. This fact is an added advantage to Colleges of Education Zambia who have limited financial resources.

In these and other ways, cloud computing is not only reducing costs, but also creating an environment where all lecturers and students can have access to high-quality education and resources. Currently, this is not the case within the Public Colleges of Education in Zambia. In this regard, [13] proposed new e-learning framework based on private cloud and virtual private network. The proposed framework was intended to help students in the university environment to access e-learning environment for resource sharing with less cost. It was noted that the framework was scalable and increased availability of resources and it was reliable. This benefit is critical for Colleges of Education in Zambia as there are little education resources for the usage of Lecturers and students within these colleges.

The other benefits of cloud computing as elaborated by [14] include increased flexibility, access anywhere, elastic scalability, pay-as-you-go and easy to implement as there is no need to purchase hardware, software licenses or implementation services. Similarly [15], posits that using cloud computing, the infrastructure is quickly available with flexibility and scalability of distributed testing environment. On the other hand, [16] submits that disaster recovery in cloud computing is another plus as cloud computing provides mechanisms for automated scheduled network wide backup systems in order to store data in off-site data centres. Mostly, Colleges of Education in Zambia have lost data because of lack of automated scheduled network wide backup arrangements. Furthermore,[1] argues that cloud computing enables all documents –projects, homework, syllabi, and collaborative exercises be updated in a centralized and systematic manner and to be modified consistently at a single point. This benefit is critical for Colleges of Education in Zambia as it will do way with duplication of work in terms of updating documents.

It should be noted that there are various studies that have been done on the benefits of using cloud computing in higher education institutions. In this vein, a study was done by [17], on the adoption of cloud computing among public universities in South Africa. The results revealed that public universities in South Africa share similar operational processes such as course offerings, admissions, enrolment, research and graduations in a more cost effective way. Currently, these processes within Colleges of Education in Zambia are done in a costly way. In this vein, adopting cloud computing applications will be cost effective. [18], carried out study in

Nigeria whose aim was to assess the impact and challenges of cloud computing adoption on public universities in South Western Nigeria. The results of the findings revealed that the adoption of cloud computing has a significant impact on cost effectiveness, enhanced availability of resources,increased operability and reduced investment in physical assets. It can thus be argued that cloud computing can beneficial to Colleges of Education in Zambia as the impacts established in the survey are not obtaining in the Colleges..

Another study was done by [19] investigated the development of an efficient cloud computing framework for the use by universities in Iraq. The framework proposed noted some characteristics such as low cost, flexibility, mobility, and business continuity as the benefits of cloud computing. However, some issues such as security, reliability, and loss of sensitive data were not dealt with adequately when proposing the framework. Furthermore, [20] studied on how cloud computing could benefit education institutions in South Africa. The findings revealed the cost saving on hardware and software and the flexibility of the cloud-based virtual computing laboratories were seen as the major benefits for adopting cloud computing applications.

Another survey done by [21], aimed at demonstrating that cloud computing plays an increasingly significant role in higher education institutions. It was noted that higher education institutions should embrace cloud-computing services because of accrued benefits such as economic advantages, increase productivity, and improve learning strategies and knowledge penetration. However, different issues such as privacy, integrity, and ownership of data were not considered in the survey. Despite the different issues that were not considered in the survey, it can be argued that cloud computing can beneficial to Colleges of Education in Zambia.

In the same vein [22], proposed a hybrid-computing model that facilitates the higher educational institutions in Saudi Arabia. It was noted that the proposed model improved the effectiveness and quality of teaching by providing support regarding course material, assessments, and projects. This should be seen as one of the reasons why Colleges of Education in Zambia should adopt Cloud Computing applications. However, the model did not consider security issues as a hindrance to the adoption of cloud computing applications. On the other hand [23] evaluated virtualized computing environments based on cloud computing using the On-demand Deployment of Infrastructures to Support Educational Activities (ODISEA) platform at a university in Spain. Findings of the evaluation demonstrated that ODISEA provided students with highly ubiquitous access to resources and strong economic benefits for higher education institutions. However, it is noted that despite the platform having a lot of flexibility, it had challenges due to the complexity of communication among its levels. Table 2 below summarizes the benefits of the adoption of cloud computing in education.

**Table 2: Summary of benefits on the usage of cloud computing**

| REFERENCES | BENEFITS | GAPS |
|---|---|---|
| Jaysena and Song (2017) | Scalability, increases availability and reliability | Limited access within campus |
| Al-Hamami and Hashem (2016) | Low Cost, Flexibility, Mobility, Business continuity | Security , Loss of sensitive data, lack of standards to enable multiple clouds to work as a single entity |
| Madhav and Joseph (2016) | Cost saving and Flexibility | Can only be used within campus |
| Khan (2015) | Knowledge at one place, improves effectiveness and quality of teaching, budget saving | Security issues where not considered |
| Alajmi and Sadiq(2016) | Increased productivity, penetration of knowledge and improves educational strategies | Integrity, privacy, security and ownership of data |
| Segrelles and Molto(2016) | Flexible platform | The complexity of communication among levels |

## 6. CHALLENGES AND CONSTRAINT OF CLOUD COMPUTING IN EDUCATION

There are various challenges that come with cloud computing in education. Some of these challenges include security, data privacy, as well as insufficient network. Furthermore, data handling, as well as privacy laws need not be taken flippantly. [24] contends that these challenges are there because of the fact that the responsibility for data storage in cloud computing is in the hands of the provider and this gives great uncertainty of the security at all levels. Such levels include network, host, application and date levels. [25] reveals that all security issues are mostly influenced by the fact that the Cloud Service Providers are not locally available ,hence the lack of specific guarantees and assurances for the security of the information. Furthermore, the ability to adequately address privacy regulations in cloud computing is not mostly adequate. In fact, security threats that are associated with cloud computing are the main challenges Colleges of Education in Zambia should face when considering cloud computing adoption.

[24], identified top threats such as abuse and nefarious use of cloud computing, insecure interfaces and application program interfaces (APIs), malicious insiders, shared technology

issues, data loss, data breaches, denial of service, insufficient due diligence and account hijacking as the major challenges associated with the usage of cloud computing applications usage. In this vein, Public Colleges of Education Zambia must take cognizant of such challenges as they consider adopting and using cloud computing applications. The other possible challenges relate to technical issues. [26] identified some technical challenges relating to the adoption of cloud computing as non availability of service and data lock-in. Additionally, the lack of scalable storage, performance unpredictability and data transfer bottlenecks are also obstacles that could limit the growth of cloud computing usage amongst Public Colleges of Education in Zambia. Furthermore, [11] submits that lack of connectivity; inadequate bandwidth and unstable power supply are some of the barriers affecting the adoption of cloud computing in developing countries. This is actually true for Zambia. In this regard, Colleges of Education in Zambia must mitigate against these challenges to get the full benefits of using cloud computing applications. On the other hand, [16] stressed that availability and accessibility to ICT infrastructure and services by staff and students in Universities of developing countries are limited,

There are various studies that have been done on the challenges of using cloud computing in higher education institutions. [27] carried out study in Nigeria whose aim was to assess the impact and challenges of cloud computing adoption on public universities in South Western Nigeria. The results of the findings revealed that major challenges confronting the adoption of cloud computing are data insecurity, regulatory compliance concerns, lock-in and privacy concerns. Another study done by [28] in Nigeria identified scarcity of ICT infrastructure and lack of access, high cost ownership ,unsteady and inadequate power supply as factors that are limiting the adoption of cloud computing.

However, the challenges that have been identified did not consider the issue of computer anxiety as one of the challenges to the usage of cloud computing applications. Based on the preceding the review of literature, the findings from previous studies on the challenges on the adoption of cloud computing a summarized in table 3 below;

**Table 3: Summary of related studies on challenges pertaining to the adoption of cloud computing.**

| REFERENCES | BENEFITS | GAPS |
|---|---|---|
| Laudon and Laudon(2016) | Great uncertainty of security at all levels | Efficacy |
| Cloud Security Alliance(2013) | Malicious Insiders, data loss, data breaches, account hijacking and denial of services | Computer anxiety |
| Mujinga (2012) | Security | Awareness |
| Truong et al (2012) | Lack of connectivity, inadequate bandwidth and unstable power supply | Perceived ease of use |
| Akin et al (2014) | Data insecurity , lock-in and regulatory compliance concerns | Computer anxiety |
| Abdulsalam and Fatima(2011). | In adequate power supply | Awareness. |

## CONCLUSION

Cloud computing ensures that education institutions spend more of their time on research and learning, rather than on implementing complex IT infrastructure. The main benefit of using cloud computing in education is financial. Using Cloud Computing, organisations do not have to make large investment in their own IT infrastructure like hardware and software but instead purchase computing services from remote providers and only pay for the amount of computing power they actually use. In this regard, education learning institutions will be able to reduce costs when it comes to maintaining the IT infrastructure including licensing, energy consumption, technical labour, as well as hardware due to virtualization that comes with cloud computing. This is possible because cloud-based applications can be run on Internet browsers. Finally, it was pointed out that there are various challenges that come with cloud computing in education. Some of these challenges include security, data privacy, as well as insufficient network. However, it was concluded and recommended that cloud computing technology as a tool can be used and it should be rapidly expanded in public Colleges of Education in Zambia. This is especially so, because cloud computing has the ability to collaboratively share, edit, process, and store huge amounts of data and this have obvious applications within the research and educational communities.

## REFERENCES

[1] C .Blurton, (1999). World Communication and Information Report: New Directions of ICT- Use in Education. Paris: UNESCO.

[2] D .M,,Watson, (2001). Pedagogy before Technology: Re-thinking the Relationship between ICT and Teaching. Education and Information Technologies, 6(4) 251-266

[3] W. Pelgrum, (2001). ICT in Education Around the World: Trends, Problems and Prospects. Paris: UNESCO

[4] UNESCO (1998). World Education Report 1998: Teachers and Teaching in a Changing World. Paris: UNESCO

.

[5] P.Mell, and T. Grance, (2011). The NIST Definition of Cloud Computing (Technical report). National Institute of Standards and Technology: U.S. Department of Commerce. doi:10.6028/NIST.SP.800-145. Special publication 800-145.

[6] A.S Weber, (2013).Cloud computing in education In Ubiquitous and Mobile Learning in the Digital Age, Springer, New York, pp. 19-36.

[7] L.M Vaquero, (2011). EduCloud: paaS versus IaaS cloud usage for an advanced computer science course in IEEE Trans. Educ., 54 (4), pp. 590-598

[8] L.A. González-Martínez,,M.L Bote-Lorenzo, ,F. Gómez-Sánchez,E and R. Cano-Parra,

[9] T.S.Behrend, F.N Wiebe, . London, J.E Johnson, E.C. (2011) Cloud computing adoption and usage in community colleges in Behavior Information Technology., 30 (2) , pp. 231-240 .

[10]N. James, and I. Weber. (2016) . Chapter 7 Cloud Computing in EducationCloud Computing in Ocean and Atmospherics pages 107-119 https://doi.org/10.1016/B978-0-12-803192-6.00007-4Get rights and content

[11] H.L.Truong,, T.V Pham, ,N Thoai, and Dustdar (2012). Cloud computing for Education and Research in Developing countries in Cloud Computing for Education and Research,IGI Global.pp.78-94

[12]K.P.N. Jayasena, and H. Song,.(2017). Private Cloud with e-Learning for Resources Sharing in University Environment, in E-Learning, E-Education, and Online Training, Springer, Cham, 2017, pp. 169–180.

[13].A. Rastogi.(2010).A Model based Approach to Implement Cloud computing in E-governance In International Journal of Computer Applications,Vol 9,no.7,pp 15-18

[14]M.Hachibozu,(2016).Cloud computing : A way to go? In The Accountant, Zambia's Accountancy Journal, no 53.pp.23-25

[15]S.F.M. Shoshtari, (2013). Cloud computing adoption in Iran as a Developing Country-A tentative Framework Based on Experience from Iran.Computing Human behavior 23(1,p 175-191

[16]M. Gerald, and K. Eduan ,(2012). Cloud computing in higher education: Implications for South Africa Public Universities and FET Colleges. Annual conference on WWW applications. Decision Processes, 50(2), 179-211.

[17]O.C.Akin, (2014). The impact and challenges of cloud computing adoption public universities in South western Nigeria in International Journal of Advance Computer Science and Applications(IJACSA),vol.5 no. 8 pp.13-

[18]H.H. AL-Hamami, . and S. H. Hashem, (2016). Sustainable Development: Proposing Cloud Computing Framework for Higher Education Ministry (HEM) in Iraq, International Journal on advanced Studies in Computer Science Engineering. Gothenbg., vol. 5, no. 11, pp. 156–163,

[19]N. Madhav. and M.KJoseph, (2016).Cloud-based Virtual Computing Labs for HEIs, in 2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech), pp. 373–377.

[20]O. Alajmi, and A.Sadiq, (2016). What should be done to achieve greater use of cloud computing by higher education institutions, in IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016, pp. 1–5.

[21]M. A Khan, (2015).A Hybrid Cloud Computing Model for Higher Education Institutions in Saudi Arabia, in Cloud Computing, pp. 255–259.

[22]J.D. Segrelles, and G. Moltó, (2016). Assessment of cloud-based Computational Environments for higher education," in 2016 IEEE Frontiers in Education Conference (FIE), pp. 1–9.

[23]Cloud Security Alliance (2013). The notorious nine cloud computing threats in Cloud Security Alliance.Competing Models. Information Systems Research, Vol. 6, pp. 144 – 176

[24] K.C.,Laudon,andJ.P.Laudon,(2016).*Management Information Systems, Managing the Digital Firm*,13th Ed.Harlow,Essex: Pearson Education Limited.

[25] M. Mujinga (2012) .Developing Economics and cloud Security: A study in Africa in Journal of Emerging Trends in Computing and Information Sciences,Vol.3,no.8.pp1166-1172

[26] M. Armbrust, A.Fox,,R. Griffith,A. Joseph,, and R.Katz,(2010). A view of cloud computing in Communications of the ACM, Vol.53,no.4 pp 50-58.

[27] Y.G.Abdulsalam, and U.Z Fatima,(2011). Cloud Computing:Solution to ICT in Higher Education in Nigeria in Advances in Applied Science Research.2(6):pp 364-369.

*[28]* O.C.Akin,(2014). The impact and challenges of cloud computing adoption public universities in South western Nigeria *in International Journal of Advance Computer Science and Applications(IJACSA),vol.5* no. 8 pp.13-19.

# Paper Title: Web Based Monitoring and Detection of Copper Cable Cuts in Fixed Access Networks

*Ndiwa Mutemwa*
**Electrical Department School of Engineering Copperbelt University Kitwe, Zambia*
*ndiwa.mutemwa@cbu.ac.zm*

*Moris Matoomana **
*morrismatoomana@gmail.com*
*Owen Hangoma**
*owen.hangoma@yahoo.com*

*Makaita Masaita**
*makapresa@gmail.com*
*Muwema Wamuyima**
*smashwamz@gmail.com*

***Abstract*** **Telecommunication service providers relying on copper cables to deliver services to their subscribers are facing massive cable thefts in Zambia leading to unreliable service delivery, loss of revenue and high network maintenance costs. The theft of copper cables has been necessitated by the rise in price and demand for copper globally. Current methods employed to monitor cable cuts in fixed access networks are not only inefficient but are also costly while replacing the copper cables with optic fibre cables might not always be appropriate. This paper therefore discusses a proposal to not only monitor the status of copper cables, but also to detect cable cuts in fixed access networks using cable capacitance and sending the captured cable status information to a centralized remote monitoring centre in real time. A copper cable exhibits capacitance between two conductors which are insulated from each other and this capacitance is directly proportion to the length of the cable. Hence by determining changes in cable capacitance, the change in cable length can be calculated. In the proposed system design a microcontroller calculates the cable distance using the cable capacitance changes. Upon detection of a cable cut in the fixed access network, the system automatically updates the web server at the remote monitoring centre with information indicating the distance from termination to the cable cut through a gateway. It also alerts personnel in the remote monitoring centre by sending an alert signal to the web server thus enabling appropriate intervention measures to be taken to avoid the cable theft and reduce outage time**

**Keywords: Remote, Monitoring, Access Network, Copper Cable, Detection, Subscriber, Capacitance**

.

## I. INTRODUCTION

Access networks provide the means by which subscribers can access services provided by telecommunication operators and also provide the means by which service providers provide services to their clients. The term access network maybe defined as a network that connects subscribers to a core/central network whilst the term wired fixed access network refers to a network between the local exchange and the subscriber and is predominately made up of copper cables based on point to point connections. Access networks maybe broadly categorized into two categories namely fixed access network and wireless access network.

In a wired fixed access network subscribers connect to the local exchange only from fixed locations contrary to wireless access networks which allow subscribers to be mobile. Therefore a fixed access network connects a user on a fixed location to a service provider. The structure of a wired fixed access network typically consists of the main distribution frame (MDF), cross connection point (CCP), copper cables of different sizes and distribution points (DP) as shown in fig 1.Subscribers are directly connected to the distribution points which are served by cross connection points before finally terminating on the main distribution frame which acts as an interface between the access network and the exchange.

.

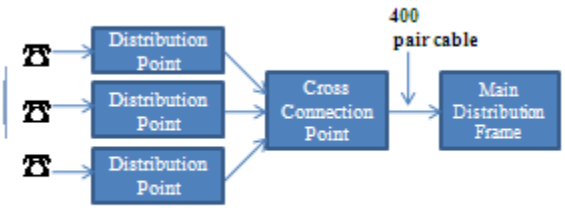



Fig. 1.The structure of fixed access network

Cable failures in fixed access network is on the rise and

hence there must be systems that can help network operators to detect those failures at time of occurrences[7].Cable failures due to cuts are less common in underground cables, this is so because underground cable systems are less accessible compared to overhead lines[4]. However, when faults occur, restoration of underground transmission lines is more difficult than overhead lines. Troubleshooting of overhead lines is fairly easy because it is readily accessible and faults are fairly easy to detect. On the contrary fault locating in underground lines can be tedious than usual.

There are various causes of copper cable failures. Some of the common causes include; corrosion due to aging infrastructure, unintentional cable cuts caused by heavy earth moving equipment, Copper cable theft and vandalism. Whatever the case of failure, there is need for solution for real time detection of faults [1].

In this paper, we are focusing on how copper cable cuts in wired fixed access networks can be automatically detected and localized automatically by analyzing and utilising the properties of copper cable itself. The advantages of this system over existing systems are that it will automate the process of cable cuts detection, enhance real time cable health status monitoring and provide timely network outage restoration. Furthermore, instant cable cut alerts can lead to apprehension of cable vandals or thieves since this is a real time monitoring system. The system is will minimize cable cut localizing durations.

## AIM OF THE RESEARCH

The aim of this research was to develop a real time web based copper cable status monitoring system to be employed in wired fixed access networks

The main objective of this research was to design a model system to monitor, detect and localize copper cable cuts and send the captured copper cable status information to a remote monitoring centre in real time.

## II. LITERATURE REVIEW

Fixed access network is a network which connects users in fixed locations to a core network through wired cables. These cables are usually made of copper conductors. To monitor such cables a pair of copper conductors is used.

[2]The pair used for monitoring purposes has finite distance of separation and are insulated from each other by a nonconductive material forming a dielectric. When potential difference is established across these conductors an electric field will exist between the conductors. This is the same effect as in capacitors. This electric field established between the conductors is the capacitance of the cable and is direct proportional to the length of the cable. The equivalent of this is capacitance in parallel along the length of the cable. The insulators separating the two conductors is not perfect therefore, losses due to leakage current will occur and is represented by parallel shunt conductance per unit length. But this effect of losses can be neglected at low voltages.

[3] Whenever current flows through a conductor, magnetic field is established around the conductor and is represented by series inductance per unit length. However, the pair of conductors under consideration is open circuit and thus only minimal current is considered establishing electric field and therefore the effect of inductance can be also neglected.

Finally, copper conductors are made up of material that has resistance per unit length. This effect can also be neglected since no appreciable will flow through the conductors and there will be no voltage drop.

Since resistance, shunt conductance, and inductance effects have been neglected the result is cable that is considered lossless. Thus the cable we used was assumed to be purely capacitive as shown in fig 2.

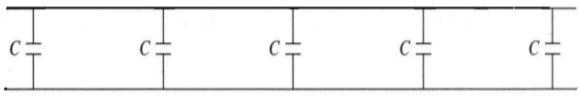



Fig. 2. Equivalent copper cable

These properties helped us to design a cable distance sensing circuit using a 555 timer IC. The 555 timer was employed as a cable capacitance detector. [3]The cable capacitance is directly proportion to length of the cable and this relationship allowed us to determine the distance of the cable. Using the following mathematical relationships, the distance to a cut of the cable was.

Let the cable have the initial length *l1* and capacitance per unit length *C1*, the initial capacitance *C1* of the cable is;

$C1=CPl1$ (1)

Equation (1) shows that the capacitance of the cable is direct proportion to its cable length *l* but inversely proportion to the frequency f of the cable according to (2)

$C \alpha 1/f$ (2)

Therefore

$f1/f2=l2/l1$ (3)

To determine the distance of copper cable cut location, (3) is expressed as:

$l2=(f1/f2)\, l1$ (4)

where *l1* is initial cable length, *f1* is the initial cable frequency, *f2* is the cable cut frequency and *l2* is the cable cut length of interest.

## III. PROPOSED SYSTEM

The system is aimed at automating the process of detecting copper cable cuts and real time monitoring of cables statuses in wired fixed access networks in order to restore networks on time without affecting much of the operations

of the service provider and its esteemed subscribers. Through instant alerts the system shall also help not only to deter copper cable thieves but also apprehend them.

This project is broken down into the following main subsystems: monitor, web-server and database as summarized in Fig 3.

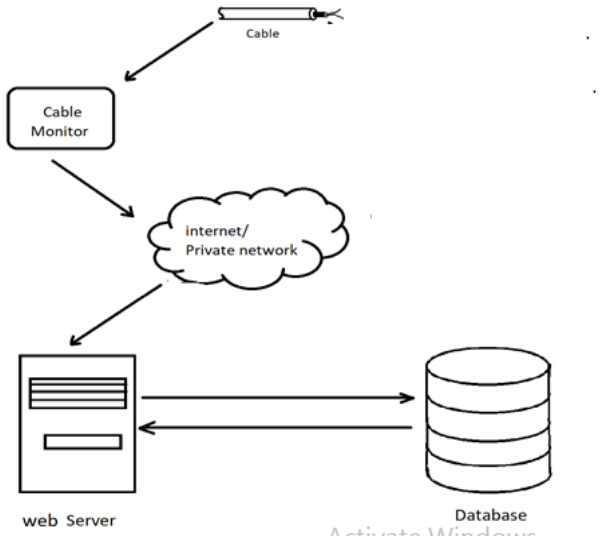



Fig 3. System summary

### *A. Copper cable monitor*

The cable monitor is one of the main subsystems of the project. The function of the monitor is mainly to detect cable cuts and is made up of a cable distance sensor, microcontroller, Network Interface Card (NIC), Keypad and Liquid Crystal Display (LCD) as shown in fig 4.

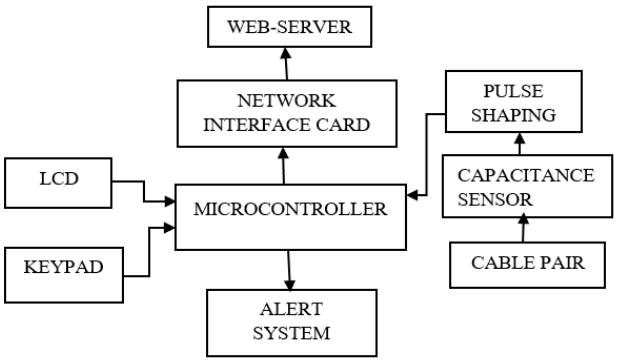



Fig 4. Cable monitoring system

Working principle

[4]The cable capacitance sensor (555 Timer) is a circuit responsible for monitoring frequency of oscillation of the cable and converting capacitance into pulses. The oscillating frequency is determined by the connected cable. The 555 Timer is an Astable multivibrator configuration which is a simple oscillator circuit that generates continuous pulses and that is what will be used to detect cable cuts in this project. The continuous pulses generated from the 555 Timer are the input to the Schmitt trigger circuit. The Schmitt trigger reshapes the pulses into square or rectangular waveforms which are the only types of waveforms the Atmega microcontroller can read. The output of the Schmitt trigger is fed to the microcontroller for the purpose of counting the pulses; the pulses are then further processed for distance calculation in the microcontroller. The microcontroller is the heart of this project and has a special function pulseIn, which enables it to determine the positive state duration or negative state duration of a particular rectangular wave. Fig 5 shows how square pulses are generated.

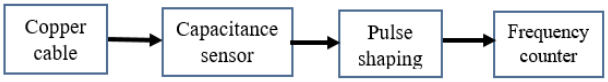
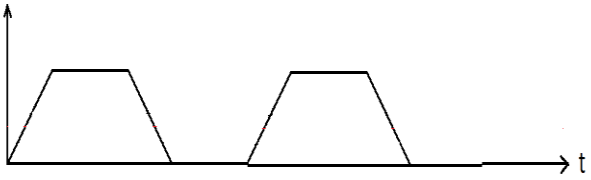


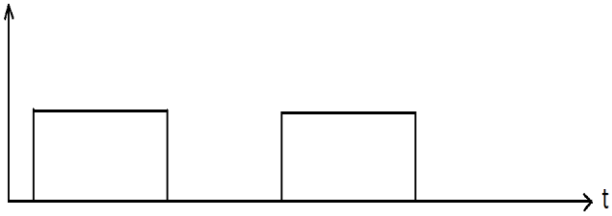

Fig. 5. Generation of square pulses

The Schmitt trigger

Generally, pulses generated by the capacitance sensor are not square waveforms. They can be triangular, sawtooth, sine waveforms and so on. With the microcontroller being able to detect only the square or rectangular waves, we need a device which could alter any signals to rectangular waves, thus we used Schmitt Trigger which is placed at the output terminal of the 555 Timer (Pulse shaping circuit).




(a)input




(b) output

Fig. 6. Waveforms on the Schmitt trigger input/output

Fig 6 shows the steps on how square waveforms are generated by the Schmitt trigger (pulse shaping) and fed to the frequency counter (Microcontroller).

When the cable is cut, the corresponding distance to the cut location and other necessary information is subsequently sent through a network and stored on the

webserver. The cable monitor can be installed at the utility company's premises where the cables are terminated such as the MDF and monitoring for cable-cuts can be done on multiple cables. For every cable to be monitored a non-active line (pair) will be used for the purpose of connection to the cable sensor. When a cable is completely cut due to either vandalism, construction works or theft the cable sensor connected to the pair will detect a cut by analyzing the change of frequency of the cable i.e. applying (4).

### *B. Web-server and Database*

Online access

For the purpose of remote monitoring of cable status, a web server and database were employed to enable online monitoring. The web server allows users to view the status of the cable wherever they may be using a browser provided they have internet connectivity. A database was employed as a memory for the web server. The server handles one request at a time and discards that information immediately. When serve handles another web request, it has no memory of what was the previous request or processed data and this necessitated the use of a database.

The web server periodically receives the cable status updates from cable monitor(s) and polls that information onto a database. This cable information polled onto the database is retrieved by the web server upon request by any user/operator. The database basically is used to store or keep the dynamic cable status updates the web server would have otherwise lost upon handling other requests. The server itself provides permanent storage in form of files stored on the hard drive of the host computer. It would not be a good idea to use such storage for dynamic data but should be ok for information that is not rapidly changing or permanent. In terms of storage, the server stores permanent data while the database stores dynamic data.

#### *User service*

In order to access the status of the cables online, a user logs on to the server and is served the page which shows the table of cables, cable-status and distances respectively. This information is fetched from the database then presented to the user. The way this happens is summarized in the flow diagram of fig 7.

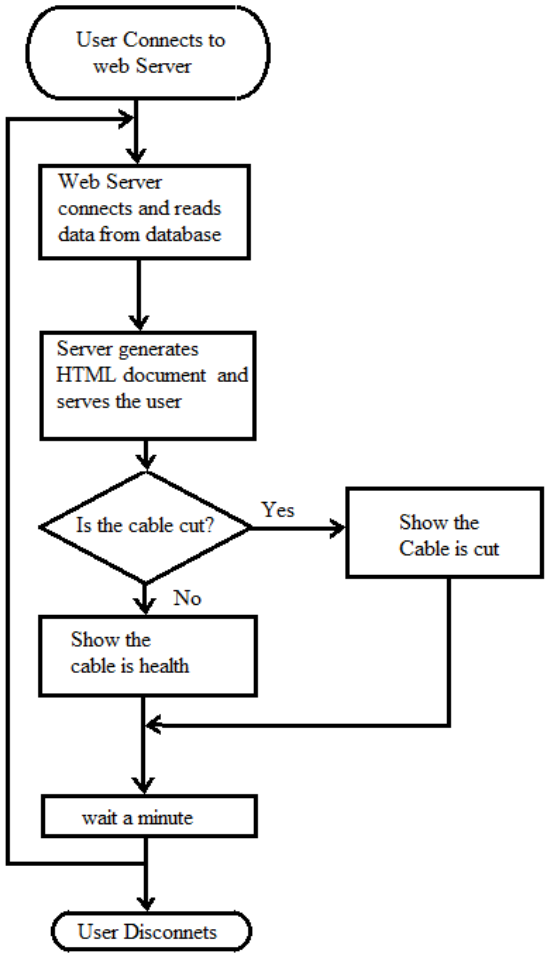



Fig.7. User service flow diagram

#### *Implementation*

Apache web server version 2.4 was employed. The web server handles web requests by users/ web clients and updates from cable monitors in the field microcontroller. Arduino mega 2560 was employed as a monitor CPU and was connected to a network via an Ethernet Shield which is a network interface card (NIC). The monitor read the cable sensor and compares the current reading against the initial reading to determine whether the cable has been cut or not. After determining the status of the cable, the monitor sent an update of the status and the current reading of the cable. The monitor is capable of monitoring more than one cable (up to 12) each cable is identified by a unique number called cable ID. The monitor also has a unique ID to distinguish it from others if more than one monitor were used.

## RESULTS AND DISCUSSION

The cable distance sensor was tested with cables of different lengths of copper cable. A telephone subscriber drop wire was used as a sample test cable

cable and table 1 shows the recorded data during the tests.

| S/n | Sample Actual Cable Length (m) | Sensor Reading Test 1 (m) | Sensor Reading Test 2 (m) | Sensor Reading Test 3 (m) | Sensor Reading Test 4 (m) |
|---|---|---|---|---|---|
| 1 | 7.3 | 6.66 | 6.94 | 7.4 | 6.80 |
| 2 | 1.96 | 1.86 | 2.22 | 2.20 | 1.98 |
| 3 | 41.8 | 40.43 | 37.98 | 39.75 | 39.82 |

Table 1 Sensor test results

From the tests conducted it was discovered that for each measurements recorded, a maximum error of +/- of 15% of the actual value resulted per sample and an average error of +/- 5% out of 100 samples.

Fig 8 is a screen shot from a web browser showing part of the database with health test cables. The health cables are shown in white background and a shade of gray depending on the rows. But each time there is change in the length of the cable being monitored either it is cut or disconnected the background colour of the row identifying the affected cable changes to red on the monitor thereby alerting the personnel of a possible cable cut in the network as shown in figure 9.

| Status | Length | Log Time |
|---|---|---|
| Health cable | 187.82 | 2018-07-16 11:42:07 |
| Health cable | 438.30 | 2018-07-16 11:42:04 |

Fig 8. browser screenshot showing health cable

| Status | Length | Log Time |
|---|---|---|
| Cable is cut | 0.00 | 2018-07-04 09:45:46 |

Fig 9 browser screen shot showing cable disconnected

## CONCLUSION

The main objectives of this research were met, which was to design, simulate and making a prototype for copper cable cuts monitoring, detection and distance localisation in wire fixed access networks.

The error can be reduced further by using components with less tolerances in the sensor circuits. A more stable power supply to the sensor can improve the sensor readings as it is sensitive to voltage changes.

The Microcontroller used has inherent frequency reading errors which can be eliminated by using hardware or a more accurate frequency counter for the sensor .

The proposed system can be incorporated into existing network operations centres thus helping in reducing networks which occur due to cuts due to real time cable detection and monitoring. For the same reasons it can also enhance security of cables as copper cables are prone to thefts especially the primary cable i.e. cables between the main distribution frames and the cross connection points.

.

## *REFERENCES*

[1] R. Shunmugam, Divya, T.G. Janani, P. Megaladevi, P. Mownisha, "Arduino based underground cable fault detector," *International Journal of Recent Trends in Engineering & Research*, vol. II, no. 04 April, pp. 1-4, 2016.

[2] P.M. Dhekale, S.S. Bhise, N.R. Deokate, "Underground cable fault distance locator," *International Journal of Innovations in Engineering Research and Technology*, vol. II, no. 4 April, p. 6, 2015.

[3] S.A. Kale, S.A. Gharpande, G.S. Darvhankar, "3-PH Underground cable fault locator using shock discharge method," *International Journal of Electrical, Electronics and Data Communication*, June 2015.

[4] V.D. Antoniello, E.C. Bascom, "Underground Power Cable Considerations: Alternatives to Overhead," in *47th Minnesota Power Systems Conference*, Brooklyn Center, Minnesota, 2011.

[5] S. Sivakami, G. Ramprabu, V Hemamalini, K. Veronica, C. Thirupoorani, "Fiber Fault Localization in FTTH Using Online Monitoring," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. III, no. 3, March 2015.

[6] *Design Practices and Products for Deterring Copper Wire Theft*, CTC & Associates LLC, May 2013.

[7] A. Chaudhari, N. Sivalenka, "Cutting the risks of Copper cable theft through the IOT," November 2017. [Online]. Available: https://www.globalrailwayreview.com/content_author/avinash-chaudhari-cyent/. [Accessed 22 May 2018].

[8] R. Jamaludin, "GSM Remote Sensing For Transmission Line Monitoring System," 2013.

# The Effects of the Vocabulary Size of Testing Data set on Isolated Word Recognition System: Insights for a Zambian Case

George Mufungulwa [1], Zita Lifelo[2], Tracy Chisanga[3], and Yoshikazu Miyanaga[4]

[1,2,3]School of Information Communication Technology, The Copperbelt University,, Jambo Drive, Riverside, Kitwe, Zambia
[4]Graduate School of Information Science and Technology, Hokkaido University,, Sapporo, 060-0814, Japan
Email ids: [1]mufungulwac@gmail.com, [2]zitalife@gmail.com, [3]tracy.chisanga@gmail.com, [4]miya@ist.hokudai.ac.jp

***Abstract*—In this paper, features for automatic speech recognition (ASR) are based on the contribution of short time energy (STE) and zero-crossing rate (ZCR) as a combined approach on one hand, and STE only on the other. The combined (STEZCR) approach emphasizes important short term frames in a signal while alleviating most noisy and silent frames. Little is known about speech feature extraction algorithms in Zambia. This paper aims to ascertain whether the vocabulary size of the testing data set as a ratio of the vocabulary size of trained data has an effect on automatic speech recognition. In addition, the paper seeks to consider whether gender of subjects does influence system performance in terms of recognition accuracy. The study applies Mel frequency cepstrum coefficients (MFCC) and linear prediction coding (LPC) feature extraction techniques respectively. The vocal frequency is reduced from 3 to 1 over the MFCC and LPC respectively. Testing results of the proposed combined speech features on isolated Japanese speech phrases was done in a clean environment. The average recognition accuracy improvement of 7.83% and 10.28% is achieved in word accuracy when the vocal frequency is reduced from 3 to 1 over the MFCC and LPC respectively. The results clearly confirm that the vocabulary size of the recognizer has an effect on automatic speech recognition. In addition, the combined STEZCR approach performs better on female subjects with the recognition accuracy improvement of between 0.20% and 1.5% while STE performs better on male subjects at 5.50% and 9.0% on MFCC and LPC respectively. However, the results show that gender of subjects has no direct influence on system performance but that a combination of the feature extraction and VAD techniques influences ASR system performance.**

***Index Terms*—MFCC, LPC, VAD, STE, ZCR, STEZCR**

## I. Introduction

Automatic speech recognition (ASR) systems have been in use for a long time now. A number of major world languages, speech phrases and speakers from a number of countries have been implemented in such systems. Despite the significance of the research area in biometric recognition, speech recognition, speaker identification, among others, this kind of research has never been done in Zambia before until now. We aim to trigger interest in this area of research using speech samples collected from Zambians. The overall objective of this paper is to ascertain whether the vocabulary size of the testing speech data set, as a ratio of trained vocabulary size, has an effect on automatic speech recognition. In addition, the authors seek to determine whether gender of subjects has an influence on system performance. In the same process, performance of the short-time-energy and zero-crossings count (STEZCR) as a combined voice activity detector (VAD) on isolated speech phrases is evaluated. The performance of proposed approach is evaluated on female and male speakers using mel frequency cepstral coefficients (MFCCs) and linear prediction coding (LPC) as feature extraction techniques.

Our work in this paper presents feature extraction framework that leverage the technique of short term energy [1] and zero-crossing rate [2] on isolated phrase recognition. In this study, the combined approach here stated is applied in time domain. The noise effect can be dealt with by selecting frames with sufficient energy as well as few zero crossings in the time domain before obtaining the cepstrum. We aim to provide insights into isolated speech recognition systems.

The rest of the paper is organized as follows. In Section 2, a literature review is given. Section 3 outlines the method description. In Section 4, simulation parameters for VAD and conditions of experiments are outlined. In Section 5, the system performance is evaluated. In the same section the results are stated and discussed. Section 6 gives the conclusion.

## II. Literature Review

Speech feature extraction algorithms have become popular in recent times due to the rapid increase in computer memory and computation speed. Speech features can be used for various applications: biometric recognition, speech recognition, speaker identification, and so on. In these applications, a good speech feature can be obtained using Mel frequency Cepstrum Coefficients (MFCC) [3]–[6], Linear Predictive Coding (LPC) [7]–[10], Time varying LPC (TVLPC) [11], Perceptual Linear Predictive (PLP) among others. Speech recognition systems often suffer from multiple sources of variability due to corrupted speech signal features [12]. In compensating

main author: George Mufungulwa
email:mufungulwac@gmail.com

for distortions, most speech recognizers use normalization methods and noise filtering techniques in conjunction with voice activity detection (VAD) techniques [13].

VAD involves detecting silence parts of a speech or audio signal. VAD system consists of a feature extraction that extracts a set of parameters from the signal and a speech/non-speech decision based on a set of decision rules. VAD methods include the ones based on energy threshold, pitch detection, spectrum analysis, zero-crossing rate, periodicity measure [14] and combination of different features. Applications include: speech coding, speech recognition, speech enhancement and audio indexing [15]. In speech coding, for example, VAD helps to avoid the unnecessary coding and transmission of non-speech fragments thus saves bandwidth and computation costs. In speech recognition, VAD is used as noise reduction in digital hearing aid devices and in real-time VoIP applications.

## III. Method Description

In order to obtain cepstrum, speech data was initially pre-emphasized and the pre-emphasized speech waveform in time domain was frame-blocked and windowed with a pre-defined analysis window. Later, fast Fourier transform (FFT) was computed. The magnitude of the output was then weighted by a series of mel filter frequency responses whose center frequencies and bandwidth roughly matched those of auditory critical band filters [16]. The FFT bins were later combined so that each filter had unit weight. From the weighted sums of all amplitudes of signals, a vector was obtained by logarithmic amplitude compression computation and subsequently transforming the result to MFCC parameter by discrete cosine transform (DCT).

Shown in Figure 1 and 2 are steps involved in obtaining speech features using the conventional approaches. The two approaches make use of FFT based MFCC and LPC respectively. The details of FFT based MFCCs and its computation are discussed in [6] [17] while those of LPC and its computation are discussed in [18] respectively. The authors decided on these method for its flexibility and its popularity.

### A. Mel Frequency Cepstral Coefficients

Figure 1 shows the MFCC feature extraction process. MFCC is used to extract the unique features of human voice. It represents the short term power spectrum of human voice. The MFCC is used to calculate the coefficients that represent the frequency Cepstral coefficients based on the linear cosine transform of the log power spectrum on the nonlinear Mel scale of frequency. In mel scale the frequency bands are equally spaced that approximate the human voice more accurately. The formula used to convert the normal frequency f to the Mel scale m is as shown in Eq. 1

$$m = 2595log10(1 + f/700). \quad (1)$$

Mel scale and normal frequency scale are referenced by defining the pitch of 1000 Mel to a 1000 Hz tones, 40 db above the listeners threshold. Mel frequency are equally spaced on the Mel scale and are applied to linear space filters below 1000 Hz to linearize the Mel scale values and logarithmically spaced filter above 1000 Hz to find the log power of Mel scaled signal [6] [19].

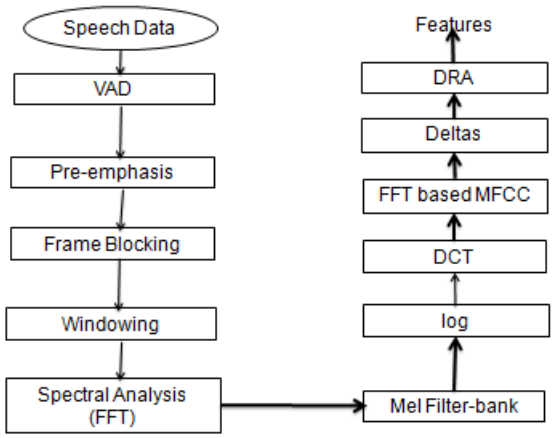



Fig. 1. Fast Fourier Transform (FFT) based Mel Frequency Cepstral Coefficients (MFCCs) feature extraction process. VAD: Voice Activity Detection, DCT: Discrete Cosine Transform, DRA: Dynamic Range Adjustment.

### B. Linear Prediction Coding Coefficients

Linear prediction coding (LPC) is based on the hypothesis that an arbitrary sample can be represented by a linear function of the preceding samples [10] as shown in Eq.2. For all-pole signal modeling, the output signal s[n] at time n is modeled as a linear combination of the past p samples and the input u[n], with G as a gain constant i.e.,

$$s[n] = -\sum_{i=1}^{p} a_i[n]s[n-i] + Gu[n]. \quad (2)$$

Therefore, so as to minimize the difference between the estimated wave obtained by applying linear prediction model to speech wave and the original speech wave, the linear prediction coefficients are determined. The spectrum obtained by LPC, is waveform that matches the human aural characteristic, and is as shown in Fig. 2.

### C. Parameters for VAD

Depending on the speech attributes, a voiced signal can be classified into speech or silence. Because speech signals vary with time, this process is done on short chunks of the speech signal called frames.

*1) Energy of a Frame:* Short term energy is a simple short-term speech measurement. It is defined as:

$$E_n = \sum_{m=-inf}^{m=inf} [x(m)w(n-m)]^2. \quad (3)$$

Since the speech signal is a nonstationary processing, the way which is used to process stationary signal cannot be used to process a speech signal. However, the speech signal in 10

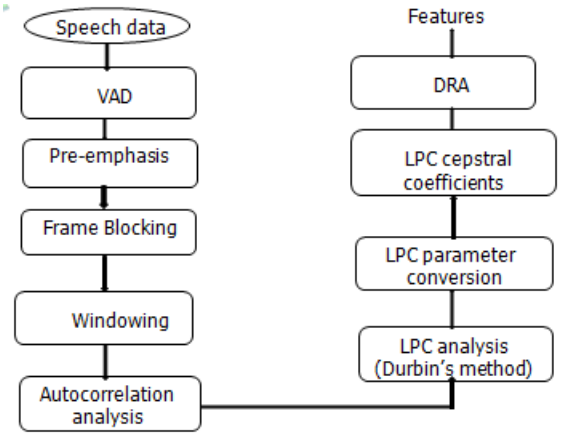



Fig. 2. Linear Predictive Coding (LCP) feature extraction stages. VAD: Voice Activity Detection, DRA: Dynamic Range Adjustment.

$\sim$ 30 ms time can be as a quasi-steady signal (as short-term steady state), because the parameters of spectrum and physical characteristics are almost invariant [20] [21]. Thus, a speech signal can be divided into many short frames and every frame is as a detecting unit. According to the energies of speech and nonspeech frame, short-term energy based VAD approach can identify endpoints of any speech signal, because the energy of speech frame is larger than that of nonspeech frame [22] [23] [24]. The samples of a waveform of input speech signal is defined as x(m), m is the sample index.

The short-time square energy of speech signal Esqr(n) is defined according to Eq.4

$$Esqr(n) = \sum_{m=-\infty}^{+\infty} [x(m)\omega(m-n)]^2 \tag{4}$$

The $\omega(n)$ is a window function which is small width in samples, representing the frame size of $n$. Usually the rectangular, Hamming and Hanning window functions are used to speech signal processing. In this paper, we make use of the hanning window.

The Hanning window function is defined as shown in Eq. 5,

$$\omega(n) = \begin{cases} 0.5(1-\cos(\frac{2n\pi}{N-1}) & 0 \le n \le N-1 \\ 0 & otherwise \end{cases} \tag{5}$$

In this paper, a simple short-time speech measurement and a Hanning window are utilized. This measurement helps in distinguishing between voiced and unvoiced speech segments, since unvoiced speech has significantly smaller short-time energy. For the length of the window, a practical choice is one between 20 $\sim$ 30 ms. The window of this length will include a suitable number of pitch periods so that the result will be neither too smooth nor too detailed.

*2) Zero Crossing Detector:* The humans pronunciation include the surd and sonant. The sonant is produced by the vibration of the vocal chords. The amplitude of sonant is high and periodicity is often visible. The surd is without vibration of the vocal chords, it is produced by the friction, impact or plosive that the suction of air into the mouth. Thus, the short-term energy is lower than that of sonant. It can be identified into nonspeech easily by short-term energy method [25]. If the nonspeech and surd segments are zoomed, we found the waveform of surd segment goes up and down so quickly around zero level value, and the number of crossing zero level value for nonspeech segment is fewer. The number of crossing zero level value can be used to distinguish the endpoint of speech signal. The method is described as zero-crossing rate (ZCR) [26] [27].

### *D. Simulation Parameters and Conditions of Experiments*

The simulation parameters shown in Table I are used in testing the isolated words. The isolated words consist of Japanese common speech data for both male and female speakers. In this study, 30 male speakers and 30 female speakers each uttering 100 common speech phrases with utterance frequency of 3 are utilized. The speech sample is 11.025 KHz and 16-bit quantization. FFT based MFCC and LPC features are separately extracted after pre-emphasis and Hanning windowing. In both cases the features are then converted to 38-dimensional features vectors, (12 statistic coefficients, 13 delta coefficients and 13 delta-delta coefficients). Frame length and shift length are 23.2 ms (256 samples) and 11.6 ms (128 samples) respectively.

TABLE I
THE CONDITION OF SPEECH RECOGNITION EXPERIMENTS

| Parameter name | Parameter value/type |
|---|---|
| Sampling | 11.025 kHz (16-bit) |
| Frame length | 23.2 ms (256 samples) |
| Shift length | 11.6 ms (128 samples) |
| Pre emphasis | $1-0.97z^{-1}$ |
| Windowing | Hanning window |
| Speech Feature vectors | $b_i(i=1,\ldots,12)$ $\Delta b_i(i=0,\ldots,12)$, $\Delta^2 b_i(i=0,\ldots,12)$, |
| Training Set | 30 male , 30 female 3 utterances each |
| Tested Set | 10 male, 10 female, 3 utterances each & 1 utterance each |
| Acoustic Model | 32-states isolated phrase HMMs |

The main methods used for speech enhancement are STE and a combination of STE with ZCR. We evaluate the adaptability of our proposed combined STE with ZCR (STEZCR) over time domain and compare the results to those of STE by gender and by feature extraction techniques under consideration,respectively.

Training sets of 30 male speakers and 30 female speakers, with each speaker uttering 100 speech phrases are utilized

for the front-end feature extraction. Each phrase is repeated 3 times. A 32-states isolated phrase hidden Markov modeling (HMM) is used in both training and recognition. Testing sets consisting of 10 male speakers and 10 female speakers (not used in training), with each speaker uttering 100 words and each word repeated three (3) times and once (1) respectively are utilized.

Frame-by-frame, 38-dimensional FFT based MFCC feature vectors are extracted after pre-emphasis and Hanning windowing. In the testing stage, we compare the performance of proposed combined VAD approach to the one based on STE.

## IV. Findings and Results

In this simulation we evaluate the performance of the conventional short term energy (STE) as well as a combination of STE and ZCR (STEZCR) on FFT based MFCC and LPC using MATLAB (R2015a) software. In the testing stage, we measure the average recognition rates of 10 independent male and female persons uttering 100 phrases. Independent speakers (not used in the training) are used in HMM recognition in both proposed and conventional approach experiments. We measure the average recognition rates of 10 independent male and female persons, uttering 100 words each repeated three (3) times on clean speech and each repeated once (1) on clean speech, respectively.

Table II shows the average recognition accuracy for male and female speech phrases on clean speech. Table III shows the average recognition improvement ( %) by gender. Table IV shows the average recognition performance increase for STEZCR as it compares with STE approach. Table V shows the combined average performance (%) by Feature Extraction technique.

### A. Simulation Results and Analysis

Analysis is carried out for the common Japanese phrase database. We consider feature extraction techniques(FFT based MFCC and LPC), VAD (STEZCR and STE) techniques, gender (male and female) and vocal counts (3 and 1) as variables. Results analysis focuses on the performance of the combined STEZCR on the various acoustic measures. No noise is considered. In this paper the model formulation is as follows: the model using FFT based MFCC and LPC coefficients consists of 38-dimensional feature vectors respectively. In both cases, the 38-parameter feature vectors consisting of 12 cepstral coefficients (without the zero-order coefficient) plus the corresponding 13 delta and 13 acceleration coefficients is given by $[b_1 b_2 \dots b_{12} \Delta b_0 \Delta b_1 \dots \Delta b_{12} \Delta^2 b_0 \Delta^2 b_1 \dots \Delta^2 b_{12}]$ where $b_i$, $\Delta b_i$ and $\Delta^2 b_i$, are MFCC, delta MFCC and delta-delta MFCC, respectively in the case FFT based MFCC.

*1) Results Explanation:* Table II shows the average recognition accuracy for 100 Japanese common female and male speech phrases on clean speech. On FFT based MFCC, STEZCR performs slightly better at 98.90% compared with STE at 98.60% for female speakers. For male speakers, STEZCR performs better at 98.60% compared with STE at 97.10%. On LPC, STEZCR performs better at 97.50% compared with STE at 97.00%. However, for male speakers, STEZCR performs slightly lower at 97.20% compared with STE at 97.30%.

Overall, the combined STEZCR VAD approach performs better than STE alone.

Table III shows the performance improvement by gender. In the first stage, experiments were conducted with three (3) vocabulary size of the recognizer and three (3) trained vocabulary size. In the second experiment, the vocabulary size of the recognizer was reduced by two (2) vocal frequencies, leaving the vocabulary size of one (1). The results from the two separate experiments were compared on both gender and VAD techniques under consideration. The results shown in Table III are differential improvement on gender with respect to VAD and speech feature extraction techniques applied. For FFT based MFCC, the female speakers on STEZCR gave 9.00% increase when the vocal frequency was reduced from three (3) to one (1), while the male speakers gave 4.00%, respectively. On the other hand, the female and male speakers on STE gave 8.80% and 9.50% increase, respectively.

When the vocabulary size of the recognizer as a ratio of trained vocabulary size is reduced from three (3) vocal frequencies to one (1) vocal frequency the recognition accuracy increases.

Results confirm that the testing vocabulary size of the recognizer as a ratio of trained vocabulary size has an effect on automatic speech recognition. Increase in the vocabulary size of the recognizer reduces the system performance while its reduction results in improved system performance.

From the same results, STEZCR performs better on female speakers while STE performs better on male speakers.

Table IV shows the STEZCR performance increase compared with STE. In this table, we show the performance increase. On FFT based MFCC, STEZCR yields 0.10% and 0.30% on female speakers for recognizer size of three (3) and size of one (1) respectively. On FFT based MFCC, STEZCR yields 7.00% and 1.50% on male speakers for recognizer size of three (3) and size of one (1) vocal frequency, respectively. On LPC, STEZCR yields -1.00% and 0.50% on female speakers for recognizer size of three (3) and size of one (1), respectively. On LPC, STEZCR yields 8.50% and -0.10% on female speakers for recognizer size of three (3) and size of one (1), respectively.

Except for LPC on female speakers with recognizer size of three (3) and LPC on male speakers with recognizer size of one (1), STEZCR performs better in six (6) of the eight (8) instances under consideration. The highest gain in accuracy is 8.50% on LPC for male speakers with recognizer size of three (3) followed by 7.00% on FFT based MFCC for male speakers with recognizer size of three (3), respectively.

Table V shows the combined average performance for each feature extraction technique irrespective of gender. In obtaining the average performance, we add the improvement (%) by gender for each feature extraction method and divide the result by 4. The purpose is to show the best of the two

TABLE II
AVERAGE RECOGNITION ACCURACY ON CLEAN SPEECH

| 1 | STEZCR | | | | STE | | |
|---|---|---|---|---|---|---|---|
| Feats. Ext | Gender | Train | Test | (%) | Train | Test | (%) |
| MFCC | female | 3 | 3 | 89.90 | 3 | 3 | 89.80 |
| | | 3 | 1 | 98.90 | 3 | 1 | 98.60 |
| | male | 3 | 3 | 94.60 | 3 | 3 | 87.60 |
| | | 3 | 1 | 98.60 | 3 | 1 | 97.10 |
| LPC | female | 3 | 3 | 87.10 | 3 | 3 | 88.10 |
| | | 3 | 1 | 97.50 | 3 | 1 | 97.00 |
| | male | 3 | 3 | 90.80 | 3 | 3 | 81.90 |
| | | 3 | 1 | 97.20 | 3 | 1 | 97.30 |

TABLE III
IMPROVEMENT (%) BY GENDER

| 1 | STEZCR | | STE |
|---|---|---|---|
| Feats. Ext | Gender | (%) | (%) |
| MFCC | female | 9.00 | 8.80 |
| | male | 4.00 | 9.50 |
| LPC | female | 10.40 | 8.90 |
| | male | 6.40 | 15.40 |

TABLE IV
STEZCR PERFORMANCE INCREASE COMPARED WITH STE (%)

| 1 | STEZCR | | |
|---|---|---|---|
| Feats. Ext | Gender | vocal | (%) |
| MFCC | female | 3 | 0.10 |
| | | 1 | 0.30 |
| | male | 3 | 7.00 |
| | | 1 | 1.50 |
| LPC | female | 3 | -1.00 |
| | | 1 | 0.50 |
| | male | 3 | 8.50 |
| | | 1 | -0.10 |

TABLE V
COMBINED AVERAGE PERFORMANCE (%) BY FEATURE EXTRACTION TECHNIQUE

| Feats. Ext | (%) |
|---|---|
| MFCC | 7.83 |
| LPC | 10.28 |

feature extraction methods under consideration. FFT based MFCC give a 7.83% while LPC give 10.28%. LPC performs better than FFT based MFCC overall.

## V. DISCUSSION OF RESULTS

In this section we discuss the findings of our experiments. We show the positive contributions in applying the combined VAD approach on isolated speech recognition.

The accuracy of a speech recognition system can be defined as the percentage of time that the recognizer correctly identifies an input utterance. Recognition errors can be generally classified as misrecognitions or as nonrecognition errors. The tendency of differences in recognition accuracy between male and female can be attributed to many factors including user characteristics (age, gender), the language (vocabulary size), and the channel and environment (noise), for example, among many others [28]. The more varied the group of speakers using the system, the more challenging the recognition process. It is more difficult for a speaker-independent system to recognize accurately both male and female speakers.

The most limiting problem of larger vocabulary sizes is the corresponding decrease in recognizer accuracy. This refers to the total number of different phrases the speech recognizer is able to identify. Therefore, the tendency of differences in recognition accuracy between the 100 Japanese phrases is due to the differences in number of times a particular words needs to be recognized among the trained wave forms. A smaller database has an increased chance of better recognition accuracy compared to a much larger database (of 900 waveforms), in this case. In the latter increased number of nonrecognitions and false recognitions are often recorded as a result, compared

to the former. To alleviate this challenge, the HMM training must be done several times to ensure definite convergence.

The combined (STEZCR) approach not only emphasizes important short time frames in a signal based on energy but also alleviates most noisy and silent frames by using zero crossings.

## VI. Conclusions

It can be concluded that when the testing vocabulary size of the recognizer as a ratio of trained vocabulary size is reduced the recognition accuracy increases. It can be ascertained that the vocabulary size of the recognizer as a ratio of trained vocabulary size has an effect on automatic speech recognition. In general, the gender of subjects does not have a direct influence on the system performance. The evaluation results have demonstrated the effectiveness of a combined VAD approach as a noise reduction as opposed to a single VAD approach. Of the feature extraction techniques considered in this paper, FFT based MFCC shows better performance than LPC. Similarly, the combined VAD approach performs better on male speakers than on female speakers. In future we intend to perform similar experiments under noisy conditions. In addition we intend to evaluate our proposed method on recognizing younger and older persons. Similar experiments would be equally beneficial to Zambia as a basis for research in biometrics recognition, speech recognition and speaker identification, respectively.

## Acknowledgment

The authors would like to thank the Graduate School of Information Communication Technology, ICN Laboratory, Hokkaido University, Hokkaido, Japan.

## References

[1] Abhijeet Sangwan, Chiranth M C, H.S. Jamadagni, Rahul Shah, R. Venkatesha Prasad, Vishal Gaurav,"VAD Techniques for Real-Time Speech Transmission on the Internet", IEEE pp. 46-50

[2] Fabien Gouyon, Francois Pachet and Olivier Delerue, "On the use of Zero-crossing Rate for the application of Classfication of Percussive Sounds", Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00), Verona, Italy, December 7-9, 2000

[3] Safdar Tanweer, Abdul Mobin, and Afshar Alam,"Analysis of Combined use of NN and MFCC for Speech Recognition," International Journal of Computer, Electrical, Automation, Control and Information Engineering, Vol. 8, No. 9, 2014.

[4] L. Muda M. Begam, I. Elamvazuthi, "Voice Recopgnition Algorithm using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques," Journal of Computing, Vol. 2, No. 3, pp. 138-143, 2010.

[5] C. Ittichaichareon, S. Suksri, and T. Yingthawornsuk, "Speech Recognition Using MFCC," International Conference on Computer Graphics, Simulation and Modelling (ICGSM2012), pp. 135-138, 2012.

[6] Anjali Bala, Abhijeet Kumar, Niddhika Birla,"Voice command recognition system based on MFCC and DTW," International Journal of Engineering Science and Technology, Vol. 2, No. 12, pp. 7335-7342, 2010.

[7] Masumi Watanabe, Hiroshi Tsutsui, and Yoshikazu Miyanaga,"Robust speech recognition for similar pronunciation phrases using MMSE under noise environments," Proc. 13th International Symposium on Communications and Information Technologies (ISCIT), 2013.

[8] J. Tierney, "A study of LPC analysis of speech in additive noise," IEEE Trans. on Acoustic., Speech, and Signal Process., Vol. ASSP-28, No. 4, pp. 389-397, Aug. 1980.

[9] Pramod B. Patil, "Multilayered Network for LPC based Speech Recognition," IEEE, 1998.

[10] S.M. Kay,"Noise compensation for autoregressive spectral estimation," IEEE Trans. on Acoust., Speech, and Signal Process., Vol. ASSP-28, No. 3, pp. 292-303, Mar. 1980.

[11] Mark G. Hall, Alan V. Oppenheim, and Alan S. Willsky,"Time-varying Parametric Modelling of Speech," Signal Processing, Vol. 5, pp. 267-285, 1983.

[12] K. Yao, K. K. Paliwal and S. Nakamura, "Model-based noisy speech Recognition with Environment Parameters Estimated by noise adaptive speech Recognition with prior," EUROSPEECH 2003-GENEVA, Switzerland, Tech. Rep., 2003.

[13] Marko Kos,"Noise Reduction Algorithm for Robust Speech Recognition Using Minimum Statistics Method and Neural Network VAD," 2007 14th International Workshop on Systems, Signals and Image Processing and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services,27-30 June 2007.

[14] R. Tucker, "Voice activity detection using a periodicity measure", IEE Proceedings-I, vol. 139, no. 4, August 1992

[15] J Ramirez, J M Gorriz and J C Segura, "Voice Activity Detection. Fundamentals and Speech Recognition System Robustness, pp 460, I-Tech, Vienna, Austria, Jun 2007

[16] Eslam Mansour mohammed, Mohammed Shraf Sayed, Abdalla Mohammed Moselhy and Abdelaziz Alsayed Abdelnaiem,"LPC and MFCC Performance Evaluation with Artificial Neural Network for Spoken Language Identification, " International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 6, No. 3, Jun. 2013.

[17] D.Sanjib,"Speech Recognition Technique: A Review," International Journal of Enginering Research and Applications, vol. 2, no. 3, pp. 2071-2087, 2012

[18] J.I. Makhoul,"Linear Prediction: A tutorial review," in Proc. IEEE vol. 32. April 1975, pp. 561:582

[19] http://en.wikipedia.org/wiki/Mel$_scale$

[20] L. Rabiner and B.-H. Juang, Fundamentals of speech recognition, 1st ed. UpperSaddle River, New Jersey, USA: Prentice Hall PTR, 1993.

[21] L. Rabiner and R.W.Schafer, Digital Processing of Speech Signals, 1st ed. Upper Saddle River, Prentice HallUSA: Rainbow-Bridge Book Company PTR, 1978.

[22] G. Hongbin, P. Weiyi, H. Chunru, and Z. Yongqiang, A speech endpoint detection based on dynamically updated threshold of box-counting dimension, in International Forum on Information Technology and Applications, vol. 2, May.2009, pp. 397401

[23] J. C. Junqua, B. Mak, and B. Reaves, A robust algorithm for word boundary detection in the presence of noise, IEEE Transactions on Speech and Audio Processing, vol. 2, no. 3, pp. 406412, Jul. 1994.

[24] L. R. Rabiner and M. R. Sambur, An algorithm for determining the endpoints for isolated utterances, The Bell System Technical Journal, vol. 54, no. 2, pp.297315, 1975

[25] Sun Xihao, PhD Thesis, Hokkaido University, Sapporo, Hokkaido, Japan , 2014

[26] D. G. Childers, M. Hahn, and J. N. Larar, Silent and voiced/unvoiced/mixed excitation (four-way) classification of speech, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 37, no. 11, pp. 17711774, Nov. 1989

[27] C Gan R. Donaldson, Adaptive silence deletion for speech storage and voice mail applications, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 36, no. 6, pp. 924927, Jun. 1988.

[28] Sherry P. Casall and Robert D. Dryden, "The Effects of Recognition Accuracy and Vocabulary Size Of A Speech Recognition System on Task Performance and User Acceptance," Industrial Engineering and Operations Research , 1988.

# DEVELOPING AN AUTOMATED FALL ARMY WORM (FAW) IDENTIFICATION AND EARLY WARNING AND MONITORING SYSTEM BASED ON ANN TECHNIQUES

Francis chulu
*The University of Zambia*
*Department of Computer Science*
Lusaka, Zambia
francis.chulu@unza.zm

Jackson Phiri
*The University of Zambia*
*Department of Computer Science*
Lusaka, Zambia
jackson.phiri@cs.unza.zm

Phillip Nkunika
*The University of Zambia*
*Department of Biological Sciences*
Lusaka, Zambia
pnkunika@unza.zm

Mayumbo Nyirenda
*The University of Zambia*
*Department of Computer Science*
Lusaka, Zambia
mayumbo.nyirenda@cs.unza.zm

Monica M. Kabemba
*The University of Zambia*
*Department of Computer Science*
Lusaka, Zambia
monica.kalumbilo@cs.unza.zm

Miyanda Moonga
*The University of Zambia*
*Department of Biological Sciences*
Lusaka, Zambia
miyandamoonga@gmail.com

***Abstract*— Since its reported presence in Africa in 2016, the fall army worm (FAW-Spodoptera frugiperda) has caused major damage to a good number of plant species including maize which is a stable food for most African countries. Their presence in Africa pauses a challenge to the food security in many African countries contributing to the already existing food problem that the continent has been facing. This possess a challenge to stakeholders such as FAO, governments, Universities and other stakeholders involved in research to come up with precise and proactive methods of monitoring and controlling the FAW pest. This paper therefore, proposes a study to develop an automated fall army worm identification and early warning and monitoring tool for the Zambian species based on Artificial Neural Network (ANN). The study will aim to address current challenges that entomologists are facing when using the pheromone traps as a way of monitoring the occurrence of FAW pests in Zambia. We will modify pheromone traps and automate them with sensors for automatic data collection. We will develop an algorithm based on ANN for identifying the FAW moth, then we develop web and mobile applications integrated with geographic information system (GIS) technology. The developed system will be able to provide some near real time FAW occurrences in Zambia. The tool will improve the accuracy and efficiency of FAW monitoring and reduce manual data collection thereby reducing the aspect of human intervention. In addition, it will act as a source of data that can be used by all stakeholders ranging from FAO personnel, government, small scale and commercial farmers in making good and informed decisions.**

[1] ***Keywords— Artificial Neural Networks, Fall Army Worm, Identification, Machine learning, Single Board Computer, Pheromone.***

## I. INTRODUCTION

There are a number of factors that can be classified as threats to the African and global Food security. These may include but not limited to climate change, droughts, emerging diseases, salt soils, fertilizer dependence and Pests (insects) [1]. Insect threats has been one of the yearly occurrences in many African countries with new pests being introduced on the continent from other regions of the world as stowaways on commercial aircrafts, brought in cargo containers or airplane holds before they are widely dispersed by wind throughout the continent [2]. One of the pests that have affected the African continent from other regions of the world is the fall army worm (FAW; Spodoptera frugiperda) which was first reported as present on the continent in January 2016. According to [3], subsequent investigations have showed that the pest has been identified in over 30 sub-Saharan African countries where it has caused extensive damage to crops especially maize fields. [3] Further states that though new pests are introduced into the African agricultural environment and pose some degree of risk, FAW possess characteristics that make it more devastating than other pests. Some of the characteristics presented in [3] includes (i). It is capable of feeding on over 80 different crop species making it the most damaging crop pest (ii). It can spread quickly across large geographic areas just like other moths in the genus Spodoptera (iii). It can persist throughout the year. FAW occurrence has negatively impacted maize yield and the economies in most maize producing countries such as Ghana, Zambia and so on. Its impact in Africa has been felt at national, continent and household level [4]. According to [4], the potential impact of FAW on the continents maize yield lies between 8.3 and 20.6 million tons per annum of total expected production of 39m tones per annum and with losses lying between $2,481m and $6,187 per annum of total expected value of $11,590.5m per annum. [4] Further states that FAW directly affects capital costs, through increased labour and increasing cost of production due to costs of control.

Owing to the characteristics of FAW, it presents a challenge on coming up with control measures that can be used alongside the conventional control methods of spraying pesticides. To apply pesticides, an occurrence of the pest in an area has to be reported but the challenge is coming up with methods that are efficient enough to act as early FAW warning systems. According to [3], there has been limited proven approaches to prevent and avoid FAW and efforts to suppress the pest is largely focused on the use of synthetic pesticides which has a high potential to damage human, animal and environmental health. Currently one of the methods being used in Africa is the use of pheromone traps to lure the male FAW moth. The field traps are monitored after a number of days and the trapped moths are counted and recorded on the data collection sheets. The pheromone moth trap counting is tedious, labour intensive, time consuming and expensive owing to the field visits and manual counting and recording of the moth. In addition, the method is prone to error as the recorded moths may be overstated or understated providing false data to interested stakeholders. There is need to improve on the methods used in the monitoring of the pest if there is to be a proactive and reactive response by stakeholders to FAW. One of the

methods that can be used is the automation of the insect identification process targeting mainly the FAW Moth. An introduction of fast systems integrated with modern image processing and analysis algorithms to accelerate the data collection process is needed [5]. Several attempts have been made to try and create a method that can accurately perform insect identification [6] and with the major technological advancements that the world has seen in recent years, we can try to improve on the methods that have already been done or come up with completely new methods. This research is therefore proposing a near accurate and efficient FAW identification method based on ANN techniques. The research will aim to automate the counting of moth thereby reducing the field visits and shorten data collection intervals to within minutes. In addition the system will provide effortless monitoring of farm areas using fewer resources. The system will employ artificial neural network (ANN) techniques for identifying the FAW Zambian species. Further the system will provide a portal integrated with GIS providing a near real time occurrence of the pest in the country.

## II. RELATED WORK

This section of the paper gives a review of literature that focuses on similar related research work as well as how similar challenges have been addressed elsewhere. The literature review mainly focuses on the identification and classification of insects.

### *A. Image based insect identification and classification*

To maintain the diversity of species within the ecosystem, various scientists in the world including computer scientists have gained an interest in the field of biodiversity and out of 1.3 million known species on this earth, insects account for more than two thirds of these known species. For many years now, there have been different kinds of interactions between humans and insects and these interactions requires that insects are observed in close contact either by taking them to the lab or capturing their images. Several attempts have been made to create a method to perform insect identification accurately mainly because accurate insect identification requires great knowledge and experience on entomology [6]. A shortage of entomologists and the labour intensive process of collecting insect samples has led scientists trying to come up with methods that will do automatic insect identification and classifications.

Feng, Bhanu and Heraty [5] states that there is a clear need to introduce fast systems integrated with modern image processing and analysis algorithms to accelerate the process of image identification. Therefore they proposed the development of an automated moth species identification and retrieval system (SPIR) using computer vision and pattern recognition techniques. They used a probabilistic model that infers Semantically Related Visual (SRV) attributes from low-level visual features of moth images in the training set, where moth wings are segmented into information-rich patches from which the local features are extracted, and the SRV attributes are provided by human experts as ground-truth [5]. For the large amount of unlabeled test images in the database or added into the database later on, an automated identification process is evoked to translate the detected salient regions of low-level visual features on the moth wings into meaningful semantic SRV attributes. They also propose a novel network analysis based approach to explore and utilize the co-occurrence patterns of SRV attributes as contextual cues to improve individual attribute detection accuracy. Using a small set of labeled training images, the approach constructs a network with nodes representing the SRV attributes and weighted edges denoting the co-occurrence correlation [5]. To detect the co-occurrence patterns as communities in the network, they used a modularity maximization algorithm. The SRV was used in their system to record the visual and semantic properties of an image and to compare image similarity. They evaluated the approach in automated moth identification and attribute-based image retrieval and found that the performance of the system was improved by the SRV attribute representation and their co-occurrence patterns and that the power of their system was mainly affected by moth with high similarities in the visual properties [5].

Mayo and Watson [7] applied data mining techniques to effectively identify images of 774 live moth each belonging to different UK moth species. They extracted feature vectors from each of the moth images and used the machine learning toolkit WEKA to classify the moths by species using the feature vectors. The ML Toolkit was able to achieve a greater level of accuracy (85%) using support vector machines without manual specification of a region of interest at all [7]. They stated that the most important factor in the success of any machine learning-based image classification system is the features that are extracted. They captured the features on the moth's wings by calculating the moth's centroid by centering a square over the centroid and taking samples from 200 patches inside the main square [7]. The method was effective in capturing the wing patterns but disadvantaged smaller species that occupied less space in the image. They suggested that while the challenge of smaller species in the image affected their work, future work should be able to address the challenge when using the same techniques that they used in their work. In their work on image identification and analysis, they showed that data mining can be usefully applied to automatic identification of species.

Boniecki, Koszela, Boniecka, Weres, Zaborowicz, Kujawa, Majewski and Raba [8] investigated the possibility of using artificial neural networks (ANN) as a tool for classification, designed to identify apple orchard pests. They presented a classification neural model using optimized learning sets acquired on the basis of the information encoded in the form of digital images of selected pests. They used Neural modeling techniques, including digital image analysis to classify the pests [8]. The qualitative analysis of neural models produced indicated that multi-layered perceptron (MLP) neural network topology achieve the best classification ability [8]. They found that the use of artificial neural modeling and image analysis methods in the identification of apple pests turned out to be an appropriate way of effectively assisting the decision-making processes that occur during the production of apples [8].

Banerjee, Kiran, Murty and Venkateswarlu [9] presented an artificial neural networks method for classification and identification of anopheles mosquitos based on the internal transcribed spacer2 (ITS2) data of ribosomal DNA string. They implemented the method using two different multi-layered feed-forward neural network model forms, namely,

multi input single-output neural network (MISONN) and multi-input multi-output neural network (MIMONN) [9]. They employed a number of data sequences in varying sizes of different Anopheles malarial vectors and their corresponding species coding in the development of the neural network models. Their results demonstrates the efficiency of neural networks models in the extraction of information.

J. Wang, L. Ji, A. Liang, D. Yuan [10] applied a content based image retrieval (CBIR) method to the automatic identification of butterfly families because of its capacity for mass processing and operability. They conducted experiments with different features, feature weights and similarity matching algorithms were compared and found that data attributes such as species diversity, image quality and resolution affected system success the most, followed by features and match algorithms. They also showed that shape features are more important than colour or texture features in the identification of butterfly families [10].

Do, Harp and Norris [11] tried to bridge the gap between professional taxonomists and non-specialists by presenting a partially automated pattern recognition system that utilizes artificial neural networks (ANNs). They trained various neural networks to identify spider species using digital images of female transform [11]. They found that the neural networks were accurate most when identifying specimens in a hierarchical system.

Gassoumi, Prasad and Ellington [12] described a computer-vision-based system to recognize and classify insects as being harmful and non-harmful to the growth of cotton. They based the recognition and classification of the cotton insect upon a neural network approach. In their work, they showed that neural networks can be applied effectively to the classification of insects and suggested that to demonstrate the viability of using trained neural networks directly in the field, further work is required [12]. The concluded their work by speculating that in the future, field deployable systems would allow farmers to identify the type of insects that inhabit a specific ecosystem and choose which biological or pesticide to use as a control [12].

J. Wang, C. Lin, L. Ji, A. Liang [13] designed an automatic identification system for insect images at the order level. They designed several relative features according to the methods of digital image progressing, pattern recognition and the theory of taxonomy and used artificial neural networks and a support vector machine (SVM) as pattern recognition methods [13]. From the results of the tests they conducted, they found that the tests conducted using ANN performed very well as compared to the tests conducted using SVM. They also concluded that to improve their insect order identification system, they need to focus on feature extraction and design of newer and more effective features from the insect order [13].

There has been a number of systems designed and suggested that can be used to come up with an effective method for insect image identification and classification and most of the methods used suggests the automation of the whole process by using machine learning techniques such as artificial neural networks (ANN) and Support Vector Machine (SVM). ANN are information processing systems constructed and implemented to model the human brain and have the ability to learn using their experience [14]. Provided with a good number of samples, ANN are capable of generalizing to other samples they are yet to encounter [14]. The use of ANN in most of the image identification systems reviewed has shown that they can effectively be trained to efficiently identify and classify images of insects. The reviewed systems showed that there are a number of factors that affect the performance of the identification system such as environmental factors i.e. temperature, humidity etc., image quality and resolution, the diversity of species, features of species such as different sizes of captured species, system placement areas and so on. Therefore to design an effective and efficient FAW identification and classification system, most of the above factors will have to be considered.

### *B. Trap based insect identification and classification*

Insects can have both positive and negative impact on the livelihood of human beings hence the reason why they receive a lot of scientific attention. For instance bees play an important role in the pollination of plants hence one of the reasons why we have food from plants. Some insects are a nuisance to humans like the FAW which has had a negative impact on the supply of food where ever they are reported. The negative impact of insects has led to scientists developing methods that can be used to monitor their presence. One of the methods that has been used from time in memorial is the traditional method of traps. According to [15], traps that are developed for capturing insects are varied according to the purpose for trapping, the targeted insect and the habitat in which they are used. Traps are used for the general survey of insect diversity or for the detection of new invasions of insects in an area and may be used as direct control measures e.g. mass trapping, perimeter trapping or can be used for suppressing population buildup of insects [15]. There are many trap types which include interception traps commonly used for faunal surveys in ecological studies which are suspended nets with an invagination along the top leading to a collecting funnel, sticky commonly used for faunal surveys in agricultural studies which are either panels, cylinders or spheres covered with sticky materials that retain insects that fly onto the panel [15].

One of the commonly used traps in the monitoring and controlling of the FAW insect are pheromone traps which can be classified as a sticky trap owing to the sticky material that is sometimes used to trap the FAW once in the trap. Pheromone traps use pheromones to attract male insects. A pheromone is a chemical secreted by (usually) a female insect to attract males for mating which can travel by air very long distances and hence are very useful for monitoring insect presence [3]. The pheromone traps are useful for detection of early pest infestations, definition of areas of pest infestations, tracking the buildup of pest population and assisting in the decision making process for pest management [3,4]. Pheromone traps are of different designs manufactured by different manufacturers as shown by Fig. 1 and Fig. 2. Fig. 1 uses a sticky surface coated with a special non-drying glue which retains the insect once it flies onto it. The sticky trap has been used in Malaysia and found to be more effective in capturing small moths [16]. The Funnel pheromone trap shown in Fig. 2 has a funnel section and a bucket. It has a pheromone dispenser holder is located on top of the funnel under an umbrella type cover. Insects attracted to the trap fly around the pheromone dispenser until exhausted and then fall into the trap as shown. Insects find it impossible to fly out of the bucket and are eventually killed

by the pesticide inside the bucket. The effectiveness of the bucket trap was tested in the capturing of male moths in corn fields [17].

The trap method of insect identification and classification is the traditional method and has been used from time in memorial and has been found to be effective in the monitoring, mass trapping and control of insects. However good this method is, it has disadvantages such as it requires regular monitoring of and counting of insects. The traps if full needs to be emptied which requires going to the fields. The method if used as a monitoring and early warning system is prone to early and takes time, day's most of the time for the warning to be past to the relevant stakeholders hence delaying the decision making process. An automation of the counting process by the use of the latest technology we can go a long way in speeding up the early warning process at the same time speeding up the decision making process.

Fig. 1. Sticky surface trap [21].

Fig. 1. Bucket pheromone trap [22].

### *C. Why Artificial Neural Networks*

Work on Artificial neural networks from its inception has been motivated by the Human brain. The brain is a highly complex, nonlinear, and parallel computer. It has the capability to organize its structural constituents, known as neurons, so as to perform certain computations (e.g., pattern recognition, perception, and motor control) many times faster than the fastest digital computer in existence today [14]. A neural network is a processing device, either an algorithm or a hardware whose design was inspired by the human brain. The neural networks have the ability to learn by example which makes them very flexible and powerful. For neural networks, there is no need to devise an algorithm to perform a specific task, meaning there is no need to understand the internal mechanism of that task [14]. The computing world and the entire world has a lot to gain from neural networks. They can successfully be applied to solve many complex problems in the world which is the reason why scientists have taken advantage of them to try and solve many problems.

Phiri, Tie-Jun, Hui and Jameson [18] used machine learning techniques, adaptive neural-fuzzy inference system, fuzzy logic and artificial neural network to implement a multifactor authentication system through a technique of information fusion. They mined identity attributes using three corpora from social networks, a set of questionnaires and application forms from the various services offered both in the real and cyberspace then they composed an identity attribute model using the generated statistical information. Then they fused the identity attribute metric values classified as biometrics, device metrics and pseudo metrics at the score level using the proposed model using each of the machine learning techniques above. They stated that due to an increase in the number of services going online, there has been an increase in the number of cyber related cases [18] therefore the proposed system help in curbing the cases.

Phiri and Agbinya [19] used artificial intelligence and biometrics to try and improve the security and privacy of internet users and service providers in a transparent, reliable and efficient way. They concluded by stating that using artificial neural networks together with biometrics forms a very effective and intelligent authentication system in identity management systems[19].

Zhou, Jiang, Yang and Chen [20] used a neural network ensemble to identify Lung cancer cells. An ensemble is a joining of several Artificial Neural Networks to solve a problem [20]. They proposed a diagnosis procedure called Neural Ensemble-Based Detection (NED) which utilizes an ANN ensemble to identify lung cancer cells in the images of the specimens of needle biopsies gotten from subject bodies. They built a two layer ensemble with the first layer being used to judge whether a cell is normal with only two outputs normal cells or cancer cells. The second layer is used to deal with cells judged as cancer cells. The second layer had five outputs four of which are different types of lung cancer cells [20]. The techniques used in this paper enabled the ensemble to achieve a high rate of overall identification and a low rate of identification which is cardinal in saving lives due to misdiagnosis [20].

Artificial neural networks have successfully been applied in so many areas in the world to solve complex problems as can be seen from the work done by the reviewed work. Their ability to generalize and learn makes them a capable tool in

solving many complex problems. In addition they are capable of handling large amounts of data. Hence this makes ANN as the preferred strategy to be used in this study.

## III. PROPOSED SYSTEM

The proposed study will be designed to be mainly quantitative in nature. The study will range from modification of bucket pheromone traps and automating them with sensors for data collection to the training of a convolutional neural network to the development of web and mobile applications to be used by stakeholders for data analysis and reporting.

### *A. Architecture*

Fig. 3 shows a simple proposed model of how the data will flow. First the data will be collected at the trap site by the modified and automated pheromone traps. Depending on the cellular network connectivity at that particular time, the collected data will be sent to the cloud server for processing. Once received, the image identification system will identify and classify the image as FAW moth and load the data in the database. Once the data is loaded onto the database, the relevant stakeholders can access it via the web or mobile interfaces that we will provide.


Fig. 3. Proposed model

### *B. Data collection*

The main instruments that will be used for data collection will be the modified traps automated with sensors and a mobile application that can be used for data collection and as a data analysis tool. The success of this project will be dependent among other things on the successful modification of the pheromone traps. The modification will involve closing two of the four trap entry points on some traps and closing three of the four entry points one other traps. The main reason for the reduction of entry points is to try and improve the accuracy and efficiency of FAW image capturing. The study will look at which modification will be most suitable and accurate between the two or three entry point closed trap. We employ the use of single board computers such as the Raspberry pi or Arduino which will be programmed to take all the necessary readings at each trap. Single board computers are credit card sized computers with I/O components, memory, a microprocessor and other necessary components on a single board. They are ideal devices to be used because of their small size which will enable them to be properly fit inside the trap, they are able to be programmed to support a good number of sensors and their low power consumption means that only a small amount of power supply will be required. Proper placement of high resolution cameras on the single board computer so that they are able to get quality and high resolution images will be another cardinal modification that will be done. The cameras will be required to be placed in a way that they are able to capture clear, quality and high resolution images of moths. Hence different angles will be tested in the trap to determine the best angle. Among sensors that will be included on the board will be the humidity sensor which will be required to take the current atmospheric humidity, temperature sensor which will be required to collect the current environmental temperature around the site and also if required the proximity sensors or motion detection sensors to detect the entry of the moth in the trap so that the cameras are triggered to capture the image of the moth. The traps will also be automated with GPS and 3G/4G capabilities. The GPS will be required to get the global position of the trap every time the image is captured and data is about to be sent to the cloud. Of course to reduce on unnecessary processing of the single board computer, we will consider the hard coding of the GPS coordinates for each trap if it's viable. Cellular network connectivity will be required to provide the single board computer with the necessary internet connectivity required to enable after automatic data collection, the data is pushed to the cloud server within seconds of collection. Also the colour of the traps will have to change to white to provide proper illumination to the camera owing to the colour of the FAW moth being brownish in colour which might hinder the quality of the images captured and hence reduce the chances of the image being identified by the image identification and classification system. The traps will also be supplied with small solar panels or a portable rechargeable batteries to supply the single board computer with the required power. To reduce on the amount of power that is required at each trap, the amount of processing that the single board computer will be required to do will be limited to the pushing of the readings and the image to the cloud server.

### *C. ANN Training*

The proposed study is going to use supervised learning to train the ANN with a training sample size of more than 100 images of FAW. The sample images will be taken from the field at the trap site at different angles in the trap and outside the trap. We intend to use convolutional neural networks since they do well when it comes to image classification and are easier to simulate and implement. Convolutional neural networks works just like a regular neural network except that they have a convolution layer at the start of a network. A convolutional network breaks up the image into tiles instead of getting the image as an array. From the tiles, the network then tries to predict what the tile is and eventually tries to predict what the image is. We will train the network until it's able to recognize the FAW image with an accuracy of 75% and above. We will first simulate the network in Matlab until we achieve the desired identification accuracy of more than 75%. We will then write a deep learning program using Tensorflow an open source deep learning framework created by google. Tensorflow will be ideal for the implementation of the algorithm because it's able to give us granular control over each node meaning we are able to adjust the weights thereby giving us good performance. It also has a lot of

libraries that we can use when it comes to deep learning and a good support community. The program that will be built will also be trained until it's able to achieve an accuracy of more than 75% then we will integrate the program in our proposed system and test and monitor its performance. We intend to bench mark the performance of the built program against google vision to see if it's able to give better results than google vision. The built program will site as a standalone application from the web interface that the study will develop.

### *D. Data analysis*

The local weather readings and images captured at the modified trap site will be sent to a cloud server where the identification and classification of the captured image will be done. Among the parameters that the site will send include the image, trap name, farm name, humidity, temperature, GPS coordinates, type of the pheromone chemical being used on the trap and so on. Once received at the cloud, the image will have to be identified by an image identification system to determine what insect species it is. The system will employ ANN techniques in the identification process. The ANN will have to be trained to identify and classify the FAW moth. After the identification process, the data received together with the image are recorded in the system database which can be accessed using a web or mobile application for reporting or analysis by stakeholders. The logged data will be accessible via a web based application which will be designed. The system will have features that will enable intended users to analyze the generated data and produce appropriate reports. In addition the system should be able to provide a detailed view of the captured image and at what trap site and in which province the image was captured using the GIS technology. The system will be developed using open software tools such as php, python, MySQL, JavaScript, css and so on. A mobile application will be developed on both android and IOs with the priority being the android platform. The mobile application in addition to being a data collection tool, will play the role of data analysis just like the web interface.

## IV. CONCLUSION

The impact of FAW on the global food security can never be underrated. It is an issue that affects all of us in the world and if we do not do anything about it, the world will starve. This has led FAO in collaboration with different regional organizations in Africa such as SADC, ECOWAS and so on to classify the FAW outbreak as a threat to food and nutrition security likely to affect the livelihoods of small scale farmers and commercial farmers if not controlled. The main problem that FAO and other stakeholders involved in the fight against FAW face is the monitoring and early warning of the pest occurrence. The monitoring is usually manual which involves weekly field visits hence hindering the monitoring process. This is the reason we proposed a model that will make use of cheap technology and do automatic data collection in the field and send the data to the cloud server where data will be easily accessible via the web interface or mobile application for analysis and reporting. The proposed system using ANN techniques will provide among other advantages like increased accuracy of monitoring at a lower costs, fewer field visits and more efficient pest control, optimized pesticide application timings, near real-time insight into a pest situation in the field, easy data sharing with advisers, effortless pest monitoring of vast area with fewer resources, automated and precise data gathering, short data collection intervals and securely archived data. The system will act as a great source of data to verify and develop pest-related models. The system once successfully implemented will provide stakeholders a platform that they can use to get the necessary information on FAW pest occurrences in the country and be able to make informed and viable decisions.

## REFERENCES


[1] "Four threats to global food security and what we can do about them," The conversation, Available online: https://theconversation.com/four-threats-to-global-food-security-and-what-we-can-do-about-them.

[2] Day, Roger & Abrahams, Phil & Bateman, Melanie & Beale, Tim & Clottey, Victor & Cock, M.J.W. & Colmenarez, Yelitza & Corniani, Natália & Early, Regan & Godwin, Julien & Gomez, Jose & González-Moreno, Pablo & Murphy, Sean & O. Mensah, Birgitta & Phiri, Noah & Pratt, Corin & Silvestri, Silvia & Witt, Arne, "Fall Armyworm: Impacts and Implications for Africa," Outlooks on Pest Management, Vol. 28, pp. 196-201, October 2017.

[3] J.E. Huesing, B.M. Prasanna, D.McGrath, P.Chinwada, P.Jepson, J.L. Capinera, "Fall Armyworm in Africa: A Guide for Integrated Pest Management," 1st ed. 2018.

[4] P. Abrahams; M. Bateman; T. Beale; V. Clottey; M. Cock; Y. Colmenarez; N. Corniani; R. Day, R. Early, J. Godwin, J. Gomez, P. G. Moreno; S.T. Murphy; B.O. Mensah, N. Phiri, C. Pratt; G. Richards, S. Silvestri, A. Witt, "Fall Armyworm: Impacts and Implications for Africa, Evidence Note (2)," September 2018.

[5] L. Feng, B. Bhanu, J. Heraty, "A software system for automated identification and retrieval of moth," Pattern Recognition, Vol. 51, pp. 225-241, September 2016.

[6] S. N. Asiah Hassan, N. N. S. Abdul Rahman, Z. Zaw Htike, S. Lei Win, "Vision Based Entomology: A Survey," International Journal of Computer Science & Engineering Survey, Vol. 5, pp. 19-32, January 2014.

[7] M. Mayo, A. T. Watson, "Automatic species identification of live moths," Knowledge Based Systems, Vol. 20, pp. 195–202, February 2007.

[8] P. Boniecki, K. Koszela, H. P. Boniecka, J. Weres, M. Zaborowicz, S. Kujawa, A. Majewski, B. Raba, "Neural identification of selected apple pests," Computers and Electronics in Agriculture Systems, Vol. 110, pp. 9-16, September 2015.

[9] A. K. Banerjee, K. Kiran, U. Murty, C. Venkateswarlu, "Classification and Identification of mosquito species using artificial neural networks," Computational Biology and Chemistry, Vol. 32, pp. 442–447, June 2008.

[10] J. Wang, L. Ji, A. Liang, D. Yuan, "The identification of buttery families using content-based image retrieval," Biosystems Engineering, Vol. 111, pp. 24-32, 2012.

[11] M. Do, J. Harp, K. Norris, ", A test of a pattern recognition system for identification of spiders," Bulletin of Entomological Research, Vol. 89, pp. 217-224, 1999.

[12] H. Gassoumi, N.R, N.R. Prasad, J.J. Ellington, "Neural Network-Based Approach for Insect Classification in Cotton Ecosystems, January 2000.

[13] J. Wang, C. Lin, L. Ji, A. Liang, "A new automatic identification system of insect images at the order level," Knowledge-Based Systems, Vol. 33, pp. 102-110, March 2012.

[14] M. Negnevistky, "Artificial Neural Networks" in Artificial Intelligence, a guide to intelligent systems, 2nd ed. Harlow, England: Pearson Education limited, 2005, Ch. 6, pp. 165–216.

[15] D. Epsky, Nancy, Morrill, Wendell, Mankin, Richard, "Traps for Capturing Insects" in Encyclopedia of Entomology, 2008, Ch. 10, pp. 3887-3901. Available online: https://www.researchgate.net/publication/303661972_Traps_for_Capturing_Insects

[16] Ahmad, S. Nurulhidayah, Kamarudin, Norman. (2011), "Pheromone Trapping in Controlling Key Insect Pests: Progress and Prospects," Oil Palm Bulletin.

[17] S. Guerrero, J. Brambila, R.L. Meagher, "Efficacies of Four Pheromone-Baited Traps in Capturing Male Helicoverpa

(Lepidoptera: Noctuidae) Moths in Northern Florida," Florida Entomologist, Vol. 97(4), pp. 1671-1678, December 2014.
[18] J. Phiri, Z. Tie-Jun, Z.C. Hui, M. Jameson, "Using Artificial Intelligence Techniques to Implement a Multifactor Authentication System," International Journal of Computational Intelligence Systems, Vol. 4, pp. 420-430, June 2011.
[19] J. Phiri, J.I Agbinya, "Fusion of multi-modal credentials for authentication in digital identity management systems," International Conference on Wireless Broadband and Ultra-Wideband Communications, Vol. 34, pp. 20, August 2007.
[20] Z. Zhou, Y. Jiang, Y. Yang and S. Chen, "Lung cancer cell identification based on artificial neural network ensembles," Artificial Intelligence in Medicine, Vol. 24, pp. 25-36, April 2001.
[21] Source:http://web.entomology.cornell.edu/shelton/swede-midge/images/DETECTION/trap2large.jpg
[22] Source:https://4.imimg.com/data4/BY/BP/MY-5420398/pheromone-trap-250x250.jpg

# An Application of Machine Learning Algorithms in Automated Identification and Capturing of Fall Armyworm (FAW) Moths in the Field

Simon H. Chiwamba
*Department of Computer Science*
*The University of Zambia*
Lusaka, Zambia
shchiwamba@yahoo.com

Jackson Phiri
*Department of Computer Science*
*The University of Zambia*
Lusaka, Zambia
jackson.phiri@cs.unza.zm

Philip O. Y. Nkunika
*Department of Biological Sciences*
*The University of Zambia*
Lusaka, Zambia
pnkunika@unza.zm

Mayumbo Nyirenda
*Department of Computer Science*
*The University of Zambia*
Lusaka, Zambia
mayumbo.nyirenda@cs.unza.zm

Monica M. Kabemba
*Department of Computer Science*
*The University of Zambia*
Lusaka, Zambia
monica.kalumbilo@cs.unza.zm Monica

Philemon Sohati
*Department of Plant Science*
*The University of Zambia*
Lusaka, Zambia
phsohati@googlemail.com

***Abstract*— As farmers' struggle with unpredictable rainfall patterns, the horror of emerging crop pests that will likely affect the quality and quantity of their harvest does not spare them. Therefore, it is very important to protect the crop by monitoring these pests. In the 2016/2017 farming season, the Southern African Development Community (SADC) food security was threaten by the outbreak of the fall armyworm (FAW) despite the normal rainfall. The FAW affected several crops including maize, the staple food in the DRC, Botswana, Malawi, Namibia, Swaziland, South Africa, Zambia and Zimbabwe. In Zambia alone, attempts to control the fall armyworms that had affected approximately 130,000 hectares of crops costed about US$ 3 million. This led SADC in collaboration with Food and Agriculture Organization (FAO) to classify the outbreak as a threat to food and nutrition security, and livelihoods of smallholder farmers in the region. This has posed a number of challenges on the stakeholders such as the University of Zambia (UNZA) School of Agricultural Sciences and Department of Biological Sciences to come up with more precise and proactive measures for monitoring and controlling the pest. This paper proposes a study to develop an automated system for identifying and capturing images of the FAW moths in the field using a supervised machine learning (ML) technique called Convolutional Neural Network (CNN). The proposed study will aim to address current challenges that the UNZA School of Agricultural Sciences and Department of Biological Sciences are facing in monitoring the FAW using the pheromone traps. The pheromone trap will be modified to include a raspberry PI, vision sensor and motion sensor to be used for data collection. The system will further integrate GPRS and 3G/4G connectivity to allow automatic data collection without field visits while GPS receiver will enable automatic position awareness of the trap. The data collected will automatically be sent to the cloud servers for immediate analysis and access by the stakeholders.**

***Keywords—Fall Armyworm, Pest Monitoring, Pheromone Trap, Machine Learning, Convolutional Neural Network, Insect Identification, Insect Capture.***

## I. INTRODUCTION

In January 2016, African recorded an outbreak of a new pest called Fall Armyworm (FAW) [1, 2] and by August, 2017 the pest had spread to cover over 75% of the continent as shown in the distribution map in fig .1. According to [2], the FAW is a moth that is indigenous throughout America where it is widely agreed to be one of the most damaging crop pests. It attacks more than 80 different plant species, including maize, a major staple food in sub-Saharan Africa on which more than 200 million people depend. According to [3, 4], the several reports received from the West and Central African countries in 2016 were a clear indication of a possible emergence of a new regional problem. The outbreak of the FAW can be one of the most difficult to control especially when it has reached an advanced larval development stage. It can cause extensive crop losses of up to 73% depending on existing conditions. The FAW affects the late-planted fields and late maturing hybrids more. In 2017, Zambia alone incurred about US$ 3 million in an attempt to control the FAW infestation that affected approximately 130,000 hectares of crops [1, 2].

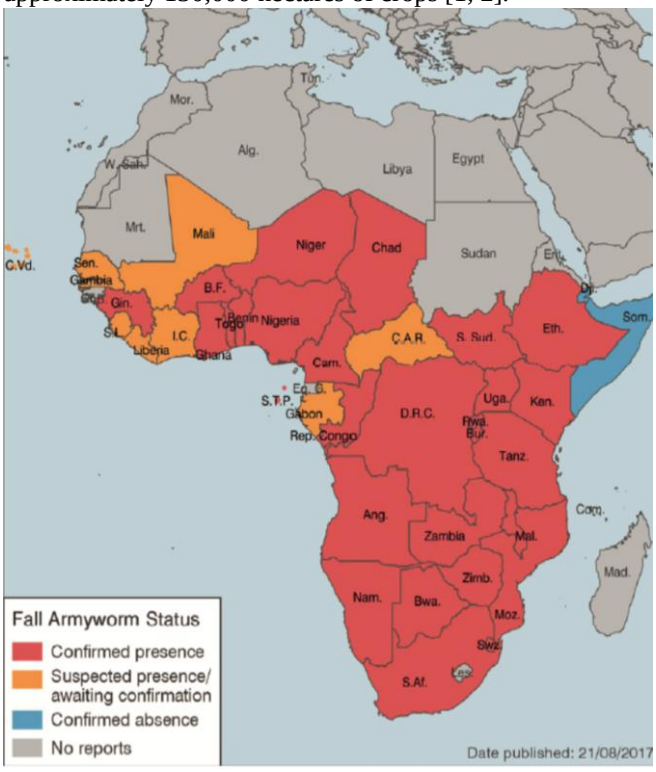



Fig. 1.FAW distribution in Africa as at August 2017. Source [2].

As explained by [4], the FAW has a life cycle of about 30 days (at a daily temperature of ~28°C) during the warm summer months and may extend to 60-90 days in cooler temperatures. The FAW are migratory in nature and the moths can fly for over 100km per night. The FAW arrive when environmental conditions allow and may have as few as one (1) generation before they become locally extinct. The female moths can lay an average of 1500 eggs and can produce multiple generations very quickly without pause in tropical environments [1, 4]. Therefore, tackling the threat of the FAW pest at an early stage and avoiding economic hardship requires quick coordinated actions, massive awareness campaigns, scientific innovations and multi-institutional collaborations as demonstrated by Southern African Development Community (SADC) and Food and Agriculture Organization (FAO) [1].

The current monitoring and early warning system used by the University of Zambia (UNZA) School of Agricultural Sciences and Department of Biological Sciences places pheromone traps in different fields to lure the male FAW moths and kill them. After a week, the Field inspectors go in these fields to inspect and count the number of moths captured. The data is recorded on data collection sheets which are later used as source document when preparing reports. This process is labour intensive due to the field visits and manual counting of the moths. It is also prone to errors due to the manual counting and physical recording of the data on the forms. The other aspect is that, the field inspectors may understate and overstate the number of moths sending force positives and negatives to stakeholders. These challenges make it difficult for stakeholders to accurately monitor, plan and provide near real-time response to the infestation of the deadly crop pest that can affect Zambia's food security.

The proposed system will reduce the field visits; shorten data collection intervals; provide near real-time insight into pest situation in the field; enable an effortless pest monitoring of vast areas with fewer resources; enable easy data sharing with advisers. This will be achieved by automating the identification and capturing of the FAW moths in the field using a supervised machine learning (ML) technique called Convolutional Neural Network (CNN) and modifying the pheromone trap to include a raspberry PI, vision sensor and motion sensor. The system will further integrate GPRS and 3G/4G connectivity to allow automatic data collection without field visits. It will also incorporate the GPS receiver to enable automatic position awareness of the trap. The data collected will automatically be sent to the cloud servers for immediate analysis and access to the intended users. The system will also be used as the primary source of data for the Cloud based Integrated Pest Management System (IPMS) to enhance pest monitoring and provide powerful analytics that will significantly increase the quality and viability of decision-making.

This paper is structured as follows: Section II devoted to literature review. Section III describes the proposed system while Section IV explains the methodology and finally, Section V provides some conclusion.

## II. Literature Review

In this literature section, we will discuss some methods and algorithms that have been used previously to monitor and identify pests.

### *A. Pheromone Traps*

[2] says that the health and well-being of humans can be impacted in a positive and negative sense by insect if not monitored and controlled. This can be confirmed by the position taken by SADC in collaboration with FAO to recognize the FAW (spodeoptera frugiperda) as a threat to food and nutrition security, and livelihoods of smallholder farmers [1]. As reported by [1] and further amplified by [3], monitoring systems are at the core of pest control and require less labour intensive, faster, inexpensive, and error free counting systems that should guide decision making in the management of the FAW. According to [5], the monitoring of pests such as the FAW has received a lot of scientific attention which has led to the development and deployment of innovative scientific tools. These tools require proper installation and maintenance in order to ensure accuracy and reliability of the data.

One of most prominent and successful scientific tool being used for pest management programmes worldwide is the Pheromone trap [5]. It is a type of insect trap that uses pheromones to lure insects. According to [5, 6, 7], this tool can be used to:

- Detect early pest infestations such as the first occurrence of migratory pests.
- Define areas of pest infestations.
- Track the build-up of a pest population.
- Help in decision making for pest management.

The Pheromone trap comes in different designs including the ones shown in fig. 2 and fig. 3. Fig. 2 is a pheromone trap that employs a sticky surface to retain the attracted insect. The insect fly into the trap and stick to the surface coated with a special type of non-drying glue. According to [5], the sticky type trap has been found to be effective in capturing small moths of the oil palm in Malaysia.

Fig. 2. Pheromone trap that employs a sticky surface. Source [18].

The Funnel (green lid/yellow funnel/transparent bucket) pheromone trap shown in fig. 3 consists of a funnel section and a bucket. A central pheromone dispenser holder is located on top of the funnel under an umbrella type cover. The trap is used in combination with pheromone dispensers.

Insects attracted to the trap fly around the pheromone dispenser until exhausted and then fall into the trap (bucket). Insects find it impossible to fly out of the bucket and are eventually killed by the pesticide inside the bucket. This type has been found to be more effective in capturing moths in corn fields [7].


Fig. 3. Funnel (green lid/yellow funnel/transparent bucket) pheromone trap. Source [19].

In Malaysia, pheromone trap have been used for monitoring, mass trapping and mating disruption of the FAW which affects oil palm the most important commodity crop. On the other hand, it has been found to be the best means of deciding on insecticide application in maize fields to control the FAW with an effectiveness of 90% larval mortality in Brazil [5, 6].

*B. Object detection and recognition*

[8] Acknowledges that automated entomology is required to meet a growing biological demand to count insects and that it has been a challenge for both the computer scientists and biologists. This has resulted in a lot of scientific attention which has led to massive research work as indicated in the work done by [8]. In this review, we will try to focus on techniques that support field based settings defined as "a capture of insects directly in cultivated fields, without any particular constraints to the images capture system" by [8]. We also define object detection as "a procedure of determining the instance of the class to which the object belongs and estimating the location of the object by outputting the bounding box around the object" [9]. Object detection can also be defined as the act of noticing or discovering something. And object recognition is a computer vision technique for identifying objects in images or videos and is a key output of deep learning and machine learning algorithms.

[9] explained the role of deep learning techniques based on CNN for object detection and assessed the techniques, frameworks, services and state-of-the-art object detection systems that use deep learning. The results where that, the deep neural architectures handles complex models efficiently than shallow networks while CNNs are less accurate for smaller data. CNNs showed significant record breaking accuracy on the large image datasets. CNNs also require large amount of labeled datasets to perform computer vision related tasks such as detection and classification.

[10] developed an insect recognition system using an advanced multiple task sparse representation and multiple-kernel learning (MKL) techniques. The system was able to combine multiple features of insect species to enhance the recognition performance. It was adapted to represent insect images so that raw features such as color, shape, and texture could be well quantified. When analyzed on 24 common pest species of field crops, the technique performed well on the classification of insect species, and outperformed the state-of-the-art methods of the generic insect categorization [10].

In UK, a study was conducted by [11] using a machine learning toolkit called WEKA to classify 774 live individual moths into 35 different UK species and the results where 85% accurate.

In [12], an algorithm called Simultaneous Detection and Segmentation (SDS) was proposed to detect all instances of a category in an image and correctly mark each instance that belonged to that category. The algorithm required segmentation of individual object instances and bounding box detection. It also used CNN to extract features on each region and trained a Support Vector Machine (SVM) on top of the CNN features to assign a score for each category to each candidate. The results of the study showed a 5 point boost (10% relative) over state-of-the-art on semantic segmentation, and state-of-the-art performance in object detection.

In [13], an automatic detection pipeline based on deep learning for identifying and counting pests images taken inside field traps was proposed. It used a sliding window based detection pipeline and applied CNN on the image patches at different locations to determine the probability of containing a specific pest type. Image patches were then filtered by non-maximum suppression and thresholding, according to their locations and associated confidences, to produce the final detections. Qualitative and quantitative experiments demonstrated the effectiveness of the proposed method on a codling moth dataset.

In [14], a new approach to object detection was presented that uses frame object detection as a regression problem to spatially separate bounding boxes and associated class probabilities. It applies a single neural network to predict bounding boxes and class probabilities directly from full images in one evaluation. The model is called You Only Look Once (YOLO) and it was able to process images in real-time at 45 frames per second while Fast YOLO had the ability to process 155 frames per second and achieve double the mAP of other real-time detectors. According to [14, 15], other methods of detection including DPM and Faster RCNN with ResNet and SSD were outperformed by YOLO.

[16] proposed a system that included a pan tilted web camera with zoom, ultraviolet lights and raspberry pi kit with WiFi connection. The autonomous monitoring system was based on a low-cost image sensor that captured and sent images of the trap contents to a control station. The ultraviolet light was attached to a square shaped yellow board where the insect got attached. Based on the time scheduled program on the raspberry pi, the camera automatically captured the image and sent an MMS to the scientist through GPRS server. In addition, the user was also able take an image without disturbing or killing the pest.

In the work done by [17], an autonomous monitoring system based on a low-cost battery-powered wireless image sensors that is able to capture and send images of the trap contents to a remote control station with the periodicity demanded by the trapping application was proposed. The proposed system was meant to cover large areas with very low energy consumption. The images delivered by image sensors were time-stamped and processed in the control station to get the number of individuals found at each trap. All the information was sent to the control station using WiFi, WiMax, 3G or 4G for storage. The wireless sensor architecture included the CC1110F32 SoC with an 8051 enhanced microcontroller unit (MCU), the CC1101 radio transceiver module. The other set of peripherals included were a 128-bit AES security coprocessor, one USB 2.0 interface, two USARTs, one I2S interface, three 8 bit timers, one 16 bit timer, 7–12 bits ADC, and 21 GPIO pins. This system proved that wireless image sensor technology can leverage the deployment of efficient pest monitoring systems with a considerably cost reduction and near zero maintenance due to the large operational life of the proposed image sensors.

## III. Proposed System

We propose an automated system for identifying and capturing FAW moths in the field using a supervised ML technique called CNN, a modified funnel/bucket pheromone trap and raspberry PI equipped with vision and motion sensors as shown in our model in fig 4.

ML is a subfield of soft computing within computer science that studies the design of algorithms that can learn. Within ML, you find a subfield called deep learning that is inspired by artificial neural networks. Inside deep learning, there is a specific kind of neural network called the convolutional neural network, also commonly referred to as CNN or ConvNet. It's a multi-layer perceptron (MLPs) and by nature a "feed-forward" because information flows right through the model. In CNN, an input image is passed through the network layers to determine its class. As mentioned earlier, the network has multiple layers with each layer learning to detect different features of an image. For each training image, filters are applied at different resolution at each layer and a convolved image is outputted. The convolved image is used as an input to the next layer and so on. The filters may include simple features such as brightness and edges. The filters may increase in complexity as the object progress from layer to layer. Our choice of CNN as the technique for identifying the FAW moth is due to the proven effectiveness on prediction problem involving image data.

The raspberry is a credit-card-sized single board computer (SBC) developed in the UK by the Raspberry PI foundation. A SBC is credit-card-sized complete computer as mentioned earlier with I/O components, memory, a microprocessor and other features required for a functional computer on a single circuit board. It has General Purpose Input and Output pins that can used to communicate with sensor other than the usual computer peripherals. It is an ideal devices for this kind of project because it can easily be fitted on top of the trap cover (lid). It is also programmable and can support a good number of sensors including vision, motion, temperature and humidity. The device is a low energy consumer, therefore, only a small amount of power supply will be required for it to run. The power will also be supplemented by a solar panel that will recharge the battery. We will install an operating system and the neccesssary algorthims for identifying and capturing of the FAW moths on this device.

The further modification will involve installing a vision sensor on the top cover (lip) next to the lure holder at an angle as shown in the proposed model in fig. 4. This will give the vision sensor an eye view of the funnel path as shown fig. 5. The placement of the vision sensor on the top cover (lid) next to the lure holder is also meant to allow enough natural light and get a clear view of the funnel path as shown in fig. 5. This will further allow the vision sensor to get a high resolution and quality image of the FAW moth.

Thereafter, a motion sensor will be installed on the funnel path to detect motion of any insect falling into the bucket. The detected motion will be used to trigger the camera to get a picture of the insect or any object in motion before it exists the funnel path and reaches the base of the bucket. The motion sensor will be placed slightly at the beginning of the funnel but lower than the lure dispenser holder in order to give enough time to the camera to take a picture. The placement will also help in counting the contents of the pheromone trap. The system will also be able to classify the FAW moth from other insects and objects therefore enhancing the monitoring at field level.

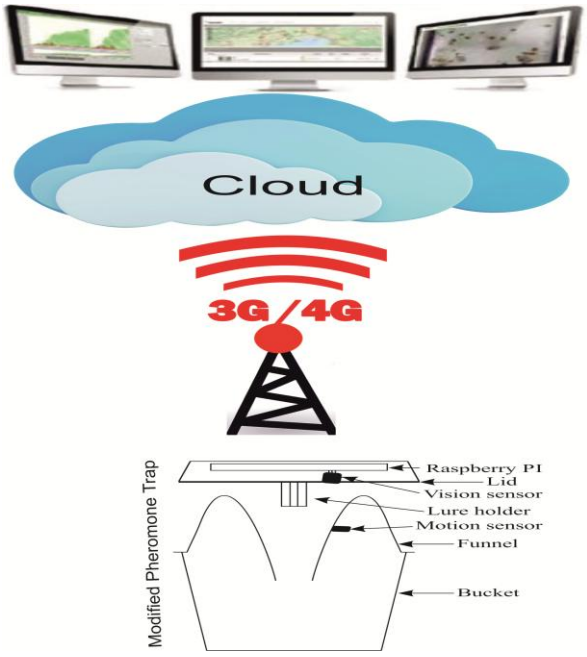



Fig. 4. Proposed model

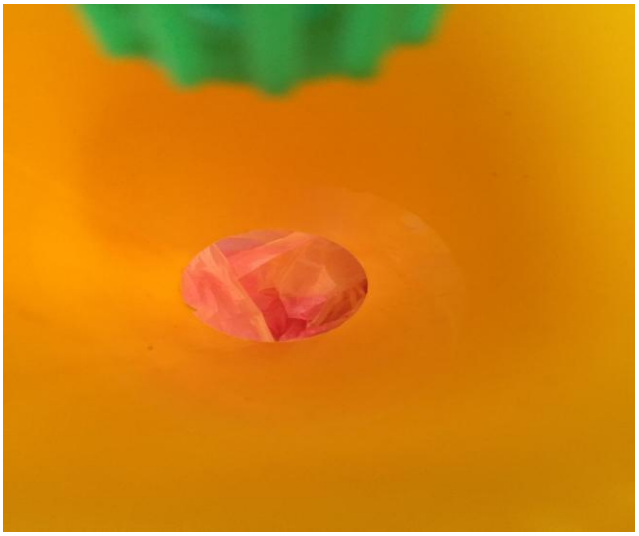


Fig. 5. Funnel Path taken using a vision sensor place next to the lure holder.

As mentioned earlier, the system will use CNN to identify and capture FAW moth that will be lured into the trap. The image will be stored on the memory card installed on the raspberry first, thereafter, sent to cloud server for further analytics whenever the connection is available. The system will use GPRS and 3G/4G to connect and send images to the cloud.

## IV. Methodology

In our proposed study, we shall pay particular attention to the image capture settings and the context of image acquisition. The acquisition mode has greater influence on the techniques used for the identification of the FAW moth, but also the context of application of such algorithms [8]. They are several analytical dimensions to be considered and the first is about acquisition conditions. This deals with both the way insect / pest is captured and the way the images are produced, that is the capture tools and the overall protocol. Our target insect in this study is a live FAW moth in motion. The motion can either be flying in or crawling on the funnel path.

The pose will be the second constraint. This is how much variating is allowed on the point of view on the insect. A constrained pose setting specifies what must be seen, for instance, only the front side or the back side of FAW moth. An unconstrained pose setting, in the contrary, induces variability in orientation of the individual and therefore in their apparent shapes, details and colours [8]. Next, closely related to the pose, the part or area of the individual is to be considered. Some insects are better recognized from their wings as others from their entire body. The FAW moth being considered under this study has very unique body features as shown in fig. 6. [20] demonstrated that these unique features can be used to identify and recognize an object. Our vision sensor will be placed on the top cover (lid) in order to give it a clear view of the FAW moth. It will also allow enough natural light for proper illumination and get the back side of FAW moth as it falls into the bucket.

The other settings include the Lab-based setting that is mainly used by entomologists in the lab to inspect them and to identify insect. Our focus in this study will be biased towards field-based setting is defined by [8]. Our pheromone traps will be placed in the field and collect data in real time.

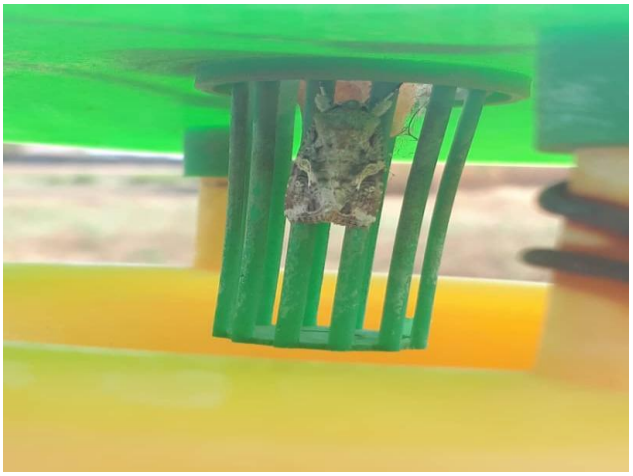


Fig. 6. Lured FAW moth.

The live FAW moth in motion will be captured using the vision sensor attached to the raspberry pi in real time. The image capture will be triggered by a motion sensor. The captured image will be classified using CNN and stored in a specific folder on the memory card installed on the raspberry pi. Thereafter, the system will check for connective to the cloud server and if connected, the image will be sent. The connection to the cloud will be via GPRS and 3G/4G. The cloud server will be used for granular identification and classification of the captured image using CNN. The data sent will include among others the image, trap name, farm name, humidity, temperature, GPS coordinates, and type of the pheromone chemical being used on the trap and so on.

To train our CNN, we shall undertake the following steps.

### *A. Step 1.Training set*

The training data set for the CNN will come from the traps currently installed at Liempa farm within the UNZA. We will take images of the FAW moth from different angles to stimulate different poses. Our target images will be 100.

### *B. Step 2. Creating and configuring the CNN.*

We will use MATLAB to create, configure and simulate our CNN.

### *C. Step 3. Training the CNN.*

We will use the data collected in step 1 to training the CNN. Our training options will include max epochs and learning rate.

### *D. Step 4. Determining accuracy and validity of the results.*

We shall continue to change the training options and network configurations till we achieve a 0.75 accurate rate.

## V. Conclusion

In this paper, we reviewed literature on pheromone traps and object detection algorithms. Thereafter, we proposed an automated system for identifying and capturing of FAW moths in motion using CNN, a modified pheromone trap and raspberry PI equipped with vision and motion sensors. We then looked at the methodology and gave details on how our CNN was going to be simulated.

## REFERENCES


[1] B. Wawa,J. Mollins, W. Njoroge, R. Nandelenga, "Multi-pronged approach – key for effectively defeating Fall Armyworm in Africa." Internet:http://www.fao.org/africa/news/detail-news/en/c/884224/, May. 02, 2017 [Oct. 02, 2018].

[2] R. Day, P. Abrahams,M. Bateman,T.Beale, V.r Clottey,M. Cock,Y. Colmenarez,N. Corniani,R.Early,J. Godwin,J. Gomez,P. G. Moreno,S. T. Murphy,B. Oppong-Mensah,N. Phiri, C. Pratt, S. Silvestri and A. Witt," Fall Armyworm: Impacts and Implications for Africa,"Outlooks on Pest Management, pp.196-201, Oct, 2017, DOI: 10.1564/v28_oct_02.

[3] G. Goergen, P. L. Kumar , S. B. Sankung, A. Togola, M. Tamò," First Report of Outbreaks of the Fall Armyworm Spodoptera frugiperda (J E Smith) (Lepidoptera, Noctuidae), abNew Alien Invasive Pest in West and Central Africa,"PLoS ONE, vol.11, no.10, Oct. 2016, doi:10.1371/journal.pone.0165632

[4] Fall Armyworm in Africa: A Guide for Integrated Pest Management, 1st ed., CIMMYT, Mexico, 2018, pp. 44–60.

[5] S. N.Ahmad and N. Kamarudin,"Pheromone Trapping in Controlling Key Insect Pests: Progress and Prospects,"Oil Palm Bulletin vol. 62, pp. 12-24,May 2011.

[6] I. Cruz, M. de Lourdes Corrêa Figueiredo, R. Braga da Silva, I. Fernandes da Silva, C. de Souza Paula, and J. E. Foster,"Using Sex Pheromone Traps in the Decision-Making Process for Pesticide Application against Fall Armyworm (Spodoptera frugiperda [Smith][Lepidoptera: Noctuidae]) Larvae in Maize,"IJPM, vol. 58, no. 1, pp. 83-90, Mar. 2012,doi: 10.1080/09670874.2012.655702

[7] S. Guerrero, J. Brambila and R. L. Meagher,"Efficacies of Four Pheromone-Baited Traps in Capturing Male Helicoverpa (Lepidoptera: Noctuidae) Moths in Northern Florida,"BioOne, vol. 97, no. 4, pp. 1671-1678,Dec. 2014, doi: 10.1653/024.097.0441

[8] M. Martineau, D. Conte, R. Raveaux, I. Arnault, D. Munier and G. Venturini,"Survey on image-based insect classification. Pattern Recognition, Elsevier, vol. 65, pp.273 - 284,jan. 2017,doi:10.1016/j.patcog.2016.12.020.

[9] A. R. Pathaka, M. Pandeya and S. Rautaraya,"Application of Deep Learning for Object Detection," ICCIDS, vol.132, pp.1706–1717, 2018, doi:10.1016/j.procs.2018.05.14

[10] C. Xie, J. Zhang,R. Li,J. Li, P. Hong,J. Xia and P. Chen,"Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning,"CEA, vol. 119, pp. 123–132, 2015,doi:10.1016/j.compag.2015.10.015

[11] M. Mayo and A. T. Watson, "Automatic species identification of live moths," KBS, vol. 20, pp. 195–202, Feb. 2007.

[12] B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik,"Simultaneous Detection and Segmentation,"ECCV,vol.7, pp. 297–312, 2014.

[13] W. Ding and G. Taylor, "Automatic Moth Detection From Trap Images For Pest Management", CEA, pp.1-17,Feb. 2016.

[14] J. Redmon, S. Divvala, R. Girshick and A. Farhadi,"You Only Look Once:Unified, Real-Time Object Detection,"May. 2016.

[15] J. Redmon and A. Farhadi,"YOLO9000:Better, Faster, Stronger, Dec. 2016.

[16] M. M. Raphael and R.Maheswari,"Automatic Monitoring of Pest Trap,"IJAREEIE, vol. 5,no. 4, Apr. 2016.

[17] O. López, M. M. Rach, H. Migallon, M. P. Malumbres, A. Bonastre and J. J. Serrano,"Monitoring Pest Insect Traps by Means of Low-Power Image Sensor Technologies,"MDPI vol. 12, no 11, pp. 15801–15819, Nov. 2012, doi: 10.3390/s121115801.

[18] Source: https://www.indiamart.com/proddetail/delta-phero-sensor-trap-9370409448.html.

[19] Source: https://www.lsuagcenter.com/~/media/system/0/5/0/8/050869493d5a066be3d4c7a5741f7c08/photo_4.jpg

[20] L. K. Musambo and J. Phiri, "Student Facial Authentication Model based on OpenCV's Object Detection Method and QR Code for Zambian Higher Institutions of Learning", IJACSA,vol. 9, no. 5, 2018, pp 88 - 94,Jan. 2018.

# A Review of Security Concerns in Mobile Application Systems: Case of USSD and Android Applications

Michael Bwalya[1] and Christopher Chembe[2]

Mulungushi University, Department of Computer Science and Information Technology,
Box 80415, Kabwe, Zambia.
1. mikob87@gmail.com, 2. cchembe@mu.edu.zm

**Abstract**
Recently, the mobile industry has experienced an extreme increment in number of its users. The Global System for Mobile (GSM) network with the greatest worldwide number of users succumbs to several security vulnerabilities. Unstructured Supplementary Service Data (USSD) is one of the technologies widely used in conducting mobile transactions. This technology has its strengths and weaknesses from perspectives of security of systems. It utilizes GSM Services and GSM Security is known to have inherent flaws in its encryption and authentication algorithms. Security is the biggest issue in the field of mobile payments and data collection, because without secure commercial information exchange and safe electronic financial transactions over mobile networks, no one will trust the USSD-Based mobile payment application and android based application. Mobile application security breach can lead to fraudulent transactions (Revenue Loss) through mobile applications, confidentiality (Users sensitive data- Credit/Debit Card Data PIN, user credentials), revenue loss through communications services misuse, brand value degradation through SIM card cloning related attacks, misuse of Enterprises Data through personal handheld devices, fraudulent transactions through USSD and DST (Dynamic SIM Toolkit) Applications. This paper proposes a solution to design a security framework which addresses USSD based mobile payment applications and android application security in the application to ensure information security confidentiality and integrity.

Keywords: **USSD, PIN, GSM, SMS, Security framework, android**.

## 1 Introduction

The use of mobile devices to establish ad-hoc communication systems is a viable solution that provides global connectivity to support a broad range of applications[1]. As mobile technologies are becoming more advanced and mobile devices are making a big impact on daily life, a new type of payment system named mobile payment (m-payment) has emerged, enabling users to pay from their wireless devices especially mobile phones wherever they go [2] . However, mobile payment is surprisingly not among the frequently used mobile services, although technologically advanced solutions exist. Unstructured Supplementary Services Data (USSD) and Short Message Service (SMS) are among the technologies widely used in conducting mobile real-time transactions [3]. USSD is a session-based, real-time communication technology for supplementary services. USSD is used in sending messages across a GSM network between a mobile client and an application server. It operates much like SMS but its session-based and interactive nature distinguishes the two. Unlike SMS, it does not operate by store-and-forward and its turnaround response time is much shorter for interactive applications than it is for SMS[4].This makes USSD much faster and very cost effective as it involves simple operations that are also handset independent (old handsets to most recent smartphones can all access the service). USSD applications are characterized by menu-driven and interactive services and a request is invoked by dialing a number that is composed of asterisks (*) and hashes (#). Examples of these services include sports updates, movies, weather information, news, stock market, reservation applications (for planes/trains/ movies, etc.), voting/polling applications, mobile account balance checking and top up, and many others. The big risk with USSD lies in the fact that data carried within the communication channel is not itself encrypted[3].

As the android market is growing, security risk has increased and thus focus should be given to the security. Android devices have gained huge market share due to the open architecture of Android, and the popularity of its application programming interface (APIs) in the developer community. Increased popularity of the Android devices and associated monetary benefits attracted the malware developers, resulting in big rise of the Android based security issues [5].

0

The apps downloaded from Google Play like True Caller, Viber, Whatsapp, IndiaLive and billions of other apps have changed the life of a typical android user with 1.2 billion people worldwide using mobile apps at the end of 2012.[6] [7]. As on December 2016, India had estimated 432 million Internet users. Report also confirms 42% of points of access for internet are mobile phone in rural India. Report by StatCounter, mobile and tablet devices accounted for 51.3 percent of internet usage worldwide compared to 48 percent on desktops. This expresses use of mobile device for accessing internet is increasing year after year because of advantages of mobility, price compare to desktop[8]. A report by Kantar IMRB & MMA about "Smartphone Usage and Behavior" indicate smartphone users are spending more time than on any other media, including TV and Print[9]. Social media and messaging apps account for the highest reach among all categories. They also account for almost 50% of all time spent on smartphones. An average user spends approximately 3 hours daily accessing the Internet on a smartphone. Active internet users spend 60% more time on mobile compared to Desktop[10].Trusted apps are available in Google's market which is self-signed by the developers but Malware has even appeared in Google's market. Two examples are Droid Dream and Droid Dream Light. Both these apps were found on the Android market in early 2011 and both applications steal personal data and are very much like traditional Trojans seen on the desktop [11]. With these many number of users using the apps downloaded from the Google play, securing the use of these apps is of paramount importance to researchers. In view of this, this research will seek to develop a security framework that addresses the concerns in mobile applications systems.

The rest of the paper is organized as follows. Next section discuses GSM security architecture, section 3 looks at android architecture. Section 4 discusses proposed security architecture while section 5 concludes the paper.

## 2 GSM and GPRS Security Architecture

Global System for Mobile Communications (GSM) is the most popular standard for mobile phones in the world. The General Packet Radio Service (GPRS) core network is an integrated part of the GSM network; it is layered over the underlying GSM network, with added nodes to cater for packet switching. GPRS also uses some of the existing GSM network elements; some of these include existing Base Station Subsystems (BSS), Mobile Switching Centers (MSC), Authentication Centers (AUC), and Home Location Registers (HLR). Some of the added GPRS network elements to the existing GSM network include; GPRS Support Nodes (GSN), GPRS tunneling protocol (GTP), Access points, and the (Packet Data Protocol) PDP Context[12].

### USSD Technologies

The Unstructured Supplementary Services Data (USSD) is a session-based, real-time communication technology for supplementary services. USSD is used in sending messages across a GSM network between a mobile client and an application server. It operates much like SMS but its session-based and interactive nature distinguishes the two. Unlike SMS, it does not operate by store-and-forward and its turnaround response time is much shorter for interactive applications than it is for SMS. This makes USSD much faster and very cost effective as it involves simple operations that are also handset independent (old handsets to most recent smartphones can all access the service)[3].

### USSD Architecture and Operation

In USSD, when the service is invoked, a real-time, interactive session is established between a client and an application server on the network. This allows data to be exchanged between the customer and the service provider until the service is completed. A session needs to be allocated to every transaction request; the response for this request and the following series of requests and responses in that session all share the same session ID until the session is closed or times out. The communication can be established even when a call is active because the two services use different communication channels. USSD services use signaling channel while call services use traffic channels [3].

### Security of USSD

Despite the convenience offered by USSD to customers in accessing the services; the technology is not without its associated security risks. When compared to SMS, USSD is considered to be relatively more secure because no copy of the message is stored on customer's phone or at the SMSC. A single session is established between the mobile terminal and the application server, and at the USSD gateway the message is encrypted preventing data to be misused between the gateway and the server. The big risk lies on the fact that data carried within the communication channel is not itself encrypted. If GSM encryption is broken, this data can be then be accessed. As it has been mentioned earlier, the A5 encryption that is used in GSM has been reverse engineered and thus leaving USSD data vulnerable to attacks because messages are not encrypted on the GSM backbone.[13] Moreover, GSM encryption is only applied between the mobile terminal and the base station; across the rest of the operator's network, the message is in plaintext[3].

### Solutions to the GSM Security Flaws

1) Using secure algorithms for A3/A8 implementations: This can thwart the dangerous SIM card cloning attack. This solution is profitable since the network operators can perform such improvement

themselves and without any need to the software and hardware manufacturers or the GSM consortium. However, this solution requires providing and distributing new SIM cards and modifying the software of the HLR. Currently, both COMP128-2 and COMP128-3 algorithms thwart the SIM card cloning and over-the-air cracking of Ki. Since COMP128-3 enhances the effective key length of the session key to further 10 bits, it allows the deployed cryptographic algorithm to have its nominal security. Although it is soon to judge on the real security of COMP128-2 and COMP128-3, they have apparent advantages over the traditional COMP128-1 that its SIM cloning apparatus are available at very low prices [14].

2) Using secure ciphering algorithms: Operators can use newer and more secure algorithms such as A5/3 provided that such improvements are allowed by the GSM consortium. The deployed cryptographic algorithms should be implemented on both BTS and mobile phones. Any change to the cryptographic algorithms requires agreement and cooperation of software and hardware manufacturers since they should perform the appropriate changes to their product.
3) Securing the backbone traffic: Encrypting the backbone traffic between the network components can prevent the attacker to eavesdrop or modify the transmitted data. Although this solution may be implemented without the blessings of GSM consortium, the cooperation of hardware manufacturers is still required[14].
4) End-to-end Security: The best, easiest, and most profitable solution is to deploy the end-to-end security or security at the application layer. Most of GSM security vulnerabilities (except SIM cloning and DoS attacks) do not aim ordinary people, and their targets are usually restricted to special groups so it is reasonable and economical that such groups make their communications secure by the end-to-end security [15] [14].

## 3 Android Security Architecture

Android seeks to be the most secure and usable operating system for mobiles by providing security measures to protect user data, system resources and it isolate applications. Google provides following security features to achieve these objectives [16].

- Robust security at the operating system level through the Linux kernel.
- The OS is sandboxed, preventing malicious processes from crossing between applications.
- Secure interposes communication
- User defined permissions
- Application signing

Despite of these attempts, it fails to address the issue of infection all together.

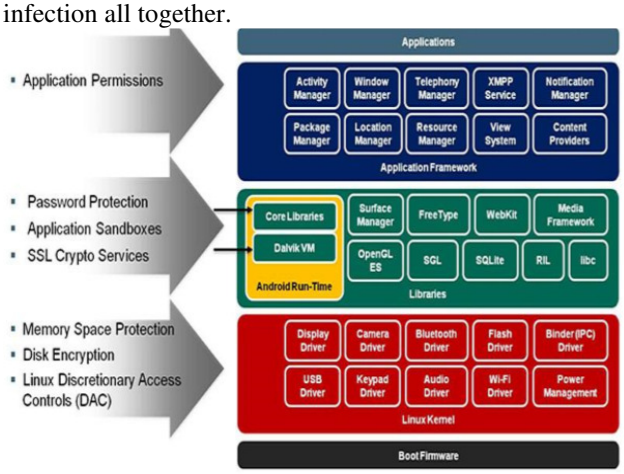



**Figure 1: Summarizes security provided at various levels of Android. Every level assumes that level below is properly secured.**

In the Linux Kernel, the main purpose of memory space protection is to prevent a task from accessing memory without proper access permissions. Without memory protection, memory segment like code and data segment are vulnerable to memory related bugs and code injection attacks. Disk encryption ensures that files are always stored on disk in an encrypted form. Security Enhanced Linux (SELinux), an access control implementation for the Linux kernel which was introduced recently has prevented multiple vulnerabilities, and now it has been strengthened even more to meet the needs of enterprise customers that have strict security requirements. All process run above the Linux kernel is restricted by the application sandbox. In the Libraries, The Android platform takes advantage of the Linux user-based protection as a means of identifying and isolating application resources. This approach is different from other operating systems (including the traditional Linux configuration), where multiple applications run with the same user permissions. This sandbox is dissimilar than the sandbox found on the J2ME or Blackberry platforms

### Security Weaknesses of Android System

Unfortunately, the increasing adoption of smartphones comes with the growing prevalence of mobile malware. As the most popular mobile platform, Google's Android overtook others (e.g., Symbian) to become the top mobile malware platform [17]. Security breach on android architecture may come in two ways from outside offensive activities; attack due to user unawareness and attack due to system defects. Most attacks exploit vulnerabilities of the

smart phone. The threats that are seen appearing on mobile are rootkits, Trojans, and even botnets. Since new malwares are appearing almost on a daily basis and it is hard to replace depicted secured device with a latest secured device in this pace by a user; user awareness can go in a great length to stop outside interference to the current device [16].

Android shows the authority information required by the system when installing application. User can check the permission required by the application and select whether to install or not. However, there is a difficulty for general users to check all permission during installation. This will eventually lead to user responsibility in case of a problem. As this authority information can be checked only during installation, this can be a security threat for those without knowledge or interest in the concept of authority. Malicious applications can make use of the address book, SMS and mobile phone information using the authority information. Also, it can manipulate or delete location information and personal information or even extort password by phishing [18]. A number of available android apps have a necessity to pass the messages over the web and therefore are considered for securing the sensitive data during the transition. Further it has been studied that these apps make use of SSL/TLS protocols. It has been estimated that 1,074 (8.0%) of the apps studied and analysed have the SSL/TLS code which is prone to MITM attacks [19]. Other numerous vulnerabilities are caused due to user unawareness. This vulnerabilities occur when a user [16]

- installs third party applications
- roots a device
- connects to a unsecured WI-FI network& maintaining unsecured Bluetooth connection
- gives remote access to PC
- connects to SD card and external device
- clicks on spam emails/sms/mms
- grants unnecessary permission to an application

There are different kinds of privacy and security threats to a Smartphone. Smartphone contains logs and logs contain information about all the activities user has done[20]

**Android Threats and Attacks**

Smart phones are under many threats and attacks. Few threats and attacks are summarized below [20] [21].

**Spam**

A spam can enter a smart phone through emails or Multimedia messages. Spam emails or messages may include links which direct users to phishing websites. Spam can also be used to cause denial of service attacks.

**Phishing**

It is a way to steal user name, password, credit card details, and other personal information by masquerading as a trusted identity. These attacks are present in social networking websites, emails, and Multimedia messages, etc and what they do is that they contain links and those links redirect the user to a website. If user fills in its username and password then attacker gets that information and can further use it.

**Snooping**

An attacker can pretend to be some other identity or a trusted identity. Attacker sends the messages but it appears to have come from a trusted party or some other identity.

**Sniffing**

It is called as eaves dropping a smart phone. It was shown that GSM's encryption for call and SMS privacy could be broken in less than a minute. There are different ways to tap a smart phone

**Pharming**

Attacker can redirect web traffic in a smart phone to a malicious website. Attacker collects smart phone's user information and after that few specific attacks are possible. For example, when a smart phone is used to browse a website, the HTTP header includes some information which can be used to start some other attacks on the smart phone.

**Denial of Service attack**

Text messages or incoming calls can be used by the attacker to commit this attack. Receiving hundreds of messages and calls could disable a smart phone.

## 4 Proposed Security Framework

The framework will focus on security enforcement on the mobile app, USSD application, systems dashboard and the server end of system.

**Android based Mobile Application**

The objective is to provide security against the Apps which are installed by the end user and is given all the permissions at the time of installation. This enhanced security has the desirable property of not disturbing a regular user in any noticeable way [11]. In fact, the user need not even be aware that the Security API has been applied. There is need to prevent the modification and access of data from mobile phones by other external malicious applications unknowingly. An API is proposed which will enhance the security of existing Android Framework by addressing the following limitations of Android Security Framework. The first step in the proposed security API will be implemented by adopting an encryption technique utilizing Advanced Encryption Algorithm (AES) and applying it to all the personal files in the Smartphone. All the files which are important for the cell phone owner should be encrypted. Nobody can break the algorithm password as AES is considered to be the strongest cryptographic algorithm among the existing ones [11]. File operations offered by the proposed Security API should aid in the detection of potentially malicious Apps whose behavior matches that of Malware. Malware recognition is usually achieved by

signature matching, heuristic analysis, or comparing hash-values. This enhancement will help achieve security on the android smartphone side.

**USSD Application**

The best, easiest, and most profitable solution is to deploy the end-to-end security or security at the application layer. Most of GSM security vulnerabilities (except SIM cloning and DoS attacks) do not aim ordinary people, and their targets are usually restricted to special groups so it is reasonable and economical that such groups make their communications secure by the end-to-end security. Since the encryption and security establishment is performed at the end-entities, any change to the GSM hardware will not be required. In this way, even if the conversation is eavesdropped by the police or legal organizations, they cannot decrypt the transmitted data without having the true ciphering key, provided that a secure enough cryptographic algorithm is deployed. Therefore, in order to avoid illegal activities, it should be transparent to both GSM operator and service provider. It may also be necessary to find solutions for a legal interception or a key screw scheme. The end-to-end security establishment has a complete flexibility to the deployed algorithms so the appropriate upgrades can be easily accomplished when necessary. However, it may be a subject to export control. Generally, the end-to-end security can be provided in the cellular systems by following one or some of the following approaches [14]:

- Exploiting the processing capabilities of mobile phone using the programming languages such as J2ME (Java 2 Mobile Edition): Supported by the most recent cellular phones and Personal Digital Assistants (PDA) with the improved processing capabilities.
- Exploiting the processing capabilities of the SIM using the SIM Application Toolkit (SAT) Not supported by all SIM cards; especial SIM cards are required; the processing resources are still limited; and operations may be so time-consuming.
- Exploiting the processing capabilities of an additional smart card, e.g. JavaCard: Not supported.

**Dashboards**

There is need to secure the dashboards so that they meet the CIA (Confidentiality, Integrity and Availability) triad and protect against various forms of attacks. Dashboards maybe susceptible to attacks like cross-site scripting (XSS). This attack occurs when an attacker injects malicious code (typically JavaScript) into a site for the purpose of targeting other users of the site. The other common attack is Cross-Site Request Forgeries (CSRF). CSRF attacks often exploit the authentication mechanisms of targeted sites. The root of the problem is that Web authentication normally assures a site that a request came from a certain user's browser; but it does not ensure that the user actually requested or authorized the request [22].

**Server side**

There are a number of vulnerabilities associated with systems. These include at the client end, in the communication channels, and on the server end. These vulnerabilities can be summarized as lack of end-to-end security for the data being transmitted, vulnerabilities in the authentication mechanisms, and vulnerabilities in the application server security policies.

The PIN characters entered by a customer on one's phone are not masked, thus being visible to someone who may be watching. Also, the PIN used is only 4 digit numerals which can be easy to guess. Moreover, there is a lack of data confidentiality as data is not encrypted between the customer and the server and there are no mechanisms for integrity checks. Additionally, the security policies in the application servers leave some vulnerabilities which can be exploited to perform attacks on the system. Periodic changes of PIN are not enforced, and the use of a system's default PIN is allowed in some systems.

To address these vulnerabilities; a model has been designed to cater for enhanced security controls. The features to be included in the model included:

- Encrypting data across the communication channel to ensure end-to-end security,
- Increasing the domain of characters to be used in a PIN (mixture of letters, symbols, and numbers) and
- Masking the PIN characters as they are being typed in.
- Along with these; the improvements to be made on server security policies include enforcing periodic change of PIN, prohibiting the usage of someone's year of birth as a PIN, and restricting the use of a system's default value as a PIN.

The proposed framework has to fulfill the following features: to ensure the integrity of an application at setup and to secure the data being sent between the mobile side and the server side. This architecture has to be able to adapt the security services according to the user needs, device characteristics and user context.

Figure 2 shown below is a proposed security framework that will be used.

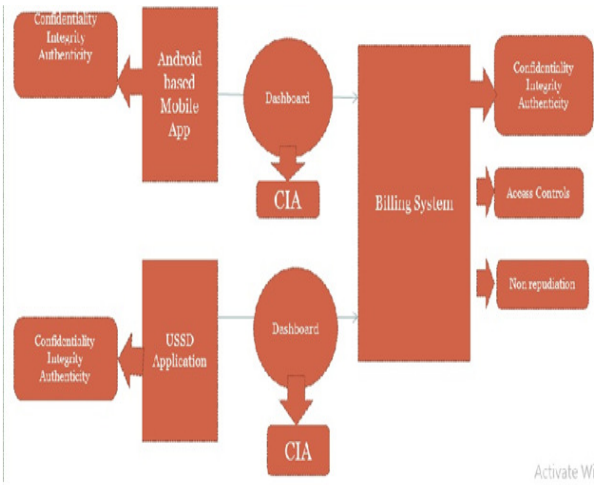



Figure 2 Security framework for Mobile Applications

## 5 Conclusion

In this paper, the security of the GSM network is evaluated, and a complete and brief review of its security problems is presented. It is proved that the GSM network has many inherent security flaws that can be misused for fraudulent purposes or for deceiving users. Some practical solutions to improve the security of currently available 2G networks are also proposed. Some solutions include the security improvement of the infrastructure while the others tend to provide the end-to-end security. It is also deduced that the end-to-end security or the security at the application layer is the best and most profitable solution for the currently available 2G systems. On the other hand, many Smartphone operating systems are vulnerable to attacks because the Smartphone user is instrumental in deciding which applications will be installed on the phone. It is not easy for a user to judge applications by their description. The Android framework is one platform that expects the user to be security conscious and implicitly assumes applications developers are not malicious. Because of this, a user may unknowingly install software that poses a security threat or is not efficient enough to handle the user's privacy issues. This paper proposed a security framework that addresses security flaws in USSD based applications and android based applications, from the mobile side up to the server side.

## 6 References

[1] D. Huang, Z. Zhou, L. Xu, T. Xing, and Y. Zhong, "Secure data processing framework for mobile cloud computing," in *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*, 2011, pp. 614–618.

[2] S. Kadhiwal and A. U. S. Zulfiquar, "Analysis of mobile payment security measures and different standards," *Comput. Fraud Secur.*, vol. 2007, no. 6, pp. 12–16, 2007.

[3] B. W. Nyamtiga, A. Sam, and L. S. Laizer, "Security Perspectives for USSD versus SMS in conducting mobile transactions: A case study of Tanzania," *Int. J. Technol. Enhanc. Emerg. Eng. Res.*, vol. 1, no. 3, pp. 38–43, 2013.

[4] S. Desai, "Mitigating Security Risks in USSD-Based Mo-bile Payment Applications," 2011.

[5] P. Faruki *et al.*, "Android security: a survey of issues, malware penetration, and defenses," *IEEE Commun. Surv. Tutor.*, vol. 17, no. 2, pp. 998–1022, 2015.

[6] M. A. Dar and J. Parvez, "A novel strategy to enhance the android security framework," *Int. J. Comput. Appl.*, vol. 91, no. 8, 2014.

[7] W. Shin, S. Kiyomoto, K. Fukushima, and T. Tanaka, "Towards formal analysis of the permission-based security model for android," in *Wireless and Mobile Communications, 2009. ICWMC'09. Fifth International Conference on*, 2009, pp. 87–92.

[8] S. GlobalStats, "Mobile and tablet internet usage exceeds desktop for first time worldwide," *Httpgs Statcounter Compressmobile--Tablet-Internet-Usageexceeds-Deskt.--First-Time-Worldw.*, 2017.

[9] K. I. &. MMA, "Smartphone Usage and Behaviour Report," 2016.

[10] P. P. Ghogare and M. P. Patil, "A Study of Security Awareness Among Android Users," *Int. J. Comput. Sci. Eng. Technol. IJCSET*, 2017.

[11] M. A. Dar and J. Parvez, "A novel strategy to enhance the android security framework," *Int. J. Comput. Appl.*, vol. 91, no. 8, 2014.

[12] K. Chikomo, M. K. Chong, A. Arnab, and A. Hutchison, "Security of mobile banking," *Univ. Cape Town South Afr. Tech Rep Nov*, vol. 1, 2006.

[13] M. Toorani and A. Beheshti, "Solutions to the GSM security weaknesses," in *Next Generation Mobile Applications, Services and Technologies, 2008. NGMAST'08. The Second International Conference on*, 2008, pp. 576–581.

[14] M. Toorani and A. Beheshti, "Solutions to the GSM security weaknesses," in *Next Generation Mobile Applications, Services and Technologies, 2008. NGMAST'08. The Second International Conference on*, 2008, pp. 576–581.

[15] J. L.-C. Lo, J. Bishop, and J. H. Eloff, "SMSSec: an end-to-end protocol for secure SMS," *Comput. Secur.*, vol. 27, no. 5–6, pp. 154–167, 2008.

[16] T. I. Mamun and L. Alam, "Android Security Vulnerabilities Due to User Unawareness and Frameworks for Overcoming Those Vulnerabilities," *Int. J. Comput. Appl.*, vol. 137, no. 1, pp. 14–21, 2016.
[17] Yajin Zhou, "Dissecting Android Malware: Characterization and Evolution," *IEEE Symp. Secur. Priv.*, 2012.
[18] J.-K. Park and S.-Y. Choi, "Studying security weaknesses of android system," *Int. J. Secur. Its Appl.*, vol. 9, no. 3, pp. 7–12, 2015.
[19] T. K. Chawla and A. Kajala, "Transfiguring of an Android App Using Reverse Engineering," 2014.
[20] N. Kaur, "Techniques Used for Detection of Mobile Spyware," *Int. J. Comput. Trends Technol. IJCTT*, vol. 11, no. 5, pp. 217–219, 2014.
[21] P. S. Dhumal, "A Review on Android Security," *Int. J. Mod. Electron. Commun. Eng. IJMECE*, vol. 3, no. 3, pp. 53–58, 2015.
[22] W. Zeller and E. W. Felten, "Cross-site request forgeries: Exploitation and prevention," *N. Y. Times*, pp. 1–13, 2008.

# Challenges of Identity Management Systems and Mechanisms: A Review of Mobile Identity

Raphael Banda[a], Jackson Phiri[b]
The University of Zambia
Department of Computer Science
Lusaka, Zambia
[a]raphael.banda@yahoo.co.uk, [b]jackson.phiri@cs.unza.zm

*Abstract* — **Digital identity and management lays the groundwork necessary to guarantee that the Internet infrastructure is strong enough to meet basic expectations for security and privacy. "Anywhere, anytime" mobile computing is becoming real; in this ambient intelligent world, the choice of identity management mechanisms will have a large impact on social, cultural, business, and political aspects of our lives. From their point of view, Privacy is a human need, and all of society would suffer from its demise; people have hectic lives and cannot spend all their time administering their digital identities. Most systems do not fulfil several of these tests; they are particularly deficient in fine-tuning the access control over identity to minimize disclosure of data.**

**We have surveyed how the requirements for user-centric identity management and their associated technologies have evolved, with emphasis on federated approaches and user centricity. Second, we have focused on related security standards as well as platforms, such as Higgins and finally, we have looked at identity management in the field of mobility with focus on the future of mobile identity management.**

**This paper looks the procedure for a defence mechanism against Man in the Middle as one of the many examples that mobile computing faces challenges in security pitfalls. We have mainly centred on how best we can create strong passwords that cannot be easily decoded by man in the middle.**
**The methodology that was adopted to write this paper was that of a literature survey. In doing so we looked at different literatures to come up with a concrete and coherence view of the problem of challenges of identity management systems of mobile users.**
**Keywords - Digital identity, mobile computing, Man-in-the-middle attack, Silo, Authentication, OpenID Stack, Biometric, Higgins, access control.**

## 1. INTRODUCTION

Identity management in an enterprise is a combination of processes and technologies to manage and secure access to the information and resources of an organisation while also protecting user profiles, including customer profiles [1]. Mobile computing is becoming easier, more attractive, and even cost-effective: Mobile devices carried by roaming users offer more and more computing power and functionalities, including sensing and providing location awareness. [2] Many computing devices are also deployed in the environments where the users evolve—for example, intelligent home appliances or RFID-enabled fabrics. [2] In this ambient intelligent world, the Internet is most likely going to generate more complicated privacy problems. [2]

Identity has become a burden in the online world. When it is stolen, it engenders a massive fraud, principally in online services, which generates a lack of confidence in doing business with providers and frustration for users [2].

Therefore, the whole of society would suffer from the demise of privacy, which is a real human need. Because people have hectic lives and cannot spend their time administering their digital identities, there is a need for consistent identity management platforms and technologies enabling usability and scalability, among other things [2].

We surveyed how the requirements have evolved for mobile user-centric identity management and its associated technologies. We realise how often people want to use their mobiles everywhere and anytime and for different reasons including keeping company data and information The security of mobiles is therefore crucial in this respect. The Android plat-form can also be used for Internet of Things (IoT), BOYD(bring your own device) to improve work efficiency and productivity and access enterprise resources [3].

## 2. DIGITAL IDENTITY DEFINITION

By definition digital identity is a representation of an entity in a specific context [2]. For a long time digital identity has been considered the equivalent of a user's real-life identity [2]. Considering the fact that the real-world identity and a digital identity is not always mandatory. In a major identity management initiative digital identity is defined as "the distinguishing character or personality of an individual. An identity consists of traits. Digital identity management is a key issue that will ensure not only service and functionality expectations but also security and privacy. [2]


*Figure 1: Identity management main components [2]*

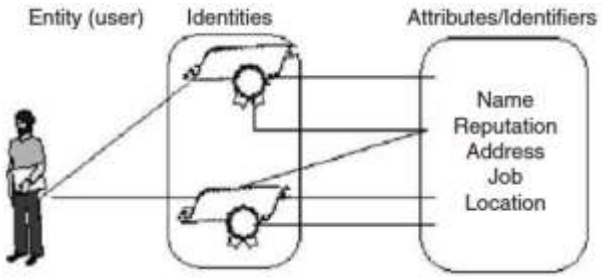



*Figure 2: Relationship among identities, identifiers and entity [2]*

The relationship among identifiers, identifiers:

- A user who wants to access to a service
- Identity Provider (IdP), the issuer of user identity
- Service Provider (SP), the relay party imposing an identity check
- Identity (Id), a set of user attributes
- Personal Authentication Device (PAD), which holds various identifiers and credentials and could be used for mobility

Authentication is the process of verifying claims about holding specific identities. A failure at this stage will threaten the validity of the entire system. Technology is constantly finding stronger authentication using claims based on the following:

- Something you know (password, PIN).
- Something you have (one-time-password).
- Reachability. The management of reachability allows a user to handle their contacts to prevent misuse of their email address (spam) or unsolicited phone calls.
- Authenticity. Ensuring authenticity with authentication, integrity, and nonrepudiation mechanisms can prevent identity theft.
- Anonymity and pseudonymity. Providing anonymity prevents tracking or identifying the users of a service.
- Organization personal data management. A quick method to create, modify, or delete work accounts is needed, especially in big organizations.

Microsoft's first large identity management system was the Passport Network. It was a very large and wide- spread Microsoft Internet service, an identity provider for the MSN and Microsoft properties and for the Internet [2]. However, with Passport, Microsoft was suspected by many of intending to have absolute control over the identity information of Internet users and thus exploiting them for its own interests. Passport failed to become "the" Internet identity management tool. [2] Since then, Microsoft has clearly come to under- stand that an identity management solution cannot succeed unless some basic rules are respected. That's why Microsoft's Identity Architect, Kim Cameron, has stated the seven laws of identity [2]. His motivation was purely practical in determining the prerequisites of creating a successful identity management system. He formulated these essential principles to maintain privacy and security: [2]

- User control and consent over the handling of their data
- Minimal disclosure of data, and for a specified purpose
- Information should only be disclosed to people who have a justifiable need for it
- The system must provide identifiers for both bilateral relationships between parties and for incoming unsolicited communications
- It must support diverse operators and technologies
- It must be perceived as highly reliable and predictable.
- There must be a consistent user experience across multiple identity systems and using multiple technologies.

## 3. PRIVACY REQUIREMENT

Privacy is a central issue because the official authorities of almost all countries have strict legal policies related to identity. It is often treated in the case of identity management because the management deals with personal information and data. We cannot run away from the fact that the future of identity is digital

## 4. IDENTITY MANAGEMENT SOFTWARE

The evolution of identity management systems is toward the simplification of the user experience and reinforcing authentication. It is well known that poor usability implies a weakness in authentication. The following are some of the identity management software:

*SAML*

The Security Assertion Markup Language (SAML) is an OASIS specification 41 that provides a set of rules for the structure of identity assertions, protocols to move assertions, bindings of protocols for typical message transport mechanisms, and profiles [2].

*OpenID 2.0*

The intent of the OpenID framework is to specify layers that are independent and small enough to be acceptable and adopted by the market. OpenID is basically providing simple attribute sharing for low-value transactions. OpenID authentication 2.0 is becoming an open platform that supports both URL and XRI user identifiers [2].

*OpenID Stack*

The first layer is for supporting users' identification. Using URL or XRI forms, we can identify a user. URLs use IP or DNS resolution and are unique and ubiquitously supported.

*CardSpace*

Rather than invent another technology for creating and representing digital identities, Microsoft has adopted the federated user-centric identity meta-system [2]. This is a serious solution that provides a consistent way to work with multiple digital identities.

*SXIP 2.0*

In 2004, SXIP 1.0 grew from efforts to build a balanced online identity solution that met the requirements of the

entire online community. SXIP 2.0 is the new generation of the SXIP 1.0 protocol, a platform that gives users control over their online identities and enables online communities to have richer relationships with their members [2].

What makes SXIP popular is its dynamic discovery. A simple and dynamic discovery mechanism ensures that users are always informed online about the Homesite that is exporting their identity data [2]. In addition to this Simple implementation. SXIP 2.0 is open source using various high-level development languages such as Perl, Python, PHP, and Java. Making it effortless to integrate into a Web site because it uses a URL-based protocol. Interoperability also makes SXIP 2.0 good because it can coexist with other URL-based protocols [4].

Figure 3 shows some of the relationships among the mentioned identity solutions.

It is clear from figure 3 that the Higgins Trust framework is one of the best solutions to identity management.

## 1. HIGGINS TRUST FRAMEWORK (HTF)

The Higgins Trust Framework (HTF) is one of the best identify framework that users have found to be very useful worldwide. The HTF or Higgins, for short is an open source project under development by the Eclipse Foundation. It seeks to make sharing of identity information easier and more secure. Some of the contributors to this project are IBM, Novell and Parity Communications.

| Requirement | Empowering the total control of users over their privacy | Usability, as users are using the same identity for each identity transaction | Giving a consistent user's experience thanks to uniformity of identity interface | Limiting identity attacks i.e. Phishing | Limiting reachability/disturbances, such as spam | Reviewing policies on both sides when necessary, identity providers and service providers | Huge scalability advantages as the Identity Provider does not have to get any prior knowledge about the Service | Assuring secure conditions when exchanging data | Decoupling digital identity from applications | Pluralism of Operators and Technologies |
|---|---|---|---|---|---|---|---|---|---|---|
| XRI/XDI | ■ | ■ | | | | | | ■ | ■ | |
| ID/WSF | ■ | ■ | ■ | | | ■ | ■ | ■ | ■ | ■ |
| Shibboleth | ■ | | | | | | | ■ | | |
| CardSpace | ■ | ■ | ■ | | | ■ | ■ | ■ | ■ | |
| OpenID | ■ | ■ | ■ | | | | ■ | ■ | ■ | ■ |
| SXIP | ■ | ■ | ■ | | ■ | | ■ | ■ | ■ | ■ |
| Higgins | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ | ■ |

*Figure 3: The Higgins ID Framework compared to other frameworks [2]*

According to Dale Olds, an engineer at Novell, the purpose of the project is to give users more control over their online identity information. [5]

The Higgins framework enables users to securely store identity information and related data and to integrate that data across multiple systems and applications [2]. Stored data can be shared anonymously among Web applications, online vendors and service providers in a controlled manner. HTF API (application program interface) can be thought of as a repository for cookie-like data that makes it convenient for users to conduct e-commerce and interact with Web sites, without the security problems inherent in conventional cookies that reside on the user's hard disk. [5]

The objective is to develop an extensible, platform-independent, identity protocol-independent software framework that provides a foundation for user- centric identity management [4]. It enables applications to integrate identity, profiles, and relationships across heterogeneous systems [4].

The HTF platform intends to address four challenges that normally characterise most identity authentication platforms [2]:

- The need to manage multiple contexts
- The need for interoperability [2]
- The need to respond to regulatory, public, or customer pressure to implement solutions based on trusted infrastructure that offers security and privacy [2].

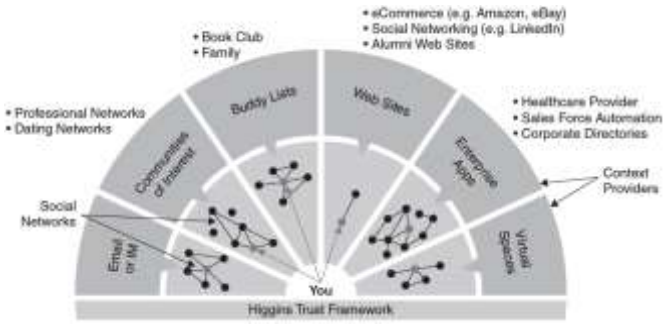



*Figure 4: The Higgins Trust Framework and context [2]*

*Mobile Web 2.0*

Mobile Web 2.0 as a content-based service is an up-to-date offering of services within the mobile network. At the moment, mobile Web suffers from lack of interoperability and usability due to the small screen size and lower computational capability [2]. Fortunately, these limitations are only temporary, and within five years they will be easily over- come [2]

## 2. MOBILITY AND MOBILE SECURITY

The mobile identity may not be stored in one location but could be distributed among many locations, authorities, and devices [2]:

- Device mobility; where a person is using the same identity while using different devices.
- Location mobility; where a person is using the same devices while changing location.

- Context mobility; where a person is receiving services based on different societal roles: as a parent, as a professional, and so on.

## 3. EVOLUTION OF MOBILE IDENTITY

Mobile identity management is in its infancy [2]. GSM networks, for example, provide management of Subscriber Identity Module (SIM) identities as a kind of mobile identity management, but they do not meet all the requirements for complete mobile identity management.

*PADs as a Solution to Strong Authentication*

A PAD60 is a tamper-resistant hardware device that could include smart cards and sensors or not. The user stores his identity in the PAD; whenever he would like to connect to a service provider:

1. He authenticates himself with a PIN code to use the PAD.
2. He chooses the password to be used for his connection to the specific service provider.
3. He launches and logs in to the specific service provider by entering his username and password.

The PAD is a good device to tackle the weakness and inconvenience of password authentication. It provides a user-friendly and user-centric application and even introduces stronger authentication. The fundamental

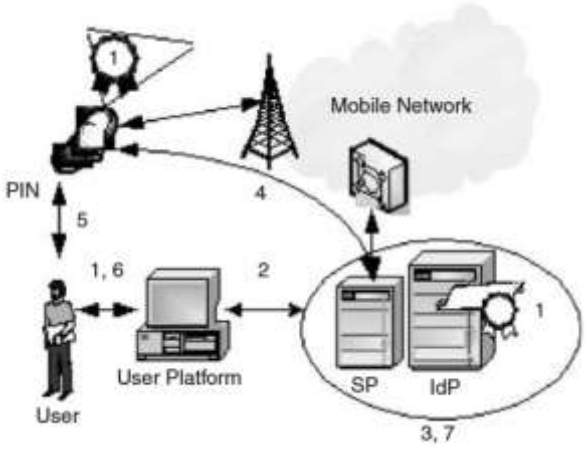



*Figure 5: SMS double-channel authentication [2]*

The scenario implemented by some banks is illustrated in Figure 5 and described as follows. First, the user switches on her mobile phone and enters her PIN code. Then:

1. The user logs in to her online account by entering her username and password (U/P).
2. The Web site receives the U/P.
3. The server verifies the U/P.
4. The server sends an SMS message with OTP.
5. The user reads the message.
6. The user enters the OTP into her online account.
7. The server verifies the OTP and gives access.

The problem with this approach is the fact that the cost is assumed by the service provider. In addition, some drawbacks are very common, mainly in some developing countries, such as lack of coverage and SMS latency. Of course, the MITM attack is not overcome by this approach. [2]

*Soft-Token Application*

In this case, the PAD is used as a token emitter. The application is previously downloaded. SMS could be sent to the user to set up the application that will play the role of soft token.

The scenario is identical to the SMS, only the user generates her OTP using the soft token instead of waiting for an SMS message. The cost is less than the SMS-based OTP. This approach is a single-channel authentication that is not dependent on mobile network coverage nor latency.

## 4. BIOMETRIC IDENTITY MANAGEMET

Biometric identities are some of the most secure kind of identities. It is for this reason that they have a strong place in computing and identity management. Biometric schemes are, generally, more secure than the traditional authentication methods mostly because metrics used in biometric authentication cannot easily be replicated [3].

In today's plain society and individual technology, based authentication methods are gradually becoming obsolete. Biometric authentication is taking over traditional passwords or ID card based authentication due to numerous advantages [6]. Biometric identification management system can be categorised according to the following attributes:

i) Who you are (Static biometrics) – authentication that depends on fingerprints, eye retina, and face. A physiological characteristic is a relatively stable human physical feature [7]. An example of a physiological characteristic is a fingerprint, retina iris pattern, or a hand geometry pattern [7].
ii) What you can do (Dynamic biometrics) – this kind of authentication uses handwriting, voice recognition and even typing rhythm. Behavioural characteristics can be identified in activities such as speech, hand-writing speed and pressure exerted on paper when writing among others [7].
iii) What you possess – authentication with what you actually possess like the, Identity card, Credit card, or even physical keys.
iv) What you know – is a form of identity that asks for what you know in form of say PIN number and/ or identity number.

Some of these personal identification management can be said to be traditional authentication methods because they mainly rely on "what you have" or "what you know" but can be reliable when a combination of these can be used [6].

### *The Pros of Biometric Security system*

*Biometrics authentication reduces maintenance costs*

Biometrics authentication reduces on the administrative and maintenance costs by:

i) reducing on the need for intense training and ongoing management costs.

ii) helping management save other costs such as the issuance of new ID cards, and replacing lost or damaged ID cards.
iii) eliminating the time consuming and resource draining need to reset passwords.
iv) Saving on IT time and cost of password resets every time an employee forgets their password.

*Enhances IT audit*
Biometric identification solution creates a concrete activity audit trail to help establish accountability. In IT audit each and every action or transaction will be recorded and clearly documented by the individual associated with it. IT biometric audit reduces the possibility of system misuse and fraud.

*Convenient*
Carrying around ID cards or physical keys can be an inconvenience. Biometric technology makes individual identification convenient without the need to carry around ID cards or remember complicated passwords [6]. Due to the fact that passwords can be forgotten or easily guessed and the fact that ID cards can be damaged, swapped, or shared, biometrics are more convenient because individual physiological attributes are always with you [6].

*Not Easy to Forge*
It is not easy to forge a Biometric attribute. For example there are no two people with the same finger prints or same eye retina. It is for this reason that it is are almost impossible to forge or duplicate a static or even dynamic biometric identity. Even if you manage to forge a biometric attribute such as a fingerprint, modern biometric devices with liveness detection have the capability to identify a fake from the original [6].

*Improved Return on Investment (ROI)*
Biometric identification management offers enhanced accuracy, improved accountability, and a reduction in opportunities for misuse [6]. Compared to traditional identification systems that may rely on passwords, ID cards, or personal identification numbers (PINs), the ROI is much higher with biometric identification systems [6].

*Seamless Integration*
Biometric identification systems can be seamlessly integrated with workforce management time and attendance systems, access control, surveillance, and visitor management solutions – all managed through a single window on a computer [6].

Biometrics provides centralized control for security administrators.

*The Cons of Biometric Security system*
Every system has its own disadvantages and Biometric authentication is not immune to disadvantages. In most cases the biggest disadvantage to the system are humans themselves who can force access to the system through illegal means or through impersonations.

Clearly, Biometric Security system seem to have more advantages than disadvantages. It is for this reason that we have seen serious mobile securities tilting towards biometric securities as they are more secure than traditional securities of passwords and codes.

## 1. MAN IN THE MIDDLE ATTACK

In mobile computing man in the middle (MITM) attack is one of the most important attacks most people face in as far as mobile attacks are concerned. There are more cases of man in the middle attacks nowadays than in the past. People are nowadays monitoring and eavesdropping others and its common phenomenon.

We now analyse and see how we can prevent MITM attack.

The MITM attack can be visualized as an active eavesdropping in which the attacker establishes separate connections with the victims and relays messages between them [8].

*Figure 6: Advantages of Biometric authentication [6]*

The majority of the defence mechanisms that are being used against MITM attacks are authentication based techniques which might be based upon the following. [8]

- Passwords
- Secret questions

- Public key infrastructures
- Voice recognition
- Biometrics ( Finger/thumb printing and Retina scan )
- Off-the-Record Messaging for instant messaging
- Off-the channel verification

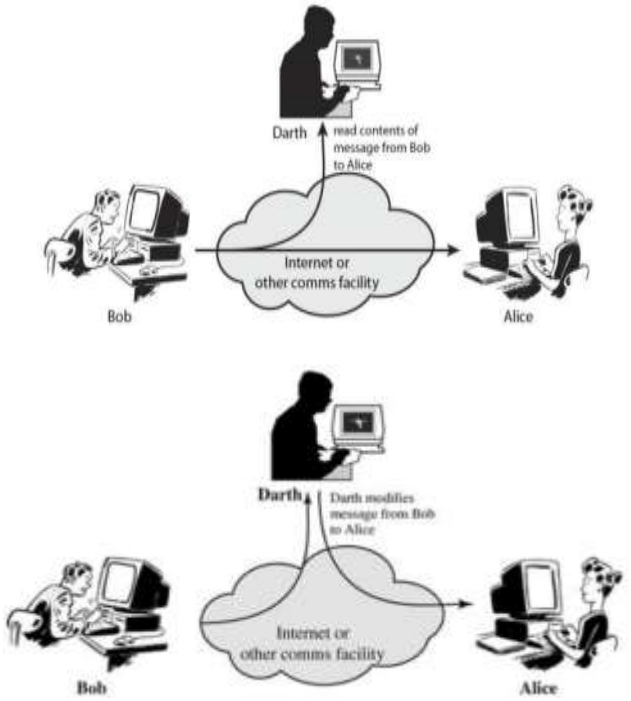



*Figure 7: Man in the Middle Attack [9]*

The following procedure for a defence mechanism against MITM has been suggested for the simplification of the process and at the same time hardening of the mechanism and vouching for the authenticity of the user a combinations of the following procedure that may be used along with the OTP scheme: [8]

We consider the following specific already known to the user code numbers which uniquely identify user:

1. A 10 digit Bank Account Number
2. A 10 digit Mobile Phone Number
3. Social Security Number of the user

*Generating the secret key*

Using the above individual details are as mentioned below:

1. 10 digit user bank account – **10 02 02 12 34**
2. 10 digit mobile number - **98 25 01 23 45**
3. 12 digit NAPSA social security no. - **ABCD EFGH IJKL**

Assuming the persons individual details are as mentioned below, the user's secrete key can be generated as below [8]:

10 digit user bank account – 10 02 02 12 34
10 digit mobile number - 98 25 01 23 45
12 digit Napsa social security no. – ABCD EFGH IJKL

Collecting the character at odd places in the 1st set of information we get

**0,2,2,2,4**

Collecting the character at even places in the 2nd set of information we get

**8,2,0,2,4**

Collecting the character at every third places in the 3rd set of information we get

**C,F,I,L,**

Adding a character in the last place in the third set of information to represent the social id type. In this case it is character "A" representing the Social Security unique id is added. Finally we have a matrix of 5 x 3

**0,2,2,2,4**
**8,2,0,2,4**
**C,F,I,LA**

Interchanging the rows and columns of the matrix for first time we get

**0,8,C**
**2,2,F**
**2,0,I**
**2,2,L**
**4,4,A**

The secrete key can be written like

**08C,22F,20I,22L,44A**

Dropping the commas we get a 15 digit unique code:

**08C22F20I22L44A**

## 2. IDENTITY MANAGEMENT - ACCESS CONTROL

What is also important in identity management is the way systems manage its control into the systems. There are four access control mechanisms that control the way users can access the system:

1. DAC – Discretionary Access Control
   This is a security access control mechanism, which controls the access permissions through data owner [10]. In DAC, the access rights of each user are performed during authentication by validating the username and password. DACs are discretionary as owner determines the privileges of access.

2. MAC – Mandatory Access Control
   This access mechanism is one of the oldest access mechanism that defines the user through the kernel. permission through the operating system or security kernel. Mac gives permission through the operating system or security kernel [10].

3. RBAC – Role-based Access Control
   In Role Based Access Control (RBAC) mechanism access rights are based on roles and privileges of the users [10]. In its basic form access control maps users with access to resources. Role-Based Access Control (RBAC) is a finer-grained method that maps defined roles with access to resources [11]. User permissions are given by different parameters of RBAC like, user-roles, role permissions and role-role relationships.

4. ABAC – Attribute-based Access Control
   ABAC defines the access control mechanism by the use of policies which determines different sets of attributes to check the access rights of each user [10]. The policies are generated using different types of attributes and based on the policies, the system determines the access permissions [10]. The considered attributes are subject attributes, object attributes, resource attributes and environmental attributes [10].

## 3. CONCLUSION

The Internet is increasingly used, but the fact that the Internet has not been developed with an adequate identity layer is a major security risk. Password fatigue and online fraud are a growing problem and are damaging user confidence [8].

Currently, major initiatives are under way to try to provide a more adequate identity layer for the Internet, but their convergence has not yet been achieved [4]. Higgins and Liberty Alliance seem to be the most promising ones [12].

In any case, future identity management solutions will have to work in mobile computing settings, any-where and anytime [4].

We have discussed the necessity of mobility and the importance of identity in future ambient intelligent environments [12]. Mobile identity management will have to support a wide range of information technologies and devices with critical requirements such as usability on the move, privacy, scalability, and energy-friendliness [4]. Biometric authentication has more advantages than other form of authentications because it can neither be invented nor destroyed unless the authorised person is involved.

## 4. References


[1] Government of the Hong Kong Special Administrative Region, *Identity Management*, Hong Kong: Government of the Hong Kong Special Administrative Region, 2008.

[2] J.-M. Seigneur and T. E. Maliki, "Identity Management," in *Computer and Information Security Handbook*, Burlington, Morgan Kaufmann, 2009, pp. 269-297.

[3] D. Kunda and M. Chishimba, "A Survey of Android Mobile Phone Authentication Schemes," *Mobile Networks and Applications - Springer*, pp. 1-9, 2018.

[4] R. Bhasker and B. Kapoor, "Information Technology Security Management," in *Computer and Information Security Handbook*, Burlingtone, Morgan Kaufmann Publishers is an imprint of Elsevier, 2009, pp. 259 - 267.

[5] M. Rouse, "Tech Target," Tech Target, September 2006. [Online]. Available: https://searchsoftwarequality.techtarget.com/definition/Higgins-Trust-Framework. [Accessed 14 October 2018].

[6] S. Shahnewaz, "M2SYS," M2SYS, [Online]. Available: http://www.m2sys.com/blog/workforce-management/the-top-seven-advantages-of-a-biometric-identification-management-system/. [Accessed 11 11 2018].

[7] L. Musambo, M. Chinyemba and J. Phiri, "Identifying Botnets Intrusion and Prevention," *Zambia ICT Journal*, vol. 1, no. 1, pp. 63-68, 2017.

[8] P. Alok and R. S. Jatinderkumar, "A Simplified Defense Mechanism Against Man-In-The-Middle Attacks," *International Journal of Engineering Innovation & Research*, vol. 1, no. 5, pp. 1-5, 2012.

[9] W. Stallings, Cryptography and Network Security Principles and Practices, Fourth Edition, London: Pearson Education, Inc., 2005.

[10] I. I, R. A. P. M and V. Bhaskar, "Identity and access management in cloud environment: Mechanisms and challenges," Elsevier, San Francisco, 2018.

[11] M. Chinyemba and J. Phiri, "Gaps in the Management and Use ofBiometric Data: A Case of Zambian Public and Private Institutions," *Zambia Information Communication Technology (ICT) Journal*, vol. 2, no. 1, pp. 1-9, 2018.

[12] R. Bhasker and B. Kapoor, "Computer and Information Security Hand Book," in *Computer and Information and System Security*, Burlington, Morgan Kaufmann Publishers is an imprint of Elsevier, 2009, pp. 259-257.

[13] A. Parkar, S. Sharma and S. Yadav, "Introduction to Deep Web," *International Research Journal of Engineering and Technology (IRJET) e*, vol. 4, no. 6, pp. 1-4, 2017.

[14] K. Christian, B. Katja, T. Markus, H. Stephan and R. Kai, "How to Enhance Privacy and Identity Management for Mobile Communities: Approach and User Driven Concepts of the PICOS Project," *Mobile Business & Multilateral Security*, 2010.

[15] K. I-Lung, "Securing mobile devices in the business environment," *IBM Global Technology Services; Thought Leadership White Paper*, pp. 2-10, October 2011.

[16] M. Ahmad and J. Parvez, "A Novel Strategy to Enhance the Android Security Framework," *International Journal of Computer Applications (0975 – 8887)*, vol. 91, no. 8, pp. 1-5, 2014.

[17] K. Kathirvel, "Credit Card Frauds and Measures to Detect and Prevent Them," *International Journal of Marketing, Financial Services & Management Research*, vol. 2, no. 3, pp. 1-8, 2013.

[18] M. S. Gaigole and M. A. Kalyankar, "The Study of Network Security with Its Penetrating Attacks and Possible Security Mechanisms," *International Journal of Computer Science and Mobile Computing*, vol. 4, no. 5, p. 729, 2015.

[19] R. P, R. G, S. K and S. R, "Defending Man In The Middle Attacks," *International Research Journal of Engineering and Technology (IRJET)*, vol. 4, no. 3, pp. 579-585, 2017.

[20] Laws.com, "Hacking," Laws.com, 2017. [Online]. Available: https://cyber.laws.com/hacking. [Accessed 03 11 2017].

# A Review of System Intrusion Prevention Techniques and Tools in Developing Countries

Yvonne Nalukui Akende
Department of Electrical Engineering
*University of Zambia*
Lusaka, Zambia
yvonne.akende@MLNR.GOV.ZM

Jackson Phiri
Department of Computer Science
*University of Zambia*
Lusaka, Zambia
jackson.phiri@cs.unza.zm

***Abstract* — The area of intrusion detection and prevention is the central concept in overall network and computer security architecture. With the expansion of the internet and e-commerce over the years, we see individuals, governments and businesses having online presence and creating huge investments onto online platforms in developing countries. In the recent past, we witness these critical infrastructures becoming prime targets and more vulnerable to major cyber-attacks than ever before. Over a million compromises of electronic data occur annually worldwide, ranging from simple system intrusions to more sophisticated and malicious attacks resulting in huge losses for businesses.**

**This paper outlines the review of the current situation as it pertains to System Intrusions and Prevention from the developing countries perspective, pin pointing the general issues and concerns in leveraging the full benefits of system intrusion prevention techniques and mechanisms discussed in this paper. The analysis of these issues can be utilized to suggest the future direction of System Intrusion Prevention and cyber security as a whole.**

***Keywords*—** *Network Intrusion Prevention, Intrusion, Intrusion Prevention Systems, Access control*

## I. INTRODUCTION

Not only has the internet become critical to the wellbeing of many people and organizations, It has also become part of several components that are critical to the wellbeing of national economies and society at large [13]. Businesses and Governments have taken advantage of the internet to provide service delivery in many countries.

Although the internet provides huge business and individual benefits, it is threatened by many risks of adverse effects on businesses, governments and individuals. These risks have the potential to harm a country's economic growth. Reports have shown that, the security breaches of information stored into ICT assets have remained difficult to solve in that major threat activities on the internet are associated with the exploitation of personal and corporate information and other assets [1], [8]. Apart from physical theft and stolen cards, the problem of unauthorized access also called intrusion to networks and systems using advanced technologies have become a big challenge in the management of information security systems.

An intrusion is an unauthorized penetration of a computer in your enterprise or an address in your assigned domain [1]. An intrusion can be passive in which penetration is gained stealthily and without detection or active in which changes to network resources are effected. Intrusions can come from outside your network structure or inside.

Whatever the goal for an intrusion, whether for malicious purposes, seeking to steal critical information on either a one-time basis or as an ongoing parasitic relationship that will continue to siphon off data, to implanting carefully crafted code designed to crack passwords, record keystrokes, or mimic your site while directing unaware users to their site, a weakness in the security of your network has been detected and exploited. And unless you discover that weakness, the intrusion entry point, it will continue to be an open door into your network environment.

Intrusion Detection was developed to identify and report the intrusion attacks by detecting hostile traffic and send alerts but did nothing to stop the attacks, which became since a major issue in network security as confirmed by several writers [1], [2], [10] and [11]. In other words, Intrusion Detection is passive. It is not able to detect all malicious programmes and activities most of the time are incompatible to integrate with control restriction to stop traffic inbound-outbound from attacking; which means it was only capable to detect attack actions, without prevention action.

Intrusion Prevention System (IPS) is primarily a network-based defense system, with increasing global network connectivity and combines the technique firewall with that of the IDS properly with proactive technique. IPSs are a proactive technique which prevents attacks before entering the network by examining various data record and detects behavior pattern recognition sensor [2],[10],[11]. When an attack is identified, intrusion prevention blocks and logs the offending data. This basic of identifying and recognizing threat with high accuracy, earliness, and active response mainly concerns enabling comprehensive attack coverage available that must exists at present.

[2] and [10] have supported a similar requirement for a system to provide early detection / warning from intrusion security violation with knowledge based as a current necessity in preventing intrusions.

## II. INTRUSION TOOLS

With the increasing quantity and sophistication of attacks Crackers today are armed with an increasingly sophisticated and well-stocked tool kit for doing what they do. Crackers today can obtain a frightening array of tools to covertly test your network for weak spots. Some of their widely used tools include the following:

### *a) Wireless sniffers*

Easy accessibility condition in wireless networks causes more vulnerability against wired networks. Not only can these devices locate wireless signals within a

certain range, they can siphon off the data being transmitted over the signals. With the rise in popularity and use of remote wireless devices, this practice is increasingly responsible for the loss of critical data and represents a significant headache for IT departments [1], [12].

*b) Packet sniffers*

Once implanted in a network data stream, these passively analyze data packets moving into and out of a network interface, and utilities capture data packets passing through a network interface.

*c) Port scanners*

These utilities send out successive, sequential connection requests to a target system's ports to see which one responds or is open to the request. Some port scanners allow the cracker to slow the rate of port scanning sending connection requests over a longer period of time so the intrusion attempt is less likely to be noticed. Common targets are old, forgotten back doors, or ports inadvertently left unguarded after network modifications.

*d) Port knocking*

Like Port scanners, Port-knocking tools find unprotected entries which are sometimes created by network administrators as secret back-door methods of accesing firewall-protected ports and implant a Trojan horse that listens to network traffic for evidence of that secret knock.

*e) Keystroke loggers.*

Spyware utilities planted on vulnerable systems that record a user's keystrokes to obtain things like usernames, passwords, and ID numbers.

*f) Remote administration tools*

Programs embedded on an unsuspecting user's system that allow the cracker to take control of that system.

*g) Network scanners*

Explore networks to see the number and kind of host systems on a network, the services available, the host's operating system, and the type of packet filtering or firewalls being used.

*h) Password crackers*

Sniff networks for data streams associated with passwords. Poor passwords can be detected by dictionary or bruteforce attacks and then employ a method of peeling away any encryption layers protecting those passwords.

## III. Symptoms of Intrusion

Crackers first look for known weaknesses in the operating system (OS) or any applications you are using to identify your network's vulnerability. Common indicators for an attack include.

*1) Large numbers of unsuccessful login attempts*

A large number of failed login attempts is a good indicator that your system has been targeted. The best penetration-testing tools can be configured with attempt thresholds that, when exceeded, will trigger an alert. They can passively distinguish between legitimate and suspicious activity of a repetitive nature, monitor the time intervals between activities (alerting when the number exceeds the threshold you set), and build a database of signatures seen multiple times over a given period. A sequence of mistyped commands or incorrect login responses with attempts to recover or reuse them can be a signs of brute-force intrusion attempts.

*2) Packet inconsistencies*

Packet direction (inbound or outbound) originating address or location, and session characteristics (ingoing sessions vs. outgoing sessions) can also be good indicators of an attack. Unusual source, an abnormal port address or an inconsistent service request could be a sign of random system scanning. Packets coming from the outside that have local network addresses that request services on the inside can be a sign that IP spoofing is being attempted.

*3) Odd or unexpected system behavior*

Though this is sometimes difficult to track, you should be aware of activity such as changes to system clocks, servers going down or server processes inexplicably stopping with system restart attempts, system resource issues such as unusually high CPU activity or overflows in file systems, audit logs behaving in strange ways, decreasing in size without administrator intervention, or unexpected user access to resources. Unusual activity at regular times on given days, heavy system use can be possible DoS attack or high CPU activity can indicate a bruteforce password-cracking attack and should always be investigated.

## IV. Network Security Best Practices

Striking an acceptable balance between keeping your network intrusion free, having a good defensive posture that still allows for access is key. It is Important to constantly audit your defenses to ensure that your network's defensive armor can meet the latest threat.

### A. *Security Policies*

Organizations need a good, detailed, and well-written security policy. The information security policy serves as a tool to provide guidance on how to manage and secure all business operations including critical assets, infrastructure and people in the organization. This guidance (e.g. usage and controls) facilitates the provisions for threat assessment and compliance based on local context [3][9]. This policy is always a work in progress and must evolve with technology, as threats continue to evolve, as will the systems designed to hold them at bay [1].

Furthermore, the three writers [1], [3] and [9] agree that a good security policy must be a conglomeration of policies that address specific areas, comprehensive and easily understood by those it affects. In as much as security policies are, Implementation and enforcement of security policies in most organisations are weak

It must spell out responsibilities for security requirements, communicate organisational expectations and lay out the role(s) for your network users. Security teams must strive to create security policies that are practical, workable, and sustainable and come up with the best plan for implementing these policies in a way that addresses both network resource protection and user friendliness. Their responsibility also extends to developing Plans for responding to threats as well as schedules for updating equipment and software and handling changes to overall network security.

Current security products are based on policies that are defined by the security team. These teams in most organisations are not inhouse, which means an added cost of consultation in meeting the demands of good security policies in organisations. The lack of in house expertise and the cost of consultation have defeated the customisation of policies to match the rapid evolving threat and technology landscape.

### B. *Risk Analysis*

You should have some kind of risk analysis done to determine, as near as possible, the risks you face with the kind of operations you conduct. Depending on the determined risk, you might need to rethink your original network design.

Lack or Unsupported IT Governance Frameworks. No Risk Management plans in place.

### C. *Vulnerability Testing*

Your security policy should include regular vulnerability testing. Some very good vulnerability testing tools, allow you to conduct your own security testing. Furthermore, there are third-party companies with the most advanced suite of testing tools available that can be contracted to scan your network for vulnerabilities.

The problem is not only having the vulnerabilities discovered, and having to provide patches for them, it is also with vendors that do not release these updates fast enough. Also constantly installing software patches when they are released is not always easy in a large system environment since patches can create other problems e.g compatibility problems, which requires expertise knowledge which is not always available within the organisation.

### D. *Audits*

Factor in regular, detailed audits of all activities, with emphasis on those that seem to be near or outside established norms. For example, audits that reveal high rates of data exchanges after normal business hours, when that kind of traffic would not normally be expected, is something that should be investigated.

### E. *Recovery*

A sound IT security recovery plan is more important in addressing the issue of recovery after an attack has occurred. You need to address issues such as how the network will be reconfigured to close off the exploited opening. This might take some time, since the entry point might not be immediately discernable.

There has to be an estimate of damage — what was taken or compromised, was malicious code implanted somewhere, and, if so, how to most efficiently extract it and clean the affected system. In the case of a virus in a company's email system, the ability to send and receive email could be halted for days while infected systems are rebuilt. And there will have to be discussions about how to reconstruct the network if the attack decimated files and systems.

This will most likely involve more than simply reinstalling machines from archived backups. Because the compromise will most likely affect normal business operations, the need to expedite the recovery will hamper efforts to fully analyze just what happened.

This is the main reason for preemptively writing a disaster recovery plan and making sure that all departments are represented in its drafting. However, like the network security policy itself, the disaster recovery plan will also be a work in progress that should be reviewed regularly to ensure that it meets the current needs.

Things such as new threat notifications, software patches and updates, vulnerability assessments, new application rollouts, and employee turnover all have to be addressed.

### F. *Educate your users*

No matter how good a job you've done at tightening up your network security processes and systems, you still have to deal with the weakest link in your armor, your users [1]. It doesn't do any good to have bulletproof processes in place if they're so difficult to manage that users work around them to avoid the difficulty, or if they're so loosely configured that a casually surfing user who visits an infected site will pass that infection along to your network.

The degree of difficulty in securing your network increases dramatically as the number of users goes up. Management views security as a drain of resources. Establishment and Placement of security personnel in organisational structures to drive security company wide security goals and policies at management level still remains a challenge. Organisations have not proritised inhouse capacity building and development of security staff as services are mostly outsourced. Lack of security trained personnel has resulted in most organisations not having security education user programmes.

## V. Tools for Intrusion Prevention

In today's rapidly changing software environment, strong security requires penetration shielding, threat signature recognition, autonomous reaction to identified threats, and the ability to upgrade your tools as the need arises.

Though no one solution can offer comprehensive security, the following discussion talks about some of the more common tools you should consider adding for preventing intrusions.

### A. *Firewalls*

The primary goal of a firewall is to protect the network behind it, and act as your first line of defense. Ssecurity/systems administrator have to configure and manage new firewalls to realize an appropriate security policy for the particular needs of the organisation [7]. [1] Recommends a Firewall that combines the five most necessary security systems, firewall, antivirus/spyware/spam, virtual private network (VPN), application filtering, and intrusion prevention/detection systems into a single appliance as most secure.

The days of relying solely on a firewall is gone; today's crackers have figured out how to exploit weaknesses in firewalls to bypass security. Another problem inherent with firewalls is misconfigurations that is presented by lack of capacity in such technologies.

### B. *Intrusion Prevention Systems (IPS)*

According to [1] a good intrusion prevention system (IPS) is a vast improvement over a basic firewall as it can be configured with policies that allow it to make autonomous decisions as to how to deal with application-level threats as well as simple IP address or port-level attacks. [3], [4] and [11] in their work agree that early detection, protection and response system is the main concept of IPS and is expanded on the functionality provided by IDS. It utilises IDS algoorithms too monitor and drop, or allow traffic based on expert analysis. The main idea is to be proactive. This early detection and intrusion response has the fundamental and part of intrusion prevention mechanism in recent network security challenge, as confirmed by several works [3], [4] and [5].

IPS can be considered another form of access control in that it can make pass/fail decisions on application content as mentioned above. IPSs must also be very good at discriminating between a real threat signature and one that looks like but isn't one (false positive). Accuracy in intrusion prevention a positive alarm is considered as an attack data, while a negative is considered to be a normal data. Additionally, more appropriately accurate mechanism keeps the number of false negative and false positive low as in work by [3], [6]. Once a signature interpreted to be an intrusion is detected, the alert management system must generate an alert for appropriate evasive action to be taken. Work by Anuar in 2010 [5], declared intrusion response as having similar function to IDS and part of it, by maintaining detection, alerting and response to security operator.

#### Features of IPS

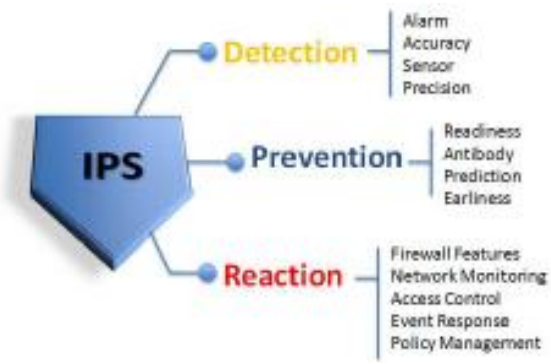



Figure 1: Features of IPS

#### Types of IPS

*1) Network-based*

Network-based IPSs create a series of choke points in the enterprise that detect suspected intrusion attempt activity. Placed inline at their needed locations, they invisibly monitor network traffic for particular network segments or devices and analyzes the network and application protocol activity to identify suspicious activity [14].

*2) Host-based*

Monitors the characteristics of a single host and the events occurring within that host for suspicious activity. They quietly monitor network traffic, system logs, running processes, application activity, file access and modification, and system configuration changes. Host-based are deployed on critical hosts such as publicly accessible servers and servers containing sensitive information. These systems are often very good at identifying post-decryption entry attempts [1], [12].

*3) Content-based*

These IPSs scan network packets, looking for signatures of content that is unknown or unrecognized or that has been explicitly labeled threatening in nature [1].

*4) Rate-based*

These IPSs look for activity that falls outside the range of normal levels, such as activity that seems to be related to password cracking and brute-force penetration attempts, for example. Protection against brute-force and DoS attacks application scanning, and flooding A regular method of updating threat lists and signatures [1].

### C. Access Control Systems

Access control is one of the mutually supportive technologies that are used in information security and management to achieve confidentiality and integrity. Authentication, Access Control and auditing provide the foundation for information and system security. Access control systems (ACSs) rely on administrator defined rules that allow or restrict user access to protected system resources. These access rules can, for example, require strong user authentication such as tokens or biometric devices to prove the identity of users requesting access.

Implementation of Access Control can be through Access Control List (ACL), which is a set of rules that define security policy. These ACLs contain one or more access control entries (ACEs), which are the actual rule definitions themselves. These rules can restrict access by specific user, time of day, IP address, function or specific system from which a logon or access attempt is being made [1].

Access control is behavior-based, that is it defines what is normal. One problem with these methods is to specify what is normal and what is not. Therefore the effectiveness of the access control rests on a proper user identification and on the correctness of the authorizations governing the reference monitor.

Access control is not a complete solution for securing a system. It must be coupled with auditing, which allows to control penetrations in which the attacker gains privileged status [16].

### D. Unified Threat Management

The latest trend to emerge in the network intrusion prevention arena are Unified Threat Management (UTM), a multilayered system that incorporate several security technologies into a single platform, often in the form of a plug-in appliance. UTM products can provide such diverse capabilities as firewall, Virtual Private Network (VPN), trusted source, Intrusion Prevention Systems, URL filtering, SSL decryption, and auditing/reporting as well as antivirus and antispam. The biggest advantage of a UTM system is its ease of operation and configuration and the fact that its security features can be quickly updated to meet rapidly evolving threats.

Though unified integrated systems are perceived as a comprehensive solution, other writers like [2] have revealed that a number of unresolved problems still exist with these integrated solutions; their effectiveness in detecting normal usages and malicious activities using heterogeneous data. The real challenge of intrusion prevention lies in the handling of unclear situations, suspicious but not conclusively malicious traffic.

## VI. Recommendations and Conclusion

Information security is an ongoing process which is influenced by many factors. While some of these factors are internal and under local supervision, many of them are external and hard to evaluate.

Network and information security is critical to cyber security implementation and many developed countries have developed and implemented cyber security protocols, Laws, Standards, Regulations and implementations for internet security and policing. In Zambia, the concept of cyber security and information security is still vague and the Internet being a borderless globally connected network, we are not immune to the risks and threats over the internet.

A Strong cyber capacity is crucial for states to progress and develop in economic, political and social spheres. In Zambia organizations must prioritize the need to integrate cyber capacity building and development policies, implementing these capacities in practice in order to achieve broader security goals [14]. This will reduce the dependence on outsourcing security products and services.

The basic of identifying and recognising threat with high accuracy, earliness, and active response mainly concerns enabling comprehensive attack coverage available that must exists in every organization. Several writers have highly recommended the requirement for a system to provide early detection or warning from intrusion and security violation with knowledge based as a current necessity in preventing intrusions [14][15]..

Unified integrated solutions could be an effective solution for building integrated systems in the industrial world. We need a coordinated approach which combines the strong points of each of the above technologies, rather than treating them as separate independent solutions [16].

.

## References

[1] J. R Vacca, 2009. *Computer and Information Security Handbook*. Newnes

[2] D. Stiawan, A. Y. I. Shakhatreh, M. Y. Idris, K. Abu Bakar and A. H. Abdulla, 2012. “Intrusion Prevention System: A Survey. A Journal of Theoretical and Applied Information Technology”, Vol. 40, No. 1

[3] C. Mu, B. Shuai and H. Liu, “Analysis of Response Factors in Intrusion Response Decision-Making,” 2010 Third International Joint Conference on Computational Science and Optimization, 2010, pp. 395-399.

[4] N. Stakhanova, S. Basu, and J. Wong, “A taxonomy of intrusion response systems”, International Journal and Computer Security, vol. 1, 2007, pp. 169-184. N.B.

[5] A. M. Papadaki, S. Furnell and N. Clarke, “An investigation and survey of response options for Intrusion Response Systems ( IRSs ),” Information Security for South Africa (ISSA), 2010, pp. 1-8.

[6] A. D. Todd, R. A. Raines, R.O. Baldwin, B. E. Mullins, and S.K. Rogers, 2007. “Alert Verification Evasion Through Server Response Forging,” vol. 4637/2007.

[7] A. Wool, “The use and usability of direction-based filtering in firewalls”, Computers & Security, vol. 23, Sep. 2004.

[8] DBIR (2014) 2014 “Data Breach Investigation Report”, Verizon Document, Tech. Rep.

[9] J. E. Mbowe, I. Zlotnikova, S. S. Msanjila and G. S. Oreku, “A Conceptual Framework for Threat Assessment Based on Organization’s Information Security Policy; A journal of Security Information, 2014” . http://dx.doi.org/10.4236/jis.2014.54016

[10] G. Ibrahim, H. Martin and P. Vaclav, “A Survey on Intrusion Detection and Prevention Systems (2014)”.

[11] M. Barkett, “Intrusion Prevention Systems,” NFR Security, Inc., 2004. http://www.nfr.com/resource/downloads/SentivistIPS-WP.pdf

[12] I. Mukhopadhyay, M. Chakraborty and S.Chakrabarti, “A Comparative Study of Related Technologies of Intrusion Detection & Prevention Systems; A journal of information Security”, 2011. http://www.SciRP.org/journal/jis.

[13] Kortjan, Noluxolo & Solms, Rossouw. (2013). Cyber Security Education in Developing Countries: A South African Perspective. 289-297. 10.1007/978-3-642-41178-6_30.

[14] L. K. Musambo, M. K. Chinyemba, J. Phiri, “Identifying Botnets Intrusion & Prevention – A Review. Zambia ICT Journal, [S.l.], v. 1, n. 1, p. 63-68, dec. 2017. ISSN 2616-2156. URL: <http://ictjournal.icict.org.zm/index.php/zictjournal/article/view/28>. Date accessed: 20 Nov. 2018.

[15] Wakwinji Inambao, Jackson Phiri and Douglas Kunda, "Digital Identity Modelling For Digital Financial Services In Zambia". ICTACT Journal On Communication Technology, Volume 09, Issue No. 03, September, 2018, DOI: 10.21917/ijct.2018.0267

[16] Johnson I Agbinya, Nazia Mastali, Rumana Islam and Jackson Phiri, "Design and Implementation of a Multimodal Digital Identity Management system using fingerprint matching and face recognition," 6th International Conference on Broadband Communications & Biomedical Applications (IB2Com), pp. 272-278, 21-24 Nov 2011.

# Security, Privacy and Integrity in Internet of Things – A Review

*Chalwe Musonda [1], Monica M.K. –Kabemba[2], Mayumbo Nyirenda[3], Jackson Phiri[4]*
*The University Of Zambia, Department of Computer Science, Lusaka, Zambia*
*Chalwe.musond@cs.unza.zm[1], Monica.kabemba@cs.unza.zm[2], m.nyirenda@unza.zm[3], jackson.phiri@cs.unza.zm[4]*

***Abstract* - This paper addresses the Data Security, Privacy and Integrity in Internet of Things (IoTs) it borrows and characterized by heterogeneous technologies, which concur to the provisioning of innovative services in various application domains. In this scenario, the satisfaction of security and privacy requirements plays a fundamental role in information system security. Such requirements include data confidentiality and authentication, access control within the IoT network, privacy and trust among users and things which are connected on the internet and the enforcement of security, privacy and integrity policies, which exist as at now. Traditional security issues countermeasures cannot be directly applied to IoT technologies due to the different standards and communication stacks involved in the technology. Moreover, the high number of interconnected devices arises scalability issues; therefore a flexible infrastructure is needed so that its able to deal with security threats in such a dynamic environment. Different vision of this Internet of Things paradigm are reported and taken researched and reviewed. And what comes out is the major issues which shall be faced by the research community. There is more research to be done in this area especially of distributed intelligence for smart objects are just the most relevant.**

## I. INTRODUCTION

There are a number of security risks associated with any electronic device that connects to a network (internet or mobile phone networks being the most common). This paper is not addressing the issues involved in the following Hacking, Phishing, Smashing, Vising, Pharming, Spyware, Viruses, and Spam. But it will concentrate on the security, privacy and integrity of IoTs. The term, internet of things (IoT) that refers to uniquely identifiable objects, thins, and their virtual representations in an internet-like structure, was first proposed in 1998 as in coming to existence. Before getting to the details of this paper below is the diagram (See Figure 1) to help in the understanding of the topic's terms and how there are linked together.

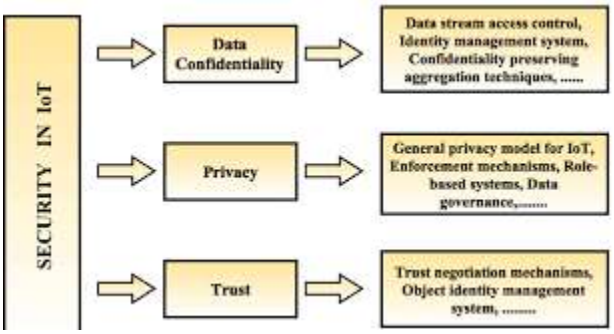



*Figure 1: Showing the diagram of the security issues in IoT. Privacy and Integrity cardinal in this discussion.*

## II. WHAT IS DATA SECURITY?

Data security is a point to take into considerations in that it's a requirement for data to be recoverable if lost or corrupted. Data Security can also be defined as a process of protecting files, databases, and accounts on a network by adopting a set of controls, applications, firewall [1] and techniques that identify the relative importance of different datasets, their sensitivity, regulatory compliance requirements and then applying appropriate protections to secure those resources. Similar to other approaches like perimeter security, file security or user behavioral security, data security is not the all end all for a security practice. It's one method of evaluating and reducing the risk that comes with storing any kind of data in the system.

### A. *Why Data Security?*

If Data Security process is just one of many different ways to structure your organization's information security systems and in the IoTs, what makes it better than competing methods when there are more devices to control? Broadly speaking, most other security processes are "user-centric": they focus on questions like:

- Is this user allowed to access this particular data?
- Is this person authorized to be on this network?
- Is this person abusing system resources?

Which is great and necessary but struggles with many real-world issues like large organizations having hundreds or thousands of servers with haphazardly applied permissions, antiquated user groups and gaps in knowing who is accessing what in the network/system. A data-centric security model is a practical way of approaching this from a different direction, of questions of who is allowed to access data, who is authorized and whether the person is abusing the resources.

### B. *Data vs. User Security Models*

Imagine a scenario where a user on your customer service team places a spreadsheet containing customer Personally Identifiable Information like Social Security Numbers or other sensitive records onto a globally accessible shared folder. When it comes to User Centric Model: this wouldn't be a problem, everyone has the proper rights to access that file. So when it comes to IoT it works well if more device are connected in the network. Data Security Model: this is a huge problem as sensitive information is now available to every intern, contractor or "coasting through their tenure until they take a new job at your biggest competitor" employee with network access. This scenario makes plain the big dependency of a Data Security approach: data classification. More in the IoTs and organizations.

## III. INTEGRITY

It is easy to define integrity of data but far less easy to ensure it. Only accurate and up-to-date data has data integrity. Any person or organization that stores data needs it to have integrity. Methods [2] that can be used to give the best chance of achieving data integrity is not covered in this paper. In a simple way, data integrity: a requirement for data to be accurate and up to date. Integrity protection includes preservation against sabotage and the use of counterfeit units or components. Another critical factor that influences data integrity is the robustness and fault tolerance capabilities of the IoT System. Sensor networks, such as RFID solutions, face also other issues that limit their capability to overcome integrity problems as many of their components spend most of the time without being attended to. Attackers can either modify the data while it is stored in the node or when it travels through the network. Read and write protections as well as authentication methods are common solutions to these issues.

Data integrity is also ensured by password-based solutions, which brings into account the shortcomings of password protection, such as vulnerabilities related to password length and randomness. Also, the resources found in common IoT systems do not support typical cryptographic solutions because of the limited resources available. Integrity for the Internet of Things not only is required to be guarded from external sources but also for internal processes, such as service integrity. Operating systems rigid process separation, known as Multi Level Security (MLS), help devices to avoid unauthorized modification from code running with high privileges. Nevertheless, MLS approaches have not been deployed widely as in some cases can be considered as expensive as well as not compatible with other IoTs software. Other approaches to guarantee integrity use hash values which are stored externally to avoid compromises. Hardware solutions have also been proposed for integrity purposes, a challenge-based solution is mentioned in by the use of symmetric or asymmetric keys known as Trusted Platform Module (TPM). Process integrity is also required by IoT devices, process integrity relies on the device, communication, and algorithm implementation integrity.

## IV. PRIVACY

Data privacy is a requirement for data to be available only to authorized users. Data privacy is about keeping data private rather than allowing it to be available in the public domain. The term 'data privacy' may be applied to a person or an organization. Each individual has an almost limitless amount of data associated with their existence. Assuming that an individual is not engaged in criminal or subversive activities, he or she should be in control of which data about himself or herself is made public and which data remains private. An organization can have data that is private to the organization, such as the minutes of management meetings and financial reports and business plans. For an individual there is little chance of data privacy if there is not a legal framework in place to penalize offenders who breach this privacy. Such laws are referred to as data protection laws and that depend on the particular country. The major aspects of data protection laws relate to personal, therefore private, data that an individual supplies to an organization. The data is supplied to allow the organization to use it but only for purposes understood and agreed by the individual. Data protection laws oblige organizations to ensure the privacy and the integrity of this data. Unfortunately having laws does not guarantee adherence to them but they do act as a deterrent if wrong-doers can be subject to legal proceedings. The IoTs are not bound by the protection laws and privacy because it uses its now environment and it directly connected to the device in the environment using the internet.

### A. *What is Data Privacy?*

Data Privacy is the branch of information security dealing with the proper handling of data concerning consent, notice, sensitivity, and regulatory concerns. Practical data privacy problems often revolve around:

1. Whether or how data can be shared with third parties.
2. If data can legally be collected or stored.

### B. *Data Security and Data Privacy what's the difference?*

Data Security and Data Privacy are often used interchangeably in many situations, but there are distinct differences: Data Security can broadly be thought of as protecting the data on your network from outsiders (and malicious insiders). Data Privacy governs how the data is collected, shared and used. Consider data that you consider to be solidly secured: it's encrypted, access to it is restricted, and multiple overlapping monitoring systems are in place. In all meaningful senses of the word, the data is secure. However, if that data was collected without proper consent that is a violation of data privacy and distinct from the actual security surrounding the data. Data privacy revolves around making sure that data is used in the correct manner with the right persons.

### C. *Data Privacy Principles*

Cavoukian, the former Information & Privacy Commissioner of Ontario, Canada says, "Privacy forms the basis of our freedom. You need to have moments of reserve, reflection, intimacy, and solitude." It is only through the freedom of play and experimentation that innovation and new ideas can emerge. You don't want to be such company to be described as creepy in the way that you leverage your customer's personal data – whether it is with passive location tracking, apps secretly absorbing your personal address book, or websites recording your every keystroke. Instead, employees should be regularly trained in security and privacy, so they understand the processes and procedures necessary to also ensure proper collection, sharing, and use of sensitive data of an organization.

## V. INTERNET OF THINGS OF TODAY

The Internet of Things (IoT) is promising to open up many new opportunities and avenues for businesses to offer new and exciting services by using the internet. However, with the myriad of devices [3] and business assets connected open to the internet, the need for a strict and reliable approach to security is essential. One such security model that has been in use for a long time is that of the CIA triad, which is not of interest in this paper. Nevertheless, it's of important to our discussion because it sets to us our basis of security. See figure 2 below:


*Figure 2 – CIA Model – security policy*

The key components of this model are based around confidentially, integrity and availability. In the context of IoT, confidentiality caters for protecting privacy of IoT devices, integrity looks after the data contained within the device while availability covers accessibility of the device. When we think about data of an IoT device we should think about not only the data being generated or used by it but also its own programming data, this including all aspects of program software, configuration parameters and operating system software. To guide the process of integrity it is helpful to consider three different states that data can exist, namely in motion, at rest and in process. Any breach of data integrity will mean that an IoT device cannot operate correctly but it also potentially exposes the device to being exploited and become a compromised platform from which other attacks can be launched. The usual method of verifying the integrity of data is by a mathematical algorithm called a hash, of which the secure hash algorithms (SHA) is most popular.

Data-in-motion requires that data be protected from modification while on its journey from sensor to cloud application. While a hash technique can be used an attacker could make a change to the message and recalculate the hash. A stronger approach is by using a data integrity check with a shared private key.

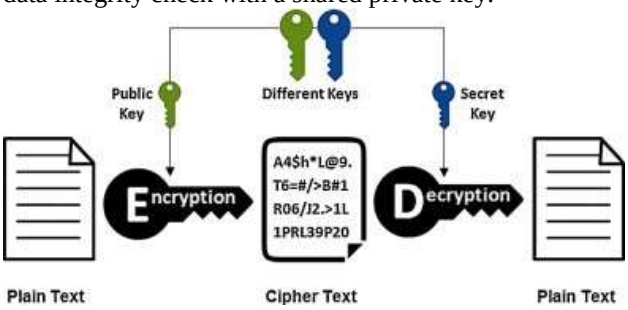



*Figure 3.– Showing the Shared private key Technology (keyed-hash message)*

This is called a keyed-hash message authentication code (HMAC), and since it needs a shared private key it must

be protected just like any other cryptographic key. If the IoT device has reliable, always-on internet connectivity, then remote attestation can be used to report and validate key boot parameters during the secure boot and trusted boot phases to a physically separate server. These parameters are typically a hash of the components of the boot process (bootloader, applications, etc.). The server than compares these measurements and determines the trustworthiness of the IoT device. The transmission of the measurements is secured and must include the identity of the IoT device.

## *VI. DATA VALIDATION AND VERIFICATION*

Data integrity can never be absolutely guaranteed but the chances are improved if appropriate measures are taken when data originally enters a system or when it is transmitted from one system to another more especially in the issues of IoT.

### *A. Validation and verification of data entry*

The term validation is a somewhat misleading one. It seems to imply that data is accurate if it has been validated. This is far from the truth. If entry of a name is expected but the wrong name is entered, it will be recognized as a name and therefore accepted as valid. Validation can only prevent incorrect data if there is an attempt to input data that is of the wrong type, in the wrong format or out of range. Data validation is implemented by software associated with a data entry interface. There are a number of different types of check that can be made. Typical examples are:

- a presence check to ensure that an entry field is not left blank
- a type check, for example only a numeric value for the month in a date.

Verification of data means confirming what has been entered. The most common example is when a user is asked to supply a new password. There will always be a request for the password to be re-entered. Clearly, if the user entered a password but did not enter it as intended, subsequent attempts at access would fail. Verification is usually an effective process but in general it does not ensure data accuracy because the wrong data could be entered initially and in the re-entry.

### *B. Verification during data transfer*

It is possible for data to be corrupted during transmission. Typically this applies at the bit level with an individual bit being flipped from 1 to 0 or vice versa. Verification techniques need to check on some property associated with the bit pattern.

The simplest approach is to use a simple one-bit parity check. This is particularly easy to implement if data is transferred in bytes using a seven-bit code. Either even or odd parity can be implemented in the eighth bit of the byte. Assuming even parity, the procedure is:

1. At the transmitting end, the number of ls in the seven -bit code is counted.
2. If the count gives an even number, the parity bit is set to 0.
3. If the count gives an odd number, the parity bit is set to 1.
4. This is repeated for every byte in the transmission.
5. At the receiving end, the number of ls in the eight-bit code is counted.
6. If the count gives an even number, the byte is accepted.
7. This is repeated for every byte in the transmission.

If no errors are found, then transmission is accepted. However, the transmission cannot be guaranteed to be error free. It is possible for two bits to be flipped in an individual byte. Fortunately, this is rather unlikely so it is a sensible assumption to assume no error. The limitation of the method is that it can only detect the presence of an error. It cannot identify the actual bit that is in error. If an error is detected, re-transmission has to be requested. An alternative approach is to use the checksum method. In this case at the transmitting end a block is defined as a number of bytes. Then, irrespective of what the bytes represent, the bits in each byte are interpreted as a binary number. The sum of these binary numbers in a block is calculated and supplied as a checksum value in the transmission. This is repeated for each block. The receiver does the same calculation and checks the summation value with the checksum value transmitted for each block in turn. Once again an error can be detected but its position in the transmission cannot be determined. For a method to detect the exact position of an error and therefore be able to correct an error it has to be considerably more complex. A simple approach to this is the parity block check method. Like the checksum method this is a longitudinal parity check; it is used to check a serial sequence of binary digits contained in a number of bytes.

### *C. Security, privacy and Integrity Challenge in the IoTs*

Any internet connected device is at high risk, and IoT devices are no exception. Providing a further layering of integrity defenses comes under the heading of AAA (pronounced triple A) which stands for authentication, authorization and accounting. Its role is to determine which other devices on the network it can communicate with and what data can be transferred. A trusted party is needed to broker the communication between the IoTs device and the distant device or server. A well-established protocol called Kerberos can be used to establish this trust between devices over the network. Once authenticated, authorization follows. Authorization is the determination of the type of access allowed to a resource within the network. A network wide security policy is used to list

what the resources are, the paths of communication to and from each resource, and to what level the access can occur. The final component of AAA is that of accounting. This is the generation of a log of events that denote security-related activities on and by the IoT device. Event logs can be stored within the IoT device or, more usually, on an external server. The security policy will denote which events need to be recorded. Reviewing the logs and associated data will help determine what led to an attack, where it came from and what happened during the attack. It is best that these events are recorded as near to real-time as possible to minimize the damage from the attack in order that any initial, front-line response as defined in the security policy can be made. Because of the large number of IoT devices that typically must be monitored in the system, along with the large number of events that need to be parsed and managed, a specialized server is required. This type of server is called a security information and event management (SIEM) server. A SIEM server is able to correlate received security event messages from IoT devices and use predictive analytics to determine if an IoT device is at risk of an attack.

### *D. Privacy Challenges in the IoTs*

Today, online service consumers are aware that when they use free online services (email, social networking, and news feeds, for example), they automatically become data sources for businesses, which can analyze this data to improve customer satisfaction. Even worse, the data can be sold to third parties for further analysis. However, in the future IoT era, it's likely that service providers will adopt one of the two following models: some consumers might willingly pay to consume services with the aim of protecting their privacy; others might offer to give away data, under some limitations and conditions, in return for consuming services free of charge. Data collected through smart wearable and smart home devices can be used to generate contextually enriched information. Device owners should remain in charge of such data at all times, even if they give access to their data to external parties temporarily to accomplish a specific task. Consequently, the IoT era poses significant privacy challenges, especially due to the IoT's sheer scale. The EU Commission report on the IoT, has identified security and privacy as a major IoT research challenge that encompasses privacy-preserving technology for heterogeneous device sets; models for decentralized authentication and trust; energy-efficient encryption; data-protection technologies; security and trust for cloud computing; data ownership; legal and liability issues; repository data management; access and use rights; rules to share added value; responsibilities; liabilities; artificial immune system solutions for the IoT; secure, low-cost devices; integration with or connection to privacy- preserving frameworks; and privacy policies management.

## VII. WAYS THE IoT COULD BE HACKED

Machine Security researchers have identified a range of other frightening vulnerabilities. Researchers have "demonstrated ransom ware against home thermostats and exposed vulnerabilities in implanted medical devices. They've hacked voting machines and power plants." Indeed, many computer security experts fear that the USB port at an airline seat could potentially be used to control the plane's avionics Clearly, the IoT offers a broad array of dangerous tools hackers can employ for a wide range of motives, including: terrorism, "national aggression," pranking, election tampering, and monetary extortion, including the car control and household devices (SmartTv, Fridge etc). Whatever the impetus for hacking in the IoT, the threats moving forward are considerable.

Hacking in the IoT will often be illegal, though the existing laws in Zambia punish conduct after the fact without addressing the vulnerabilities that facilitate hacking. Hacking with the Intention of Controlling an Object Consider the following hypothetical. For example: Mr. Phiri has a grudge against his neighbor Mr. Mulenga. Mr. Phiri discovers a security vulnerability in one of the many electronic control units (ECUs) of Mr. Mulenga's late model sedan, and he hacks in through the internet and enters commands that enable him to take control of Mr. Mulenga's car. Mr. Phiri's actions are increasingly plausible as cars become ever more connected and automakers struggle to update outmoded software. The hypothetical identifies a fundamental aspect of the IoT: the hackers' target is not the computer, but the object connected to the computer and basically data. Though the motives varied: the fourteen-year-old hacked a train system for a prank in Japan; the Iranians hacked a dam apparently as an act of terrorism; the extortionists attacked the system for money; and the disgruntled employee hacked into cars sold by his former employer for revenge. All sought to achieve their goals by controlling a remotely accessible object in the IoT. In the IoT, a major objective of remote access will be to control the "things." Thus, a key question is whether the current legal regime covers this relatively new threat, governing scenarios like the one involving Mr. Phiri and Mr. Mulenga. Well, in Zambia it's included in the ICT policy but not that activated it needs the reinforcement and probably as devices in IoT becomes many.

## VIII. CONCLUSION

In conclusion, in order for us to benefit from the huge potential of connected IoT devices without too much worrying, there is a need for a strict and reliable approach to security, privacy and integrity are essential. The Confidentiality, Integrity and Availability (CIA) triad provides a very simple and convenient model of both discovering and representing the security needs of your IoT device today. Maintaining data integrity is certainly one of the key aspects of implementing a set of security

policies, and the same considerations and approach needs to be made for that of confidentiality, Integrity and availability. In the IoT, data owners must have full control of data and be able to delete or move data from one service provider to another at any time. Unfortunately, existing IoT solutions in the marketplace provide only limited access to users, until we reach at that level. Moreover, users should be able to choose hardware devices and software components from different vendors to build their smart environments (for example, a smart home that includes Smart Television, Fridge, Stove and many internet devices). This gives users full control and freedom of choice. Consequently, users must decide on what kind of data they will share, with what access rights for service providers. Users should also have the ability to withdraw or change previous user consents. And be able to switch or transfer data from one service provider to another.

## IX. REFERENCE


[1] Tsai, C., Lai, C., & Vasilakos, V. (2014). Future internet of things: Open issues and challenges. *ACM/Springer Wireless Networks*, doi: 10.1007/s11276-014-0731-0.Google Scholar

[2] Wan, J., Yan, H., Suo, H., & Li, F. (2011). Advances in cyber-physical systems research. *KSII Transactions on Internet and Information Systems, 5*(11), 1891–1908.CrossRefGoogle Scholar

[3] Sipiwe Chihana, Jackson Phiri and Douglas Kunda, "*An IoT based Warehouse Intrusion Detection (E-Perimeter) and Grain Tracking Model for Food Reserve Agency*" International Journal of Advanced Computer Science and Applications (IJACSA), 9(9), 2018. http://dx.doi.org/10.14569/IJACSA.2018.090929

[4] Hachem, S., Teixeira, T., & Issarny, V. (2011). *Ontologies for the internet of things* (pp. 1–6). New York: ACM. Google Scholar

[5] Sundmaeker, H., Guillemin, P., Friess, P., & Woelfflé, S. (2010). *Vision and challenges for realising the internet of things*. *Cluster of European Research Projects on the Internet of Things—CERP IoT*.Google Scholar

[6] Hamad, F., Smalov, L., & James, A. (2009). Energy-aware security in M-Commerce and the internet of things. *IETE TechmeM review*, *26*(5), 357–362.CrossRefGoogle Scholar

[7] Tsudik, G. (2006). YA-TRAP: Yet another trivial RFID authentication protocol. In *Proceedings of fourth annual IEEE international conference on pervasive computing and communications workshops* (pp. 196–200).Google Scholar

[6] Mathur, S., Trappe, W., Mandayam, N., Ye, C., & Reznik, A. (2008) Radio-telepathy: Extracting a secret key from an unauthenticated wireless channel. In *Proceedings of MobiCom* (pp. 128–139).Google Scholar

[7] Xu, X. H. (2013). Study on security problems and key technologies of the internet of things. In *Proceedings of the IEEE international conference on computing and information sciences (ICCIS)* (pp. 407–410).Google Scholar

[11] Ouafi, K., & Vaudenay, S. (2009). Pathchecker: An RFID Application for tracing products in supply-chains. In *Proceedings of the workshop on RFID Security–RFIDSec* (vol. 9, pp. 1–14).Google Scholar

[12]Blass, E. O., Elkhiyaoui, K., & Molva, R. (2011). Tracker: security and privacy for RFID based supply chains. In *Proceeding of the 18th network and distributed system security symposium*. Google Scholar

[12] Elkhiyaoui, K., Blass, E. O., & Molva, R. (2012). CHECKER: On-site checking in RFID-based supply chains. In *Proceedings of the fifth ACM conference on security and privacy in wireless and mobile networks*. Google

[13] Ye, T., Peng, Q. M., & Ru, Z. H. (2012). *IoT's perception layer, network layer and application layer security analysis.* http://www.iot-online.com/jishuwenku/2012/1029/22888.html. Accessed 12 Nov 2018.

[14] Liu, B., Chen, H., Wang, H. T., & Fu, Y. (2012). Security analysis and security model research on IoT. *Computer & Digital Engineering*, *40*(11), 21–24.Google Scholar

[15] Suo, H., Liu, Z., Wan, J., & Zhou, K. (2013). Security and privacy in mobile cloud computing. In *Proceedings of the 9th IEEE international wireless communications and mobile computing conference* (pp. 655–659), Cagliari, Italy. Google

[16] Wan, J., Chen, M., Xia, F., Li, D., & Zhou, K. (2013). From machine-to-machine communications towards cyber-physical systems. *Computer Science and Information Systems*, *10*(3), 1105–1128.CrossRefGoogle Scholar

[17] ITU-T. Recommendation Y. 2002. (2010). *Overview of ubiquitous networking and of its support in NGN.* Geneva: ITU. Google Scholar

[18] Want, R. (2006). An introduction to RFID technology. *IEEE Pervasive Computing*, *5*(1), 25–33.CrossRefGoogle Scholar

[19] Yang, G., Xu, J., Chen, W., Qi, Z. H., & Wang, H. Y. (2010). Security characteristic and technology in the internet of things. *Journal of Nanjing University of Posts and Telecommunications (Natural science)*, *4*, 20–29.Google Scholar

[20] Wan, J., Zou, C., Ullah, S., Lai, C., Zhou, M., & Wang, X. (2013). Cloud-enabled wireless body area networks for pervasive healthcare. *IEEE Network*, *27*(5), 56–61.CrossRefGoogle Scholar

[21] Wan, J., Zhang, D., Sun, Y., Lin, K., Zou, C., & Cai, H. (2014). VCMIA: A novel architecture for integrating vehicular cyber-physical systems and mobile cloud computing. *ACM/Springer Mobile Networks and Applications*, *19*(2), 153–160.CrossRefGoogle Scholar

[22] Lv, B. Y., Pan, J. X., Ma, Q., & Xiao, Z. H. (2008). Research progress and application of RFID anti-collision algorithm. In *Proceedings of the international conference on telecommunication engineering* (vo1. 48, no. 7, pp. 124–128).Google Scholar

[23] Blaskiewicz, P., Klonowski, M., Majcher, K., & Syga, P. (2013). Blocker-type method for protecting customers' privacy in RFID systems. In *Proceedings of the IEEE international conference on cyber-enabled distributed computing and knowledge discovery (CyberC)* (pp. 89–96).Google Scholar

[24] Chen, M., Gonzalez, S., Zhang, Q., & Leung, V. (2010). Code-centric RFID system based on software agent intelligence. *IEEE Intelligent Systems*, *25*(2), 12–19.

[25] Toumi, K., Ayari, M., Saidane, L., A., Bouet, M., & Pujolle, G. (2010). HAT: HIP address translation protocol for hybrid RFID/IP internet of things communication. *TUNISIA: International conference on wireless and ubiquitous systems* (pp. 1–7).Google Scholar

[26] Lakafosis, V., Traille, A., & Lee, H. (2011). RFID-CoA: The RFID tags as certificates of authenticity. In *Proceedings of the IEEE international conference on RFID* (pp. 207–214).Google Scholar

[27] Chen, M., Lai, C., & Wang, H. (2011). Mobile multimedia sensor networks: Architecture and routing. *EURASIP Journal on Wireless Communications and Networking*, *2011*(1), 1–9

# Assessment of the Impact of Social Networking Sites Usage on Students' Academic Performance: *A Systematic Review*

Beauty Lweendo
School of Engineering,
Department of Information Technology,
Mulungushi University, Kabwe, Zambia
Email: lweendobeauty@gmail.com

Douglas Kunda
School of Engineering,
Department of Information Technology,
Mulungushi University, Kabwe, Zambia
Email: dkunda@mu.edu.zm

***Abstract***

**The implementation and adoption of contemporary Internet services in Zambia has seen the greatest number of netizens originating from Zambia spend their time online on social networking sites (SNSs). Social networking sites are online platforms that provide individuals with an opportunity to manage their personal relationship and remain updated with the world. This study will focus on assessing the impact of social networking sites usage on the students' academic performance in the selected Colleges of Education in Zambia. This is with the view of establishing whether the use of these SNSs has positive or negative impact on students' academic performance in the modern education arena. The study will be based on a research model with six hypotheses. Quantitative approach will be used. The sample size will be 500 students from Colleges of Education and the researcher will use stratified random sampling to select participants because of the uneven enrolment at Colleges of Education. Data for the study will be collected through self-administered questionnaire and data will be analysed using Pearson's correlation and SPSS tool will be used. The six hypotheses will be tested using Pearson's correlation and Cronbach's alpha will be used to test validity and reliability of the instruments.**

**The general objective of this study is to assess the use of SNSs by college students in selected Colleges of Education in Zambia with the view of determining if the use of SNSs has impacted upon their academic performance and if so, how. It is hoped that the information to be generated from this study will first and foremost help stakeholders like colleges to re-assess the importance and significance of the use of SNSs by students in their educational studies and help not only students but also other individuals throughout the country on how to use SNSs responsibly and for the development of their knowledge and skills.**

***Keywords: Internet, Social Networking Sites, Students, Academic Performance***

## I. INTRODUCTION

It has been acknowledged globally, that electronic media is the most effective tool to bring about the socio-economic development and it has the potential for individuals to communicate with each other no matter where they may be located [20]. The Internet is more than just a means of seeking information as people have discovered that the Internet could be used to connect with other people, whether for business or commercial purpose, make new friends, reawaken old friends and long lost relatives [24]. Therefore, the emergence of SNSs simplify the whole process as they are easier to use and navigate, as it does not require advanced knowledge and experience of the internet and are made up of a wide array of different formats and topics hence anyone can connect.

Consequently, young generation of Zambians especially the students from universities and colleges are the major users of the internet facilities mainly because Eduroam provides connectivity to the Internet where the service is deployed for every user registered in the member institution within the country with no charge to the user [6]. Reference [29] also observes that since social media has become a popular and fundamental aspect of the social lives today, most Zambian students are using it. The Zambian government has built learning institutions with the aim to share and impart knowledge and skills to those who become the part of these institutes. These Learning institutions measure the learner's academic performance by their examination outcomes held in the educational institutes themselves. . It is vivid that the advanced and improved usage of social networking platforms such as Facebook has become a worldwide phenomenon for quite some time. It started as a hobby for many computer enthusiasts and transformed into a social norm and existence style for students around the world. We seem unaware of the personal impact brought by SNSs on one's economic stand view with time constant. Reference [17] observes that through the Internet, communication oriented Internet sites such as social networking sites, blogs and sites for sharing photos and videos, are offered.

In Zambia, like many other countries, access to Internet is also available in many other higher institution of education apart from the University of Zambia and the use of Internet has been increasing rapidly amongst university students [1]. The growth of Internet use in Zambia is currently estimated at 63% from 13.47% in 2012, and this has been attributed to widespread use of personal mobile devices and adequate Internet service provision which has covered almost all areas of the country [26]. Recent survey conducted by the Zambia Information and Communications Technology Authority (ZICTA) found that 63% of Internet users in Zambia spend their time online on SNS while about 71% of those that own

smart phones use their devices to access WhatsApp, Viber, Facebook, Skype and Twitter for communication using instant messaging or voice calling [29]. Due to this vital position it is commanding in education sector, the Internet has become one of the most sought services in most institutions of higher learning, accounting to a large chunk of most college budgets. SNSs, such as Facebook, MySpace, Friendster and many others, have become instrumental for providing free, user friendly access for communicating with others over the Internet. SNSs can be used by group members to distribute news or opinions and share materials publicly through the News Wall [11].

Reference [22] also observes that as much as SNSs are leisure focused, there is growing emphasis on exploiting the sites for education. Reference [3] writes that SNSs allow individuals to construct a profile either public or semi-public within a bounded system in order to articulate a list of other users that share a connection and view as well as traverse their list of connections and also those made by others within the system. Reference [16] further indicates that SNSs connect students through shared activities as members can create personal profiles, join interest groups and upload videos, pictures and music. One notable observation is that girls spend more time on SNSs as compared to boys and that the youths, especially students, are the most users of SNSs, therefore, making the study of SNSs among high school and college students an area of interest amongst many researchers [3].

To date, the research on the relationship between time spent on Facebook and academic performance has provided mixed results [20]. This study seeks to assess whether there is a positive or negative relationship between the use of online SNSs and college students' academic performance in selected Colleges of Education in Zambia. In the face of the ever rapid increasing of online social networking amongst students, it is necessary to explore the risks that may result out of excessive use of these SNSs by college students.

## II. LITERATURE REVIEW

### *A. Social networking sites (SNSs)*

Social Networking Sites are the web-based services that gives individuals the opportunity to create either a public or semi-public profile within a bounded system, add a list of others with which they share a connection and view and transverse their list of connections and those made by others within the system[14]. Reference [21] defined SNSs as web based services that enable individuals to construct a semi-profile within a bounded system, articulate a list of other users with whom they share connections and views". SNSs is a web-based services that allow individuals to construct a public or semi-public profile with a view to articulate a list of other users with a shared connection in order to view and traverse the list of connections within the system. SNSs as online platforms that provide individuals with an opportunity to manage their personal relationship and remain updated with the world. In other words, SNSs can be defined as an online community of Internet users who want to communicate with other users about areas of mutual interest. Reference [14] further observes that, the term "social network site" is usually used to describe this phenomenon and "social networking sites" also appears in public discourse, and both are often used interchangeably. It is further important to realize that "networking" emphasizes relationship initiation, usually between strangers as one of the main activities in the course of SNSs usage. Simply put, Social networking sites are the online services whose core focus is to allow involved parties share data and information in real time [4]. Social networking sites are more common to the young generation worldwide. Reference [29] adds on that social media has become ubiquitous and almost inescapable, revolutionizing the way students communicate, socialize hence becoming an integral part of their social and cultural fabric. The usage of social networking sites has become the trend among the youth world over [5]. This is because social networks tend to help people a lot in a positive manner. Some examples of SNSs are Twitter, Friendster, MySpace, Facebook and Orkut.

In recent years, SNSs use has become very popular in many areas especially in education where many schools now have access to Internet. In Zambia, the use of SNS has been increasing rapidly amongst university and college students [1]. In recent years, SNSs use has become very popular in many areas especially in education where many schools now have access to Internet. In Zambia, the use of SNS has been increasing rapidly amongst university and college students [1]. This is the case because young generation of Zambians especially the students from universities and colleges are the major users of the internet facilities mainly because many such learning institutions have modern technology supportive infrastructure in place. The following discussion outlines SNS issues and challenges:

### *B. Social networking sites (SNS) as a tool for communication*

The evolution of internet technology by man has turned the entire world into a "global village" and millions of social networking sites have transformed the thought of global village into a reality where billions of people communicate through social networking sites [23]. SNSs are online communities of the internet that enables students to communicate with other users so as to promote their interactions with others.

As such, it has been acknowledged globally, that electronic media such as SNSs being an effective tool for communication has the potential for individuals to communicate with each other regardless of distance [20].SNSs are tools students use to get in touch with each other sharing the same interests while facilitating information sharing as well as opinion exchange [3]. Social networking tools are changing the environment and possibilities for education as they harness the collective intelligence of students, promoting collaboration and the sharing of knowledge [16]. Reference [12] notes that

social networking and media tools provide students the opportunity to communicate, access information, research and chat with friends. Reference [24] writes that social software refers to a set of tools designed to support sharing, collaboration and socialization resulting into development of multiple forms of social capital. SNS are web based services that give individuals an opportunity to create their own personal profile and preferred list of users whom they would connect with in public forum that provides them with features like chatting, video calling, mobile connectivity as well as video/photo sharing [15]. As such, [25] observes that social networking tools are being adopted in education.

### *C. Social networking sites and the youth*

The youth worldwide have adopted SNSs enthusiastically, mostly because of so much freedom in the physical world, their degree of digital literacy, the technical design of SNS which is easy to manage settings and of course its transparency regarding use of personal information [11]. Reference [7] observes that teenagers learn to assimilate, customize the look of their online profiles to present themselves as well as identify with their peers on SNSs. They further note that newcomers mostly struggle with demands of socializing virtually. SNSs like Facebook, Twitter and MySpace are gaining popularity due to their attractive features, as such, youths of today's generation are fascinated towards them [15]. Social software for SNSs helps youth build communities of interest on a wide range of subject matter through collaboration. One notable observation is that the youths, especially students, are the most users of SNSs, therefore, making the study of SNS among high school and college students an area of interest amongst many researchers [3].

### *D. Social networking sites and the student's academic performance*

Academic performance is defined as"...how students deal with their studies as well as how they cope with or accomplish different tasks given to them by their teachers"[14]. It is the friendship networks that usually enable students to access information and knowledge either directly or indirectly, and as a result has an effect on student academic performance. Learning institutions should realize that involvement of students in different forms of activities such as making friends on social networks is a way of enabling them to access up to date information that is relevant and that can be channelled towards improving student's academic performance. As it may be so, this entirely depends on the ability and willingness of a particular student to be able to harness that opportunity in order to cope with academic related stress. Reference [14] further observes that, a student who records a high creativity on social networks has the tendency to make lots of friends online and also may translate same to his normal daily academic life.

Reference [22] observes that SNSs enables students to discuss and share their educational knowledge, ask for help from their fellow students, participate in online communication as well as interact for active learning. He further observes that among SNSs, Facebook is more popular because it offers opportunities to give and receive constructive peer-to-peer feedback in a community that matches the social context of teaching and learning in a collaborative venture like a school, college and university. Reference [25] also notes that Facebook is an easy-to-use and familiar tool for students, as such, it is used to share and exchange ideas among peers, classmates and public. Reference [7] write that some students use their social network to work out emotional situations, leading to increased academic productivity. Students use SNSs with regards to communication abilities, technical skills development and effective use of new technology. By the use of SNSs, the students can retrieve the required information within the short period.

Reference [12] observes that, social media allow students to get together outside their classrooms as they collaborate and exchange ideas about projects and assignments. They can have access to e-books and e-journal with which not only provide them with the needed information but also improve their learning. Reference [22] observes that SNSs have been utilized in education field because of its accessibility, efficiency and ease of use, individual affordances and engagement. SNSs allows for the exchange of information and for interaction among individuals making transport and distance no longer a problem [10].

Reference [19] writes that SNSs enable students to upload their content to the web in form of text, voice, images and videos. Reference [25] in their research, observe that integrated use of SNSs increases the efficiency and productivity in academic settings and Information and Communication Technology (ICT) usage providing a rich platform for teaching and learning, enhancing students' ability to teach and learn, a positive outcome in educational settings. In this way, students use the SNSs among other things for academic purposes such as sharing information needed for school projects and other academic assignments. This in turn can improve the academic performance of students and also help them to develop the necessary skills and technical literacy that will be vital in the future. SNSs is one of the greatest advancement in the world of information technology on which students depend on either socially or academically as shown in fig. 1.


***Fig. 1: Cloudworks social networking site used in education***
**Source: [16]**

Cloudworks is an example of SNSs which aim at providing a dynamic environment to assist those in education to share and discuss teaching ideas and designs. It is therefore important to provide social as well as academic support mechanisms amongst students themselves and with faculty.

*E. Social networking sites used by boys and girls*

Reference [12] indicates that there is a significance difference between boys and girls use of SNSs. Boys spend less time on SNS as compared to girls. They further write that boys use music sharing sites whereas girls use Facebook and twitter to share videos. Boys usually use SNSs to make new friends and share 'self-promoting' pictures while girls use SNS to communicate with friends and usually post 'cute' pictures.

*F. Student's perception of social networking sites (SNSs) usage*

The perception of use of SNSs depends on an individual as it has both positive and negative influence upon academic performance of students. If more time is spent to communicate with faculty members, academic resource sharing of materials, collaborating with lecturers and educational video sharing, then it has positive influence, but if more time is spent on SNSs for non-academic usage such as updating profiles, gossiping etc, then it has negative influence [14].

*G. Benefits of students using social networking sites*

There are several researchers who have researched on social networking sites and student academic performance across the world. They have observed that SNSs can have positive impact on student academic performance depending on how one uses them, of which even students' themselves are aware of (Oye et al., 2012). (Table 2) shows different writers' views on benefits of SNSs on student's academic performance and simply deduces that there are similarities of researchers towards benefits of SNSs on student academic performance.

***Table 2: Benefits of SNS on Students Academic Performance***

| Reference | Benefit: For Obtaining Information | Benefit: For Communication |
|---|---|---|
| [29] | SNSs have benefits like facilitating interactions, sharing of information, and promoting connections among health professionals. | SNSs have been developed as useful platforms for communication |
| [1] | Students use SNSs to obtain new information from their friends | SNSs help students to keep in touch with old friends, family and to strengthen relationships. |
| [4] | | Students are always online for social communication and interaction. |

*H. Challenges students using social networking sites*

Education is very essential part of an individual's life and as such, for every teenager education is very important than anything else. Currently, majority of the youth have various accounts of the networking sites and that is why their academic performance is poor. Students might blame about the poor excellence of the teachers while ignoring their social networking sites' fever. The time that students should give to learning and academic research is crushed by the passion for making new friends and other sluggish issues. Henceforward, most of the students' academic performance suffer due to the use of social networking sites as it distracts them from their academic goals [5].

Impairment of educational performance and internet dependency are correlated by utilizing synchronous communication programs including internet sites and forums. Electronic media usage is negatively associated with grades. Two third of the students, use social networking sites in classroom, while studying and doing homework. Social networking sites may badly affect academic life and learning experiences of the students as social network sites grab the attention of the students and then diverts it towards the non-educational activities.

Today teenager shows very much interest for using social networks but unfortunately Social Networks affect education badly [6]. Social Networking Sites grab the total attention and concentration of the students and diverts them towards non-educational, unethical and inappropriate actions such as useless chatting, time killing by random searching and not doing their jobs. As SNSs has introduced many attractive tasks like gamming, advertisements etc. so that people can never get enough of these things. The SNSs addict becomes a useless node for parents, friends and other associated people. They cannot succeed because they have no sense of upcoming future and competitions in their careers [23].

Due to a great percent of times spent on social networking activities, the performance and the success of students suffer setbacks, which could lead to student's poor performance, and they might fail to create a balance between the social media and academic activities. Furthermore, it is about their addiction, the loss of time leading to consequences in their academic development, and the social networking sites influence them largely negatively, because their attention is focused on chatting and music while their academic activities are neglected [7].

To date, the research on the relationship between time spent on SNS like Facebook and academic performance has provided mixed results [20]. Communication oriented Internet sites such as social networking site is one of the major factors affecting the academic performance and social life of students. The findings of these studies show that the number of hours spent on SNSs affect the grades of students depending on if the SNS are used for study or social purposes.

(Table 1) show different writers' views on challenges of SNSs on student's academic performance and simply deduces that there are similarities of researchers towards challenges of

SNSs on student academic performance as long as SNS is negatively used.

***Table 1: Challenges of SNS on Students***

| Reference | Challenge: Addiction/Bullying | Challenge: Risk to Sexual Activities |
|---|---|---|
| [23] | Lots of students are now addicted to online rave of the moment with Facebook chat, WhatsApp, while lectures are on. | |
| [12] | Social media cause depression, cyber bullying leading to poor academic performance | Sexual harassment leading to poor academic performance |
| [15] | The rapid popularity of SNS causes students to distract themselves from their studies as they continuously use it | |
| [16] | Students lack the motivation to communicate online once they face bullying | Risk to students such as sexual solicitation |
| [25] | Students worry about their privacy and invasiveness of social networking | |
| [11] | SNS usage is associated with bullying, harassment and exposure to harmful content | SNS usage is associated with sexual grooming and violent behaviour |
| [17] | SNS addiction by students due to excessive use | |

## III. LITERATURE REVIEW METHODOLOGY

The literature review methodology follows a purposeful data selection methodology and data is primarily selected from Journals and conference proceedings. A standard literature review methodology, which was proposed by [18] is applied and will be modified by the researchers to suit the topic at hand which is to assess students on the impact of academic performance at Colleges of Education in Zambia. Three stages namely; planning the review, selection and execution will be taken into account as shown in fig.2.

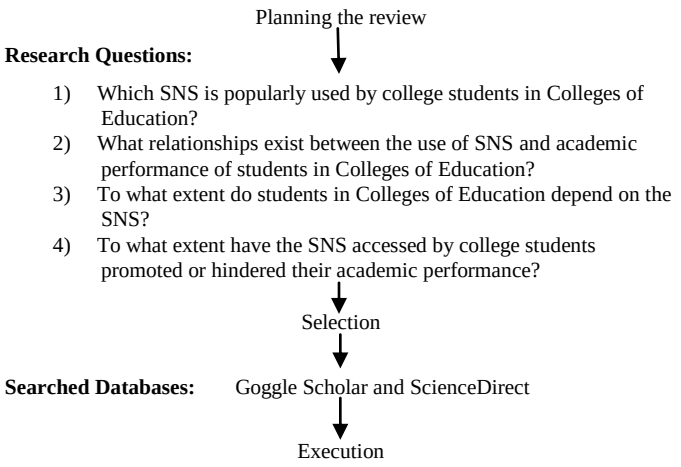



**Fig. 2: Systematic Literature Review Steps**

### *A. Planning the review*

The initial step in planning the review is to define the research questions, presented here as:

1) Which SNS is popularly used by college students in Colleges of Education?
2) What relationships exist between the use of SNSs and academic performance of students in Colleges of Education?
3) To what extent do students in Colleges of Education depend on the SNS?
4) To what extent have the SNS accessed by college students promoted or hindered their academic performance?

A search was conducted in June and September 2018 only for literature published in English language prior to developing a review's protocol. The search string was *"social networks", "social networking sites", "impact"," academic performance" and "social networking tools".*

### *B. Conducting the review and reporting*

#### *1) Selection*

The search selection was based on the pre-determined key words which are; *"social networks", "social networking sites", "impact"," academic performance" and "social networking tools".* The initial results for each search were high due to the search criterion which was wide because the search was based on selecting articles containing any of the key words. Thereafter, the criteria were refined using advanced search to narrow the search selection criterion. We looked for articles containing a pair of key words, either "social networks and academic performance", "social networking sites and impact on academic performance", and "social networking tools and academic performance". The time periods were from 2008 – 2018 in order to capture journals and articles of not order than ten years ago.

- **Searching Literature**

This systematic literature review was limited to the following databases; Google Scholar and ScienceDirect.

- **Practical Screening**

In order to test conformance to set criteria, articles and journals were systematically analyzed by scanning titles, reading abstracts, articles published in English language and of course those published between 2008 and 2018. Also, the articles and journals selected were those containing social networking sites/social networks and impact on academic performance but excluding those without references.

### *C. Execution*

#### *1) Data reporting*

Data reporting section presents a summary of the final publications that were selected for detailed analysis. They are reported in Table 2 in terms of author, the title of the research paper, research approach, and location of the research. Research findings and in some cases the gaps in the research findings are also highlighted.

*Table 2: Summary of previous studies on social networking sites and academic*

| Author | Title of the research paper | Research Approach | Location | Research Findings/gaps |
|---|---|---|---|---|
| [1] | Students' Social Media Use and its Perceived Impact on their Social Life,": A Case Study of the University of Zambia | Survey | Zambia | Findings of the research from the sampled University of Zambia students, indicates that most students use WhatsApp more than any other to obtaining new information, keeping in touch with old friends, family and to strengthen relationships than they would for academic purposes. The paper further indicates that most of these students find social media "irresistible" to the extent that they check their accounts every now and then before doing something else. This causes their academic work to suffer. |
| [13] | Usage of online social networking sites among school students of Siliguri, West Bengal, India | Cross-sectional descriptive | India | The study findings indicated that age, gender, accessibility, etc are socio-demographic factors that are associated with the SNS addiction among adolescents. The amount of time spent using social media has a positive association with the SNS addiction, meaning the more time one spends on social media, the more they will show addiction, which takes over time for studying. |
| [27] | Use of Social Network Sites for Communication Among Health Professionals: Systematic Review | Systematic Reviews and Meta-analyses guidelines | China | The paper indicated that SNSs have been developed as useful platforms for communication among users with significant benefits in education and professional networks. The paper further indicates that SNS users in the reviewed studies considered SNSs as user-friendly, easy-to-use, free, and fast tools for communication. The researchers also observed that SNSs has benefits like facilitating interactions, sharing of information, and promoting connections among health professionals. |
| [4] | Social Networking Sites as Communication, Interaction, and Learning Environments: Perceptions and Preferences of Distance Education Students | Quantitative cross-sectional | Turkey | This paper explored the use of SNSs for communication, interaction and learning by distance education students at Anadolu University, of which it was discovered that they use mobile devices to access these social tools. The paper indicates that distance education students usually connect to SNSs on a daily basis and most of them are always online for social communication and interaction. The paper further found that gender is significant in terms of communication patterns and degree of interaction. |
| [23] | Impact of social networking on academic achievement of undergraduate social studies students in rivers state. | Descriptive Survey | Nigeria | Thee paper indicates that students are addicted to social networking hence they are negatively affected when it comes to academic performance. The paper further writes that this is due to the fact that students are always are pinging and Facebooking while lectures are on. It was also discovered that, mostly teenagers are the ones that mostly use social networks. |
| [12] | School performance, social networking effects and learning of school children,": Evidence of reciprocal relationships in Abu Dhabi | Quantitative cross-sectional | Abu Dhabi | This study shows that social media networks compete with academic work for students' attention as such each student needs to be responsible in order to make the right decision in relation to the use of social media networks. The writers stated that "learning outcomes of students are influenced by the student's decision on the choice of situation (social media networks and participation) and peers (friendship networks), hence could make the right decision in the usage of these media to acquired better academic performance. |
| [2] | The Impact of Social Networking Sites' Usage on the Academic Performance of University Students of Lahore, Pakistan | Descriptive survey | Pakistan | The findings of this study reveal that Facebook is the most popular social networking site used by many students followed by WhatsApp Messenger and then YouTube. The paper further indicated that there is an impact of social networking sites on the student's academic performance because spend a lot of time on these sites as a result they get distracted. |
| [8] | The impact of social media on students' academic performance- a case of malaysia tertiary institution | Descriptive survey | Malaysia | The paper found that research, time appropriateness and health addiction has a stronger significant influence on students' academic performance. The reason is that, time management plays an important role in determining the success or failure of an individual. As a result, students who lack time management can easily fall prey to the negative impact which social media platforms present to its uses. |
| [9] | The Impact of Social Networks on Students' Performance | Quantitative | | The findings of the paper show that most students are aware of the social media advantages and disadvantages. On the advantage part, students can access various social networking sites such eentertainment with friends, posting personal issues, photos as they believe that social networks enable them post different information as they get informed through social networks what happens in university activities. On the disadvantage part, students fail to create a balance between the social media and academic work. It is for this reason that students should create a balance between chating and other socialization aactivities. |

## IV. ANALYSIS OF FINDINGS

The literature review indicates that information represents an evolution because it changes the way information is accessed and that it is important to know where to collect information than knowing the information itself [16]. This is as a result of the e-world that has taken its toll over the world and as such, SNSs have been developed as useful platforms for communication among users because they are user-friendly, easy-to-use, free, and fast tools for communication [27]. Among these social networking sites, Facebook is the most popular followed by WhatsApp, Messenger and YouTube [2]. These social networking sites have both positives and negatives. Literature review shows that students are even aware of these positives and negatives of social networking sites on their academic performance [14] but it entirely depends on an individual student to get the best or vise versa out of the social sites.

SNSs are very beneficial in the academic sector as long as they are used positively. This paper will indicate some of the benefits of SNSs observed from the literature review. Students can use SNSs positively for academic purposes just like Egyptian students who spend between one to more than six hours daily on social media platforms, but they still manage to get good grades [28]. SNSs can be used to communicate with both lecturers and fellow students basically for academic purposes such as collaboration in classroom by using technology in order to have active learning.

On the negative part, social networks has negatively affected academic performance of students, reason being that, they spend more time on doing un-academic work than they spend time on academic related work [20]. The literature further indicates that teenagers are the ones that use social networks frequently which is very sad. Also, mobile phones are owned by almost each and every student, instead of using them for academic work, they use these phones to access social sites. Students keep checking their accounts even when lessons are going on and at times when they are supposed to be studying. Social sites are irresistible if not handled properly and may cause academic work to suffer [1]. Once these social networks become irresistible, then students become addicted. Some students become addicted simply because they are homesick or unfamiliar to the new environment while others cannot just resist to using these social sites not knowing that they are losing social values.

## V. RECOMMENDATIONS

Literature showed that social media networks compete with academic work for students' attention. Also, learning outcomes of students are influenced by the student's decision on the choice of using social media to socialise with peers, therefore it is advisable that each student becomes responsible in order to make the right decision in relation to the use of social media networks because students who lack time management can fall prey to negative impacts of the social network sites. Another recommendation is that, College management should not be blocking access to social sites but integrate social media into classrooms, assignments and projects and drop ideas for discussions. This will help to indirectly discourage them from nonacademic aspects of SNSs when they are online and in due course, management will be educating students on how positively they can use these sites in order for them to appreciate use of SNSs in education.

## VI. CONCLUSION

It has been observed that most researchers in the literature review are in agreement that the world of education has completely changed since the development and growth in form of social networking sites and educators are seeking for their potential use in education, have the conscious that social networking sites have the capability to endorse both collaboration and active learning. Nevertheless, literature review indicated that social networking sites usage has both positives and negatives towards academic performance of students as it depends on how they are used. When these social sites are used for academic purposes such as communication with lecturers and friends, they have a positive impact but when they are used for social activities like gossiping, entertainment etc, they have a negative impact on academic performance.

It was also observed that it is high time management started to integrate SNSs in education for students to discuss as a way of encouraging them to use SNSs for education purposes unlike the other way round.

## REFERENCES

[1] Akakandelwa and G. Walubita, *"*Students' Social Media Use and its Perceived Impact on their Social Life,": A Case Study of the University of Zambia", The International Journal of Multi-Disciplinary Research, ISSN: 3471-7102. pp. 1-14. 2017.

[2] Ali Waqas, Muhammad Afzal, Fakhar Zaman, Muhammad Sabir, "The Impact of Social Networking Sites' Usage on the Academic Performance of University Students of Lahore, Pakistan," *International Journal of Social Sciences and Management*, vol. 3, no. 4, pp. 267-276, 2016.

[3] Anusha, J. J. Parappilly and A. K. Sangaiah, "A New MCDM Approach Integrating QFD, Dematel with Topsis for exploring the effect of Social Network Usage on Academic Performance," vol. 6/4, 32-42, 2015.

[4] Aras Bozkurt, Abdulkadir Karadeniz, Serpil Koçdar, "Social Networking Sites as Communication, Interaction, and Learning Environments: Perceptions and Preferences of Distance Education Students," *JL4D*, vol. 4, no. 3, 2017

[5] Bithika M and Sara Selvaraj, "Impact of social media on student's academic performance," vol. 2, no. 4, 2013.

[6] D. Kunda, C. Chembe and G. Mukupa, "Factors that Influence Zambian Higher Education Lecturer's Attitude towards Integrating ICTS in Teaching and Research," 2018. doi.org/10.3926/jotse.338.

[7] Greenhow, B. Robelia and J. E. Hughes, "Learning, Teaching and Scholarship in a Digital Age," Educational Researcher 2009 38: 246. Doi: 10.3102/0013189X09336671.

[8] Nizam, "The impact of social media on students' academic performance- a case of malaysia tertiary institution," International Journal of Education, Learning and Training, vol. 1, no. 1, 2016.

[9] Zekiri, "The Impact of Social Networks on Students' Performance," *Academic Journal of Business*, vol. 2, no. 3, pp. 182-193, 2016.

[10] Sangaiah, X. X. Gao and A. Abraham, "Exploring the Antecedents of Social Network Usage on Academic Performance,": A Combined GP-Topsis Approach, vol. 78:1, pp. 55-67, 2015. elSSN 2150-3722.

[11] Livingstone, Sonia, Brake and R. David, "On the rapid rise of social networking sites,": new findings and policy implications. Children and Society. vol. 24. No. 1. pp. 75-83, 2010. DOI:101111/j.1099-0860.2009.00243.x.

[12] MasoodBadri, A. AlNuaimi, YangGuang and A. AlRashedi, "School performance, social networking effects and learning of school children,": Evidence of reciprocal relationships in Abu Dhabi. Https://doi.org/10.1016/j.tele.2017.06.006.

[13] Medha Raj, Abhijit Mukherjee and Sharmistha Bhattacherjee, "Usage of online social networking sites among school students of Siliguri, West Bengal, India," *Indian Journal of Psychological Medicine*, vol. 40, no. 5, pp. 452-457, 2018.

[14] D. Oye, M. A. Helou and N. Z. Rahim, "Model of Perceived Influence of Academic Performance Using Social Networking Sites," International Journal of Computers & Technology. vol. 2, 2012. ISSN: 2277-3061.

[15] N. Khurana, "The Impact of Social Networking Sites on the Youth," J Mass Communicat Journalism 5: 285, 2015. doi:10.4172/2165-7912.1000285.

[16] N. Mccarroll and K. Curran, "Social Networking in Education," Research Gate, DOI:10.4018/jide.2013010101.

[17] N. Xiaoli, Y. Hong, C. Silo and L. Zhengwen, "Rapid Communication": Factors Influencing Internet Addiction in a Sample of Freshmen University Students in China. vol. 12. no. 3. DOI: 10.1089/cpb.2008.0321.

[18] Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *Working Papers on Information Systems*, *10*(26), 1–51. https://doi.org/10.2139/ssrn.1954824

[19] P. Kanthawongs and P. Kanthawongs, "Perception of Primary School Students, Parents and Teachers toward the Use of Computers, the Internet and Social Networking Site," Social and Behavioural Sciences Symposium, 4th International Science, Social Science, Engineering and Energy Conference 2012 (I-SEEC 2012). doi: 10.1016/j.sbspro.2013.08.507.

[20] Glass, J. Prichard, A. Lafortune and N. Schwab, "The influence of Personality and Facebook use on Student Academic Performance," Issues in Information Systems, vol. 14, pp 119-126, 2013.

[21] Serpell, "Promotion of Literacy in Sub-Saharan Africa,": Goals and Prospects of CAPOLSA at the University of Zambia. Human Technology, vol. 10. no. 1, pp. 22-38, 2014.

[22] S. Hayato, Abrahim, M. A. Bilal and X. Masahiro, "Exploring Academic Use of Online Social Networking Sites (SNS) for Language Learning": Japanese Students Perceptions and Attitudes Towards Facebook. J Inform Tech Softw Eng. vol. 8. No. 1, 2018: 223. ISSN: 2165-7866.

[23] S. B. Tete and C.C. Ezinne Abe, "Impact of social networking on academic achievement of undergraduate social studies students in rivers state," Nigeria. vol. 2; Issue 5; September 2017; pp. 47-50. ISSN: 2455-5746.

[24] S. McLeod, "Likert Scale," 2008. [Online]. Available: https://www.simplypsychology.org/likert-scale.html. [Accessed 24 September 2018].

[25] Saul McLeod, "Questionnaire," Simply Psychology, 2018. [Online]. Available: https://www.simplypsychology.org/questionnaires.html. [Accessed 25 September 2018].

[26] Anderson, Poellhuber and R. McKerlich, "Self-paced Learners Meet Social Software," An Exploration of Learners Attitudes, Expectations and Experience. Online Journal of Distance Learning, vol. XIII, Number III, University of West Georgia, Distance Education Center, 2010.

[27] Lei, T. Krilavicius, N. Zhang, K. Wan and K. L. Man, "Using Web 2.0 Tools to Enhance Learning in Higher Education,": A Case Study in Technological Education, ISBN:978-988-19251-9-0, 2012.

[28] Waqas Tariq, Madiha Mehboob, M. Asfandyar Khan and FaseeUllah3, "The Impact of Social Media and Social Networks on Education and Students of Pakistan," IJCSI International Journal of Computer Science Issues, vol. 9, no. 4, 2012.

[29] Windy SY Chan, and Angela YM Leung, "Use of Social Network Sites for Communication Among Health Professionals: Systematic Review," *JMIR*, vol. 20, no. 3, 2018.

[30] Y. Hashem, "The Impact of Social Media on the Academic Development of School Students," vol. 1 pp. no. 46, 2015. DOI: 10.5430/ijba.

[31] ZICTA, CSO, Republic of Zambia Ministry of Transport and Communications, "ICT Survey Report- Households and Individuals," Lusaka, 2015.

# A Review of Identity Attribute Metrics Modeling based on Distance Metrics

Felix Kabwe
The University of Zambia
Department of Computer Science
Lusaka, Zambia
fmlkabwe@yahoo.co.uk

Jackson Phiri
The University of Zambia
Department of Computer Science
Lusaka, Zambia
jackson.phiri@.cs.unza.zm

***Abstract*** **- The growth in the use of services on the World Wide Web has brought about in the proliferation of online fraud. This is hinged on the fact that cyber fraudsters and criminals would hide their online identities to steal services and other valuables. Work has been done in the past on strengthening of identity management systems as a way to arrest this growing problem. This study considers past work on the subject matter and builds on developing the metrics models in order to provide quantitative analysis to quantify the credential identity attributes in online services. Metrics models will be explored that would quantify the credential identity attributes which will help in uniquely identifying the real owners of the digital identities before services and other assets could be issued to requesters of the same. A review of literature in this area of interest has been done and therefore, the identified area of interest adds value to the resolution of the said problem.**

**Keywords: Identity Attributes, Metrics Model, Digital identity, Authentication, Online Services**

## 1 INTRODUCTION

Many computing devices are also deployed in the environments where the users evolve—for example, intelligent home appliances or RFID-enabled fabrics. [1] In this ambient intelligent world the Internet is most likely going to generate more complicated privacy problems. [2]

Digital identification has become a challenge in the cyber space. A stolen digital identity is a serious risk to assets, fraud, safety, and privacy; this reduces confidence in electronic business, cyber social networking, and communication in general.

Privacy is a real human need which needs to be protected. People have busy lives and should not spend their time administering their digital identities. There is need to build technology that would help users to enable and secure usability.

## 2 BACKGROUND INFORMATION

This part of the paper reviews the literature that focuses on similar research work, how the issue of identity management issues have been addressed as well as how the proposed solutions to the challenges have been implemented from previous works. The literature looks at Identity Management Mining, Metrics Composition, Information Fusion, Identity Quantitative Analysis, Development of a Metrics Model, Multifactor Authentication System, Fuser Block Technologies Performance, and Design and Implementation of Multimodal Digital Identity Management System. The other literature review covers Using Artificial Neural Networks to Implement Information Fusion in Digital Identity Management Systems as discussed in [3][4][5] and [6].

With the introduction of internet banking, mobile sending of money through Mobile phone service providers like Airtel, MTN, and Zamtel, Zambia has identified theft of digital identities. Customers would send money electronically, using mobile phones, but their digital identification has been stolen, and sometimes people have been misrepresented or asked to reveal their identities so that their money could not reach the intended beneficiaries but get into wrong hands. Developing techniques of identification of the real owners of the digital identification would help resolve this problem.

Zambia has also been rocked with identity management issues when it comes to distribution of farmers' inputs by the government. People pose to be the true clients of the government to receive government support to receive assistance. Government has spent billions of Kwacha thinking the recipients are the intended beneficiaries when in fact not. When it comes to a lot of Government Ministries, it is a big problem to identify the registered needy for them to receive government support. This has also led to huge government expenditure. Identifying officers who qualify to have fuel from the fuel stations is another area that gobble huge expenditures, officers use the card to fill in the tanks of their vehicles with very poor identification using electronic smart cards.

A project like this would help government reduce unnecessary expenditures and in the saving of the limited

resources which could be channeled to deserving areas of the economy.

## 3 DIGITAL IDENTITY DEFINITION

In a major identity management initiative digital identity is defined as "the distinguishing character or personality of an individual. An identity consists of traits. Digital identity management is a key issue that will ensure not only service and functionality expectations but also security and privacy". [11]


*Figure 1: Identity management elements [11]*

Identity management elements include the user who has to sign in for issuance of a service, or for authentication for access to an electronic or non-asset. The identification request has to be verified by a single or multi-mode verifier. The authorisation of the identification permits the requester access to the service or asset. Authentication is an assumption of trust. Therefore, it is at this point, in this process, where fraud can either prevented or allowed.

The previous works have looked at identification, techniques that could be combined to come up with a multimode identification, methods of mining the attributes that could be used in metric model quantification. In this case, an entity (which could be a person or thing) can be identified using specific identification like user name, objects like electronic cards, biometrics like voice etc. The frequency of the attributes or identifiers which could be used in the identification would help in establishing metric models quantifications which could further help in uniquely identifying the entity. The diagram below shows the relationship amoung identities, identifiers and Attributes/identifiers which would be used in the metric model quantification.

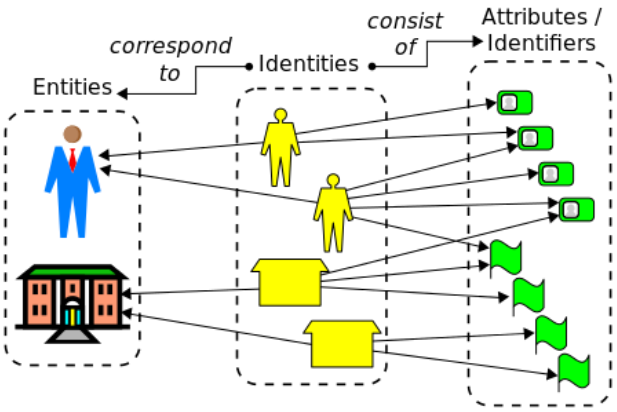



*Figure 2: Relationship among identities, identifiers and Attributes/identifiers [11]*

The relationship among identifiers, identifiers:

- A user who wants to access to a service
- Identity Provider, the issuer of user identity
- Service Provider, the relay party imposing an identity check
- Identity, a set of user attributes
- Personal Authentication Device (PAD), which holds various identifiers and credentials and could be used for mobility

Authentication is the process of verifying claims about holding specific identities. A failure at this stage will threaten the validity of the entire system. The technology is constantly finding stronger authentication using claims based on [12]:

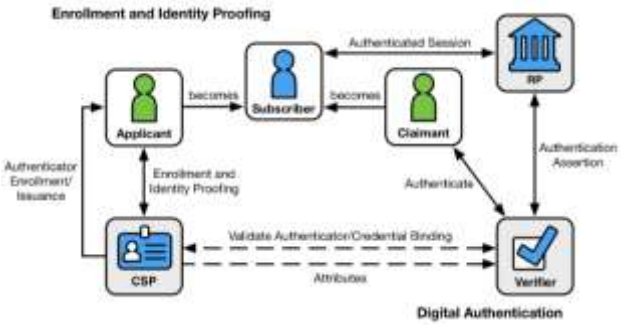
- Something you know (password, PIN)
- Something you have (one-time-password)
- Reachability. The management of reachability allows a user to handle their contacts to prevent misuse of their email address (spam) or unsolicited phone calls.
- Authenticity. Ensuring authenticity with authentication, integrity, and nonrepudiation mechanisms can prevent identity theft.
- Anonymity and pseudonymity. Providing anonymity prevents tracking or identifying the users of a service.
- Organization personal data management. A quick method to create, modify, or delete work accounts is needed, especially in big organizations. [13]

## 4 PROCESS OF DIGITAL IDENTIFICATION

In most cases, digital identification requires that the requester for identification must be on the database for identification. The entity must have made an application to be enrolled for recognition. The application could be done by filling in a form where details of personal identification are captured on a database. The applicant becomes subscribed and a token of identification is issued. The subscriber becomes a claimant until specific identification can be made by the applicant that is when authentication has to be made. A session of authentication takes place and verification of details that

were submitted by the subscriber are thoroughly checked. The identification of the subscriber is done as verification of the personal details of the subscriber; the verifier has to be a secure mechanism. The attributes of the subscriber are the subject of authentication. It is therefore, important that a lot of attention is paid to the attributes of identification as they are the pillar of the security of identification of a subscriber. In our work, we will consider the application forms and the attributes identified which we shall apply in our work to strengthen further the process of identification.

When identification is done, the subscriber can be authenticated.

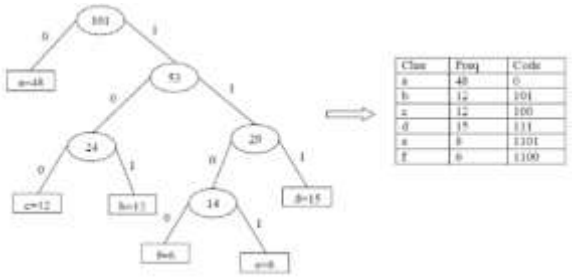



## 5 7. METRIC MODELS QUANTIFICATION

The thrust of this research will be on adding knowledge to metric models quantification. This quantification would add value to raise a bar in unique identification of the right entity for authentication.

At this stage, we will explore Huffman Code Procedure and Genetic distance for gene-frequency data to build our metric models. We will further explore if other methods could be added for metric models quantification. This paper is a precursor of the activities that we intend to carry and is thus preparatory work for the research activities. The paper should be considered as a research proposal which will assist in the exploration of the research interests. The paper is therefore not conclusive at this stage but prepares the platform of our work. It should be mentioned early that the identified methods have not been thoroughly explored and have therefore, room for change and inclusion of valuable information depending on the findings of the research. Huffman Code has not been conclusively explored at the moment.

In order to obtain the encoded bit stream, initially, we obtain the frequency of the residual coefficients that are arranged in ascending order. Then, two nodes that contain lowest frequency are selected to merge and the addition of two values is given into the new node. Subsequently, the same process is repeated for all nodes until we obtain a single node. Finally, the binary value is assigned to every node in accordance with the location (left or right) of the node. Then, each value obtains one code vector. [14]



*Huffman Code procedure*

We will explore Huffman Code procedure to develop an algorithm to identify the frequency of attributes that would show from a set of large number of application forms that would have the probability of the applicant being the right identity.

The Huffman Code Model's algorithm is shown below:

*Huffman Coding Algorithm*

i. Start with a list of symbols and their frequency in the alphabet.
ii. Select two symbols with the lowest frequency.
iii. Add their frequencies and reduce the symbols.
iv. Repeat the process starting from step-2 until only two values remain.
v. The algorithm creates a prefix code for each symbol from the alphabet simply by traversing the symbols back. It assigns 0 and 1 for each frequency value in each phase.

*Example*

Input: Six symbols and their respective Probability / Frequency Values

| Original Source | | Source reduction | | | |
|---|---|---|---|---|---|
| Symbol | Probability | 1 | 2 | 3 | 4 |
| a2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.6 |
| a6 | 0.3 | 0.3 | 0.3 | 0.3 | 0.4 |
| a1 | 0.1 | 0.1 | 0.2 | 0.3 | |
| a4 | 0.1 | 0.1 | 0.1 | | |
| A3 | 0.06 | 0.1 | | | |
| a5 | 0.04 | | | | |

*Frequency of attributes*

*Steps:*

i. Arrange the symbols in descending order of frequency values.
ii. Add the two lowest values and remove the symbols
iii. At the end we will be left with only two values

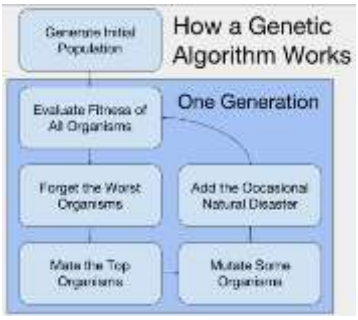
## 6 VARIABLE LENGTH CODES

*Steps:*

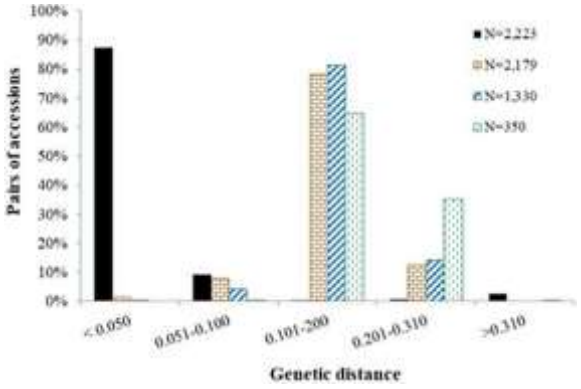
i. Moving backward assign the bits to the values.
ii. With each succession phase (reverse) add bit values using step 1 Result: Symbols with variable length codes.

## 7 GENETIC DISTANCES FOR GENE-FREQUENCY DATA

The other model we shall explore is the Genetic distances gene frequency method. This method displayed strong conformity of genes in very similar populations. This not the case when conformity was poor. [14]




*Genetic Algorithm*




## 8 PRIVACY REQUIREMENT

Privacy is a central issue due to the fact that the official authorities of almost all countries have strict legal policies related to identity. It is often treated in the case of identity management because the management deals with personal information and data. Therefore, it is important to give a definition. Alan F. Westin defines privacy as “the claim of individuals, groups and institutions to determine

## 9 TRUST FRAMEWORKS

Some Trust Frameworks are being developed by different establishments. The Higgins Trust Framework (HTF or Higgins, for short) is an open source project under development by the Eclipse Foundation that seeks to make sharing of identity information easier and more secure. IBM, Novell and Parity Communications are among the organizations contributing to the project. According to Dale Olds, an engineer at Novell, the purpose of the project is to give users more control over their online identity information. [12]

The Higgins framework enables users to securely store identity information and related data and to integrate that data across multiple systems and applications.

Stored data can be shared anonymously among Web applications, online vendors and service providers in a controlled manner. HTF API (application program interface) can be thought of as a repository for cookielike data that makes it convenient for users to conduct ecommerce and interact with Web sites, without the security problems inherent in conventional cookies.[13]

## 10 CONCLUSION

Quantification of the attributes of the digital identification metrics model will add value the unique identification on online activities. This will help in curbing fraud which has pervaded online businesses, electronic social interactions, and communications. This would confidence to electronic services and electronic commerce.

## REFERENCES

[1] J. I. Agbinya, N. Mastali, R. Islam, J. Phiri, "Design and implementation of multimodal digital identity management system using fingerprint matching and face recognition", Proceedings of the 7th International Conference on Broadband Communications and Biomedical Applications, Syden, Australia, pp. 272-278, 2011

[2] J. Phiri, T. Zhao, "Identity attributes quantitative analysis and the development of a metrics model using text mining techniques and information theory", Proceedings of IEEE International Conference on Information Theory and Information Security, Beijing, China, pp. 390-393, 2010

[3] J. Phiri, T. Zhao, J. Mbale, "Identity Attributes Mining, Metrics Composition and Information Fusion Implementation Using Fuzzy Inference System." Journal of Software, Volume 6, Issue Number 6, pp. 1025-1033, 2011.

[4] J. Phiri, T. Zhao, H. C. Zhu and J. Mbale, "Using Artificial Intelligence Techniques to Implement a Multifactor Authentication System," International Journal of Computational Intelligence Systems, vol. 4, no. 4, pp. 420-430, 2011.

[5] J. M. Seigneur and T. E. Maliki, "Identity Management," in Computer and Information Security Handbook, Burlington, Morgan Kaufmann, 2009, pp. 269-297.

[6] W. Stallings, Cryptography and Network Security Principles and Practices, Fourth Edition, London: Pearson Education, Inc., 2005.

[7] K. Christian, B. Katja, T. Markus, H. Stephan and R. Kai, "How to Enhance Privacy and Identity Management for Mobile Communities: Approach and User Driven Concepts of the PICOS Project," Mobile Business & Multilateral Security, 2010.

[8] K. I-Lung, "Securing mobile devices in the business environment," IBM Global Technology Services; Thought Leadership White Paper, pp. 2-10, October 2011.

[9] R. Bhasker and B. Kapoor, "Information Technology Security Management," in Computer and Information Security Handbook, Burlingtone, Morgan Kaufmann Publishers is an imprint of Elsevier, 2009, pp. 259 -267.

[10] R. Bhasker and B. Kapoor, "Computer and Information Security Hand Book," in Computer and Information and System Security, Burlington, Morgan Kaufmann Publishers is an imprint of Elsevier, 2009, pp. 259-257.

[11] K. Kathirvel, "Credit Card Frauds and Measures to Detect and Prevent Them," International Journal of Marketing, Financial Services & Management Research, vol. 2, no. 3, pp. 1-8, 2013.

[12] M. S. Gaigole and M. A. Kalyankar, "The Study of Network Security with Its Penetrating Attacks and Possible Security Mechanisms," International Journal of Computer Science and Mobile Computing, vol. 4, no. 5, p. 729, 2015.

[13] L. E. George, Faisel G. Mohammed, and Ibtisam A.: “Effective image watermarking method based on DCT,” Iraqi Journal of Science, 2015, vol.56, No.3B, pp

[14] A. Brodacki, J. Tarkowski, and J. Flis,:"Genetic distances in hens estimated with protein genes frequencies and procedures of DNA analysis",Department of Biological Bases of Animal Production. 2003, Electronic Journal Of Polish Agricultural Universities Volume 6, Issue 2

# *Identity Management Based on Frontal Facial Recognition for Voters Register in Zambia*

*Lubasi Kakwete Musambo*[1]
*School of Engineering*
*Dept. of Electrical & Electronics Engineering*
*The University of Zambia*
*Lusaka, Zambia*
[1]*e-mail: lubasimusambo@gmail.com*

*Jackson Phiri*[2]
*School of Natural Sciences*
*Dept. of Computer Science*
*The University of Zambia*
*Lusaka, Zambia*
[2]*e-mail: jackson.phiri@cs.unza.zm*

***Abstract*— biometric technology offers a great opportunity to identify individuals, authenticate individuals and separate individuals. Using these advantages, an election or voting model can be developed to perform elections for a country such as Zambia. Zambia currently uses a manual based voting or election model that heavily relies on paper presented documents that must be physically verified and or matched to existing prior collected information before an individual is allowed to participate in an election or a voting system. This paper proposes a frontal facial election based biometric model that can be used to rid the current election system of redundancy and introduce a paperless, accurate and efficient identification, authentication and voting process. A baseline study conducted shows that biometric authentication based on this proposed model improves a work related process such as a voting system. We start by introducing the elements that make a biometric model ideal, we then give an insight into the Zambian based election system and then we review various biometric technologies available and then finally introduce our biometric model.**

***Keywords— identification, biometric, election, frontal-facial, authentication, model***

## I. INTRODUCTION

A Biometric is a measurement and statistical analysis of people's unique physical and behavioral characteristics. Biometrics can be collected from either a physiological characteristic or a behavioral characteristic [1]. The essence of biometrics is to accurately distinguish an individual by their inherent physical or behavioral feature. A physiological characteristic is a relatively stable human physical feature. An example of a physiological characteristic is a fingerprint, retina and iris pattern, or a hand-geometry pattern. Physiological measurements are static and non-alterable. This type of measurement is unchanging and irreversible or permanent apart for deformity caused by external significant duress such as ailment or physical injury [2] [3]. A behavioral characteristic on the other hand attempts to resemble a person's psychological makeup. This is affected by a person's build stature and gender among others. Behavioral characteristics can be identified in activities such as speech, hand-writing speed and pressure exerted on paper when writing among others [4]. Four methods of biometric authentication systems were reviewed employing both physiological and behavioral characteristics. These have been reviewed in terms of basic operation, advantage and disadvantage of implementation.

Developments in biometrics entail that within-person variation factors have been taken into account at development as incoherencies can be determined with a level of accuracy by applying:

$$\cup ij = \frac{|a_i - a_j|}{r_{ij}} \quad (1)$$

where |ai−aj| is the magnitude of the vector difference between the two feature variations or drifts $a_i$ and $a_j$ while $r_{ij}$ is the distance between the corresponding feature locations (the variation). The combined potential energy of the drift map characterized by *K* feature drifts is given by [5] :

$$C = \sum_{i=1}^{K} \sum_{j=i+1}^{K} U_{ij} \quad (2)$$

It therefore follows that *'the lower the potential energy C'*, then it is more likely that the images belong to the same person [6] [5].

This then ensures that a biometric feature has longevity of integrity as long as the subject is alive. This consistency of a biometric feature is tied to the fact that a biometric signal is constant in time save for exogenous circumstances like injury or illness. To achieve biometric consistency, a match which uses a raw signal or fresh input (the biometric template or BT) must be collected from the signal directly at feature matching (the biometric signal or BS). Therefore the biometric, B governed is by BS, BT and B. It therefore follows that a stable biometric signal is a function of [6]:

$$〚BT〛_s = f(B) \quad (3)$$

## II. UNDERSTANDING VOTING SYSTEMS IN ZAMBIA

Applying a secure biometric infrastructure is key in ensuring that organisational or private data is well managed and accessed only by the intended party. It is important that a possibility to authenticate only those individuals that are registered as voters in the Republic of Zambia exists [2] [4] [7].

Elections in the Republic of Zambia are held every 5 years [8] [9]. These elections are held so that the citizenry can elect or choose their preferred leaders. Leaders are categorized into:

a. President,
b. Member of Parliament,
c. Mayor,
d. Council Chairperson and
e. Councilor.

This allows for free choice on the part of the voter. This systems allows the voted for, to run government affairs on behalf of the citizens for a period of 5 years unless under exigency circumstances like death mental health sicknesses and others that may dwell on thieving among others [8].

This type of governance in Zambia was introduced circa 1990 and has been this way to date [10].

The current system of voting introduces issues of ethics and among them is an issue of Identity (ID) theft. ID theft is stealing one individual's personal details that are used to identify and authenticate that the bearer of the details is indeed who they purport to be. Reasons for this theft span from gaining advantage in various forms such as by use of another's identity for fraud purposes or simply to bar the owner of the identity from exercising certain activities such as access to certain facilities [11] [12]. ID theft may damage an individual's reputation [13] and breed war if not countered in events such as elections. A need to stop this activity arises.

The development of Information Systems (ISs) for government operations has enabled a better service delivery in certain sectors of the Zambian economy [14]. There is however a need to ensure that all other government institutions are introduced to e-governance. One such institution with a great impact on the greater Zambia is the ECZ (Electoral Commission of Zambia).

The ECZ is mandated to conduct National, Parliamentary, Mayoral, Councilor and Council Chairmen/Women for the Republic of Zambia. A by the way mandate is to hire the ECZ to conduct other elections such as the FAZ (Football Association of Zambia) or political party elections [9].

In all these election activities of the ECZ, one thing is paramount; a Zambian is involved directly as a participant in an election. Due to this, a need arises to ensure that:

a. An individual cannot vote twice in the same election for the same participant,
b. An individual must be eligible to participate in that election,
c. An individual cannot be represented in proxy,
d. No ghost individual must participate in an election and
e. An individual who participates once in an election cannot deny ever participating.

An element of ID theft is present in the 5 issues highlighted. To ensure the 5 issues above are addressed, it is recommended that frontal facial biometrics is introduced primarily as a tool to eliminate identity theft in voting.

It is the purpose of this research to focus attention on the subject of ID management in an attempt to eliminate ID theft in voting in Zambia.

This study focuses on the registration, storage and authentication of a voter who enrols to participate in elections in Zambia.

It is therefore envisioned in this paper that an optimal business process map for the election process must be premised on optimum authentication. The researchers hold the view that authentication is the fundamental element that yields integrity and defines the ethics in voting.

We, therefore, present our argument by first understanding various literature present in the area of biometrics and voting and then present the issues present in the current literature and finally present our considered view of a biometric identity model that is built on security features that do not compromise performance but still deliver in terms systems expectations.

## III. LITERATURE REVIEW

In his paper citing Alexander Trechsel and Kristjan Vassil from their writing "Internet Voting in Estonia: A Comparative Analysis of Four Elections since 2005', European University Institute, 2010", [15] raises concerns about the security of a voter's detail if electronic mean are to be used to deliver an electronic vote [15]. The security question here borders around the concerns of whether Electronic systems can be attacked through various schemes such as denial of service, spoofing, viruses, and man-in-the-middle efforts.

This position held by [15] can be countered however by a having an encrypted authentication of a biometric nature for each voter. It is impossible to having matching biometric data in the form of hand geometry and fingerprint [16]. This feature of a biometric datum is enough to allay fears of ID theft. Secondly once a biometric datum is retrieved from an individual, cryptography takes over and this guarantees that an encrypted datum is non-understandable even if it is stolen because it is in firm that a third party cannot understand [1]. Once captured it is the duty of ECZ to ensure a third party has no physical access to the system.

In the "Analysis of an Electronic Voting System" paper by [17] there is a fear raised about a possibility of a man in the middle

attack which results into the theft of electronic data as it propagates [17]. Though this is a true possibility, a fully secured end-to-end encrypted system can be developed with a MAC (message authentication code). A MAC would counter the issues related to live data theft and subsequent tempering of the same by applying its image resistance properties [1].

[18] holds the view that in order for a voting system to be fair for political parties and voters; 7 elements must be present as follows:

a. **Authentication**: Only authorized voters should be able to vote;
b. **Uniqueness**: No voter should be able to vote more than once;
c. **Accuracy**: Voting systems should record the votes correctly;
d. **Integrity**: Number of casted vote must not be modified;
e. **Verifiability**: Possible to verify that votes are correctly;
counted in the final tally.
f. **Auditability**: Reliable and demonstrably authentic election records.
g. **Reliability**: Systems should work robustly, even in the face of numerous failures.

In this work, [18] determines that authentication is a central element in elections and should that authentication for some reason have a problem and result into a failed authentication then a failed election may exist. [18] further states that a failed authentication may result into non acceptance of an elected government. It is our position that to counter [18] a biometric authentication can be used. Because most biometric systems use a single trait to perform the authentication as pointed out by [19], fusion method of a digital identity as defined by [2] , [4], [3] and [19]can be used to ensure the authentication utilises multiple biometric units that can counter single biometric entity use.

## IV. Reviewing Biometric systems Currently in Use

*1. Fingerprint Authentication:* Fingerprints are made up of ridge patterns on a person's fingers. These ridge patterns have capacity to uniquely distinguish and identify individuals. Fingerprint features are made up of arches, loops, and whorls. An individual fingerprint will exhibit at least one of these major features. The minor details that are collected from these fingerprint features are referred to as minutiae. Figures 1 and 2 show a finger print sample and finger print features. The authentication processes is an automated method of verifying a match among different human fingerprints [20] .

*Advantages:*

i. Individualistic features guarantee authentication of subject [3].
ii. Systems are relatively inexpensive to purchase and install.
iii. Longevity of life of the fingerprint pattern's individualistic feature composition guarantees long term usage [3].
iv. Once in use a subject does not have to rely on memory for passwords as fingerprint authentication will guarantee access.
v. A fingerprint identity point cannot be spoofed [21].

*Disadvantages:*

*i.* Limitation of capture is reduced to an individual finger with further limitation of capture reduced to a section or part of that finger only and not the entire finger.
*ii.* Susceptible to FAR (false acceptance error) whereas a wrong subject is enrolled and allowed access.
*iii.* Hand injury (fingers included), chemical prone jobs and labour prone activities such as brick-laying or metal fabricating present a within-person variation that makes the reading and capture of finger prints difficult.
*iv.* Washing with a soap detergent or submerging a finger in water for period of time (approximately 30 minutes) works as a contraceptive to finger-print scanners and this may impede the scanners from capturing or enrolling the finger prints until the finger reverts to its original form it was in during capture or enrolment [1].

*Figure 1: Fingerprint Image Sample*

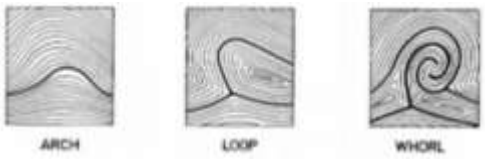



*Figure 2. Fingerprint features [23]*

*2. Retina Authentication:* This is one of the two forms of eye biometrics; the other being iris

recognition. This form of biometrics is one of the most secure authentication systems in place today. The installed technology requires that an impression of a retina pattern must be taken and stored. The authentication process involves evaluating a subject's retina with a stored version (impression enrolled) of that subject's retina. Retina recognition has a low FAR (false acceptance error) as well as low rejection rates [22]. An image sample of an eye is shown in figure 3.

*Advantages:*

i. Different even in identical twins.

ii. Highly specific with unique structure shape and limits the possibility of fake retina presentation.

iii. Longevity of structure throughout life time of subject.

iv. Wearing of glasses or contact lenses does NOT work as a contraceptive to technological accuracy.

v. High accuracy and High recognition process speed.

*Disadvantages:*

i. Eye injury or sickness may render this biometric system ineffective.

ii. Intrusive technology and may not be welcomed by many individuals.

iii. Lighting may affect the accuracy of the reader.

iv. Fairly expansive to acquire when compared to other systems of biometrics.


*Figure 3. Eye Image Sample – for iris Recognition*

*3. Voice Authentication:* This technology allows the conversion of voice or sounds from human voice into an electrical signal that can be coded. Voice recognition software is designed to identify an individual via their unique voiceprint. Voiceprints are generated from physical characteristics of an individual's throat in conjunction with their mouth. Research indicates that no two voices are the same and therefore voice biometrics provide a rare opportunity to use one's voice to authenticate or identify individuals [7]. A sample of a voice pattern is shown in figure 4 below.

*Advantages:*

i. No need for user training as users can simply speak into the voice biometric reader.

ii. Voice communications is a natural activity for human beings.

iii. Voice communications eliminates the need to learn keyboard operations (and in this way helps to bridge the gap between the able-bodied and individuals who experience restricted capabilities in hand based motion activities such as writing). By eliminating the learning aspect, voice overcomes the need to learn how to operate some complex biometric technology's operations.

iv. It eliminates the need to be accurate in written statements as is for password based authentication.

v. Because one uses voice, the speed of operation is enhanced. People generally speak faster than they are able to write.

*Disadvantages:*

i. Impulse noise may affect the accuracy of the voice signal and render the system ineffective.

ii. Microphone proximity must be precise for the system to work well.

iii. A pre-recorded audio may by-pass this system.

iv. A person may speak different languages and this may affect the accuracy of the device should that individual use a different language or dialect.

v. Certain words have a homonym characteristic, this may affect the accuracy of the device.

vi. The learning curve for the system may be long as it is trained per voice.

vii. Most voice controlled biometrics are expensive.

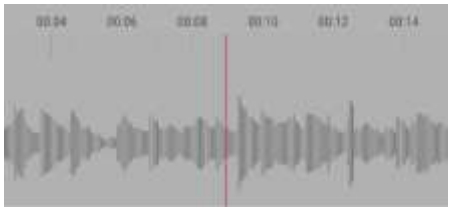


*Figure 4. Voice Print. Adapted from [20]*

*2. Face:* Facial biometrics divides into two aspects namely the face detection and face recognition programs. Face recognition extracts a face from a given image while face

recognition compares a captured face against saved faces in order to match the face. The entire process is run by a series of complex algorithms. One of the options of face recognition is to select features of a face and match those features to a face. Figure 5 below shows a facial image sample with facial image mapping that is used to collect facial features. The facial features or dataset is normally stored in a database. In ideal situations this database must be encrypted to achieve sufficient security [23].

*Figure 5. Facial Image Sample with facial map.*

*Advantages:*

i. Non-intrusive technology and can be performed stealthily without the subject knowing, therefore, proves ideal for investigation purposes.

ii. Certain algorithms can be adjusted to scan a large scale of a population and thus this technology proves ideal in crowded environments.

iii. Ideal for person tracking and incident reporting.

iv. User friendly as far as users are concerned as no need of complex training for the subjects to be captured.

v. Can be developed and run from a basic computer camera without buying any other tools. This proves to be one of the strongest advantage and reduces the cost of this technology exponentially.

vi. Some easy to install ready to use pre-trained facial calibration tools are available. This again reduces cost of setup.

vii. Facial biometric algorithms have a within-person variation calculation that can detect aging and basic facial deformity and reduce a face to a known variable [24].

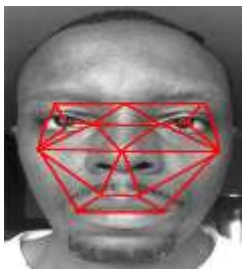
*Disadvantages:*

i. Certain algorithms may NOT work well on black faces.

ii. Light conditions and camera capabilities may affect the accuracy of the technology.

iii. Within-person variations may affect the accuracy levels of the technology [25].

iv. When used for security purposes, extra equipment to provide lighting can increase cost of setup.



## V. METHODOLOGY

The methodology is divided into 10 sections as follows:

(1) Baseline which address the process undertaken to determine the information requirements. (2) Descriptive research design which address aspects of research that are within the region of illustrative research concerning the research. (3) Target group which speaks to the identified target for this study. (4) The sample size and why it was taken. (5) The data collection tools which explains which tools were used in the study. (6) The data analysis which explains which analysis tool has been used. (7) The ethical consideration which explains the ethical position of the research. (8) The limitation of the baseline study which states how far the baseline study was stretched. (9) The presentation of the findings section to demonstrate our findings concerning our biometric model. (10) The system design detailing the technical aspects of the system.

### *A. Baseline Study*

Baseline study was used in order to investigate the awareness levels and understanding of biometrics among individuals and organizations and to determine if a biometric standard that defines use and management of biometrics in Zambia is in use. Additionally, the use of the baseline study was to establish if organizations utilize biometrics for one function or another do so within a framework that is defined by government regulators of ICTs. The result of the baseline was used to develop the biometric software model. The model was validated to ensure it would fit into the use of conducting an election through frontal facial biometrics. As a result, the study used a mixed methods research methodology to analyze the data from the respondents.

The researchers hold the view that mixed methods research is the type of research which involves the use of more than one approach to or method of design, data collection or data analysis within a single program of study (e.g. both qualitative and quantitative research), is ideal as it integrates the different approaches or methods occurring during the program of study [26]. Mixed methods approach to research, helps researchers to incorporate methods of collecting or analyzing data from the quantitative and qualitative research approaches in a single research study. Similarly, researchers can collect or analyze numerical data which refers to quantitative research coupled with narrative data which is the standard for qualitative research such that research question (s) are addressed as defined in any typical research study. Mixed methods designs also provide pragmatic advantages when exploring complex research questions.

Qualitative data was used to deepen understanding of survey responses while the statistical analysis was done to provide detailed assessment of patterns of responses.

*B. Descriptive research design*

Descriptive research is meant to provide a picture of a situation as it naturally happens. As such, it could be used to justify current practice and make judgment and also to develop theories. As a matter of fact, descriptive research [27]. A descriptive research design is used to explain the state of affairs at present. The researchers used it to obtain pictures of the current prevailing Election registration systems of registration in the Republic of Zambia.

*C. Target group*

The study was made up of eight types of target groups of the biometric authentication ecosystem comprising: ICT regulators, Standardization bodies, Consumer protection authorities, students in higher education institutions, banks, Government Ministries and departments, Health Support Institutions and general users. The mentioned respondents were sampled from the: University of Zambia (UNZA), Matem University, Bank of Zambia, Proflight, Stanbic Bank, Zambia Bureau of Standards (ZABS), Zambia Information Communication Technology Authority (ZICTA), Ministry of Home of Affairs (Passport Office and Citizen Registration Office), John Snow Initiative (JSi), Ministry of Commerce – National Technology Bureau, Ministry of Information and Broadcasting Services, Competition and Consumer Protection Commission (CCPC), Zambia Development Agency (ZDA) and the study area comprised Lusaka.
The significance of targeting the mentioned groups was meant to capture primary data from the mentioned area through purposive sampling. Purposively sampling signifies how the researcher sees sampling as a series of strategic choices about whom, where and how one does one's research.

*D. Sample size*

A total number of 100 respondents were randomly selected for interviews. The sample size was manageable and wide enough for valid generalization to the biometric ecosystem in Zambia.

*E. Data collection tools & Systems Design*

*1) Self-administered questionnaires*

The self-administered questionnaires were used to collect information from all the respondents. The use of questionnaires was not only simple to administer, but questionnaires were also relatively inexpensive to analyze. When alternative replies are provided in the questionnaires, respondents are able to understand the meaning of questions more clearly [26].

To validate the software, a validation question answered in tandem with software operations was done.

*F. Data analysis*

Data analysis for the study was done by computer based software known as Microsoft Excel. Microsoft Excel is a paid for computer program that is developed and maintained by the Microsoft Corporation [28].

*G. Ethical consideration*

Ethical clearance through authorization was awarded to the researchers by the institutions where the research was conducted from, by means of introductory letters which were given to authorities and respondents. Similarly, all questionnaires administered, did not allow respondents to disclose their names or any information that would review their status and ultimately compromise on confidentiality.

*H. Limitation of the baseline study*

The prototype is designed to enhance the Election registration processes in the Republic of Zambia and as such live tests can only be performed at the Ministry of Home Affairs and partly with the Ministry of Health. Getting permission for a live test with these institutions implies collecting citizen data. This was inhibitive. The other limitation was from some target groups like: commercial banks and some government offices that deal with citizen data who entirely refused to take part in the survey for fear of disclosing the data they collect to the general public.

*I. Presentation of findings*

The findings have been presented in the section labelled, "*Findings*".

*J. System design*

The system design is arranged into 4 sections including the system design as follows:

The first is explanation of the Haar Cascade algorithm, the second is the presentation of the Current business process in the Election voting process (which includes registration) and highlights the current problematic areas. The third is the proposed Election voting business process. The fourth presents the overall business process flow for the proposed model. The fifth section introduces the system interaction that a typical user encounters when they interact with the system illustrated using UMLs' Use Case and Interactive Sequence Diagrams.

*i. Understanding the Haar based Frontal Face Biometric Algorithm*

Based on a rapid object detection scheme based on boosted cascade of simple feature classifiers introduced by Paul Viola and Michael Jones, a facial biometric model can be developed based on Haar-like features and implemented to detect and recognise a student's face. This recognition facility allows for authentication. Facial features to form a Haar classifier are collected after a facial mapping as shown in figure 8 above. The biometric model utilises Haar basis features as used by Papageorgiou et al [29].

An adaption of the algorithm based on an OpenCV Open Source technology which is readily available from OpenCV has been used. This algorithm uses Haar like features and OpenCV pre-trained classifiers for face detection. A classifier is a program that can decide whether an image is positive or not. A positive image is an image face (image having a face) while a negative image is a non-face image. Classifiers are trained from a huge volume of faces (both positive and negative images) to learn how to classify a new image correctly. This is a machine learning concept. The classifiers used for this student authentication is the HaarClassifier which is earlier developed by Viola et al [30]. Haar Classifiers process data in grey scale (non-colour). Colour is inconsequential in determining whether an image has a face or not.

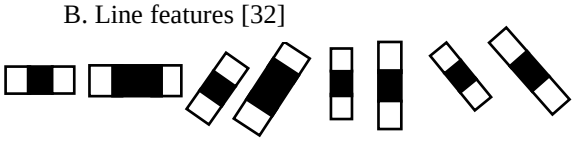
*ii. Haar Classifier function logic*

Viola et al states each object has features that are unique and can be used to identify and recognize that object. Haar features can be picked out from edge, line, center and diagonal features of an object as shown in figure 6.

A. Edge features [32]

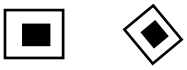
i ii iii iv

B. Line features [32]



i ii iii iv v vi vii viii

C. Center-surround features [32]



D. Diagonal line Features [37]

i ii i

*Figure 6. Example feature determination for extraction [31]*

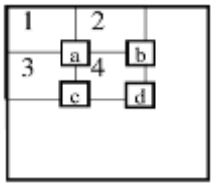
Edge features are characteristics of an image that are unique and at unique distances from each other. No two people share the same features. The features can be mapped by placing an object identifying feature. A biometric model developed to pick up the readings from the facial recognizer can pick up the features and collectively store them to perform identification and recognition. The features can be collected into small elements referred to as a weak classifier which when collectively used identify and recognize an object [31]. Feature collection is done via rectangles. Haar like features consist of two or more rectangular regions enclosed in a template. Each of the rectangles is a window that is placed on an image as shown in figure 7 that is to be captured and recognized. A feature is extracted from subtracting the sum of pixels under the white part from the black part of that window (rectangle).

In determining the haar like features an understanding that the area around the eyes have a darker area then the nose bridge is used. This view is also held for the cheeks (brighter than other areas), though the data from the cheeks is not necessarily used.

Rectangles are placed on an image so as to pick the features using a weak classifier. The features of a rectangle are computed using an integral function of the form:

$$ii(x, y) = \sum_{x' \le x, y' \le y} i(x', y'), \qquad (4)$$

In this function an object or image at location *x, y* contains




the sum of pixels above and to the left of *x, y* inclusive.

Where *ii (x, y)* and *i (x, y)* is the original image. Using the following pair of recurrences:

$$s(x, y) = s(x, y - i) + i(x, y)$$

$$ii(x, y) = ii(x-i, y) + s(x, y)$$

(Where *s(x, y)* is the cumulative row sum, *s (x-a) = 0,* and *ii (-i, y) = 0).* Using the integral image any rectangular sum can be computed in four array references [31] [32] [30].

The rectangle itself can be understood to have an object of pixels *W x H* (i.e.to say width x Height) [30]. Figure 8 below shows the determination of a rectangular region of an integral image.

*Figure 8. Rectangular regions of an integral image [37]*



*Figure 7. Feature Determination. Adapted from [38] [39]*

To determine the sum of pixels, the logic can be deduced as follows:

$$a = sumRec(pixels) \quad (5)$$

$$b = 1 + 2,$$

$$c = 1 + 3$$

$$d = 1 + 2 + 3 + 4$$

The sum is then derived as $d + a - (b + c)$.

Using the OpenCV library of face detectors and recognizers a function can be developed into a web based biometric application that can perform an online web authentication during elections where an individual would be participating in a voting process.

### *iii. Business Processes*

a. Current business process

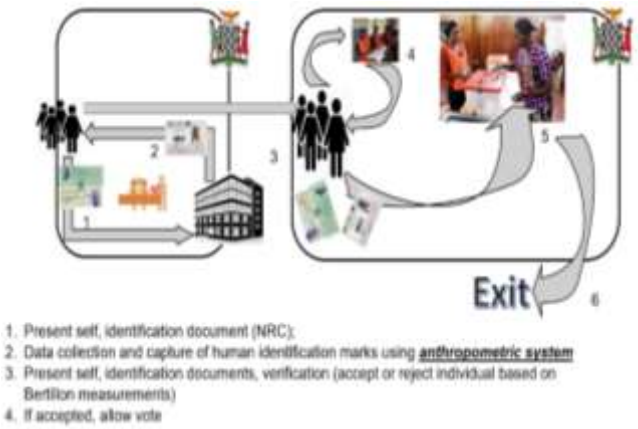



*Figure 9. Current business process – Voting Process*

A pseudo-code of the current business process is as follows:

*START*
*Get individual*
*Present identification documents*
*Collect personal data using Bertillon system*
*WHILE individual available*
*Verify identification documents*
*Authenticate based on Bertillon measurement*
*IF Bertillon measurement = true*
*Then allow vote process*
*ELSE*
*Reject vote process*
*ENDIF*
*ENDWHILE*

b. Proposed business process

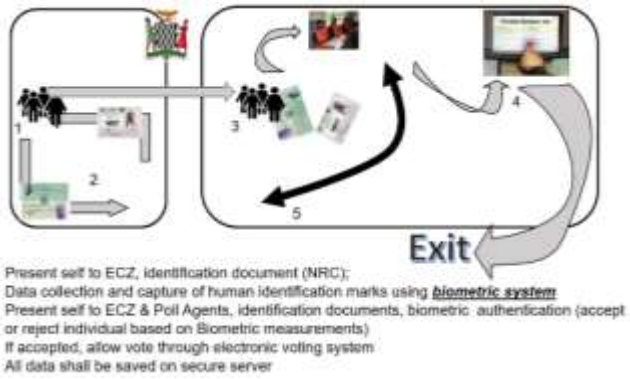



*Figure 10. Proposed Business process – Voting Process*

A pseudo-code of the proposed business process is as follows:

*START*
*Get individual*
*Present identification documents*
*Collect personal data using biometric system*
*WHILE individual available*
*Authenticate individual using biometrics*
*Authenticate individual using biometric measurement*
*IF biometric measurement = true*
*Then allow vote process*
*ELSE*
*Reject vote process*
*ENDIF*
*ENDWHILE*

Figure 9 above shows the current business process for a voting process. The challenges in the current system can be identified as:

- Bertilloni systems are not accurate measure to identify people [33];
- Individuals may lose identification documents through theft and others;
- Slow process;
- Untrustworthy among the political players [34] and
- Defaced documents may result into a reject of a vote process.

To overcome the challenges identified above, a proposed business process as shown on figure 10 can be implemented.

The proposed model would have the following advantages:

- A biometric is constant [6];
- An individual may lose identification documents or identification marks may be defaced but an individual can still be allowed to participate in a vote process;
- Lessens paper and

- May increase trust due to reduction in human to human interaction.

The methodology used for the analysis, design and development of the software system is the object-oriented systems development methodology (OOSDM) [35] [36]. This research study utilized some of the diagrammatic representations that are present in the unified modeling language (UML) in order to visualize the system from various perspectives [35].

The object-oriented system development (OOSD) approach that was used in the system development process is one that is use case driven. The object-oriented system development life cycle (OOSDLC) was used for the system development in this research study in order to show multiple iterations to be carried out throughout the entire development cycle for the system to be gradually built in small modular increments [36].

## VI. Results

In order to ensure this software model we propose can function adequately and meet election or voting processes in Zambia, a validation which is the process of ensuring that a piece of software system or module meets its' systems specification and delivers on its' intent. Aspects of quality control may need to be undertaken to ensure this is met [36]. The validation process has been undertaken by allowing specialist stuff to run the software model and after which respond to questions set out in questionnaire form.

### *A. System Reliability, Usability, Performance and Portability*

#### *1) Learning Curve*

Research participants were requested to a question probing how the participant viewed the learning curve of using the election system. In Figure 11; SA = Strongly Agree, A = Agree; N = Neither Agree nor Disagree; D = Disagree and SD = Strongly Disagree.

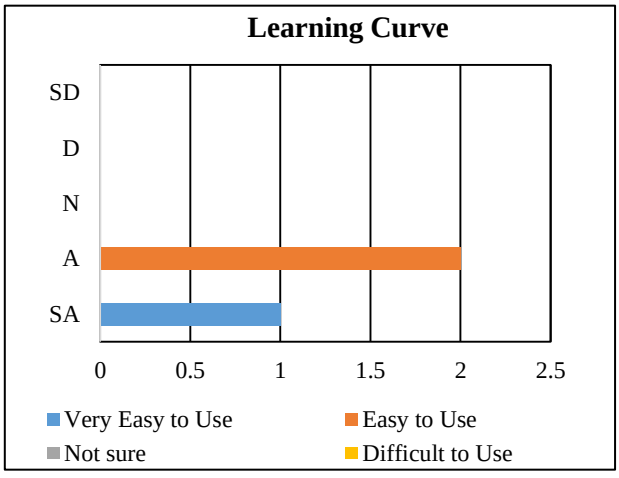



*Figure 11. Election System Validation – Learning Curve of System Use*

#### *2) Satisfaction of Use*

Research participants were requested to state the satisfaction level they obtained from system use of the election model. Figure 12 below illustrates this result.

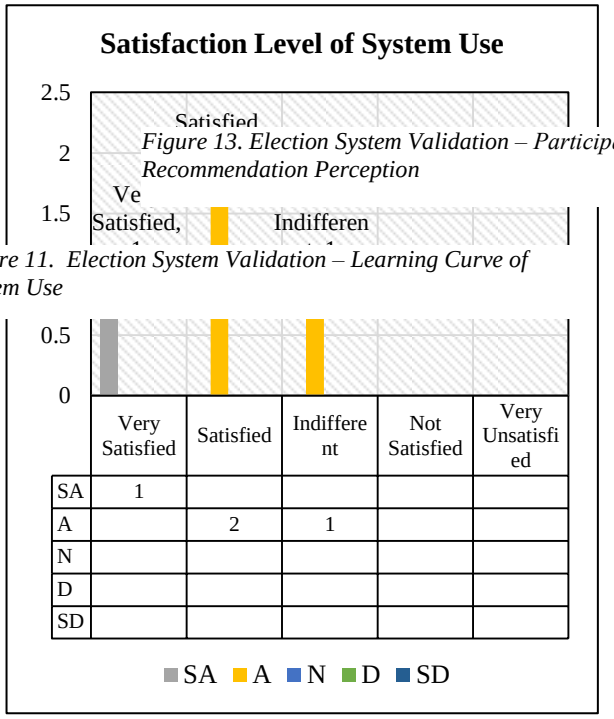



*Figure 12. Election System Validation – Participant' Satisfaction Level of usage of the System*

#### *3) Recommendation to use system for Elections*

Figure 13 shows responses the research participants gave to the question asking them to recommend the Election registration system for Election registration processes in government processes. 2 out of the 6 participants said they would recommend the system, 1 was indifferent, and the other 2 were silent.

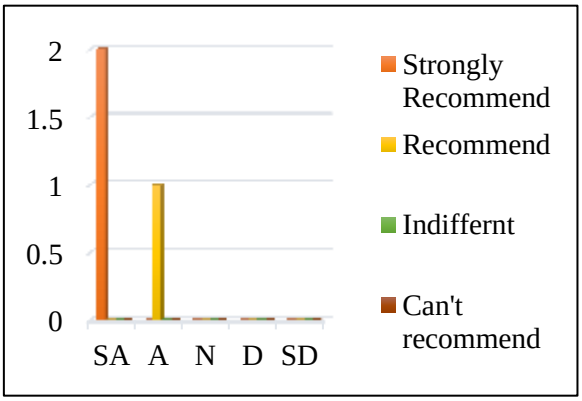



*Figure 13. Election System Validation – Participant's Recommendation Perception*

*4) Interest to use the Election Registration Processes*

Figure 14 below illustrates the research participant's interest to work with the Election registration system often. As can be seen 2 out of the 6 participants said they would be very must interested to work with the system often, 1 participant held the view that they would like to work with the model while 1 participant was not sure.

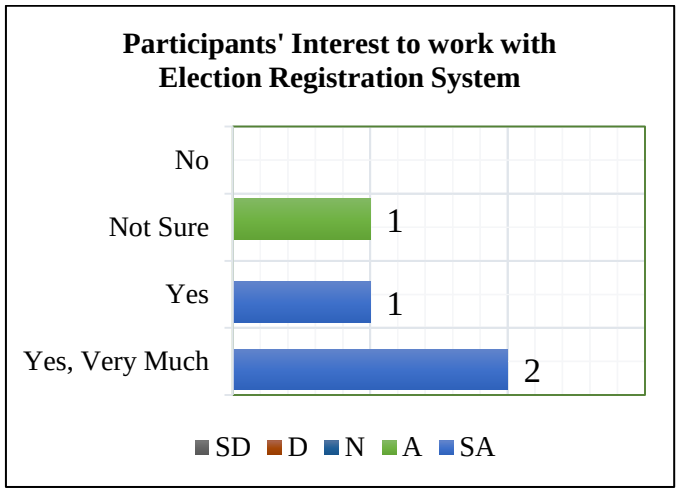



*Figure 14. Election System Validation – Participant's Desire to work with System*

*5) Election System's Ability to meet Job Function*

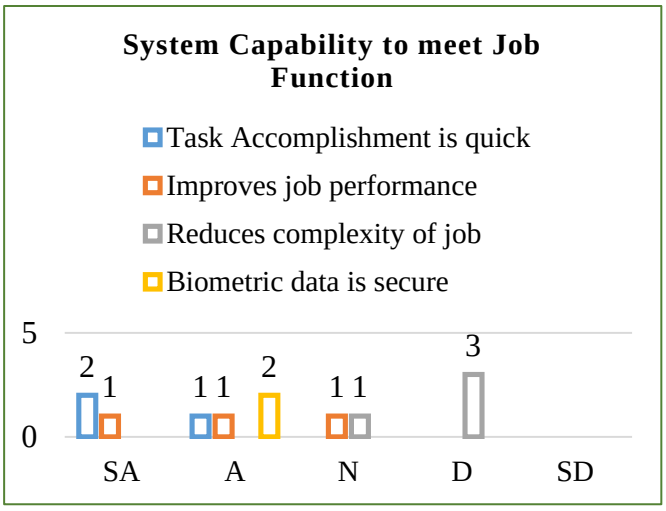



*Figure 15. Election System Validation – System Capability to meet job function*

The research validation participants were requested to comment on how they viewed the system's ability to meet job functions that a Election registration officer would undertake. Figure 15 above shows that out of the 6 research validation participants

*6) System Portability*

Given that the Election Registration system can only run if Python, OpenCV's Boost Cascade and Xampp control which are open source programs are installed; the research validation participants were asked to determine if the system was portable enough. Figure 16 below gives the responses to this question. As can be seen, 67% of the participants stated that the system was portable enough while 33% could not respond to the question and none of the participants stated that the system was not portable.

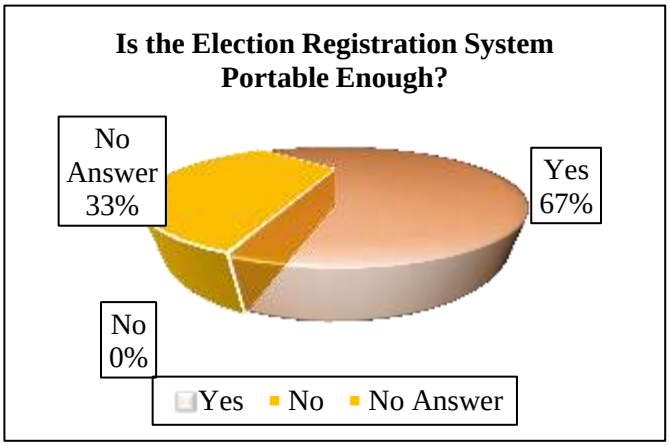



*Figure 16. Election System Validation – System portability Status*

*7) System Resources*

Given that the Election Registration system requires systems resources of a computer with at least 4GB RAM, Hard disk capacity of at least 500GB and a web camera with a resolution of at least 0.9MP 16:9 (1280 x 720); the system validation research participants were asked to state whether these resources are too ambitious. Figure 17 below shows the responses and as can be seen 3 participants stated that the resources were not too ambitious, 2 participants could not respond to the question and 1 participant was not sure.

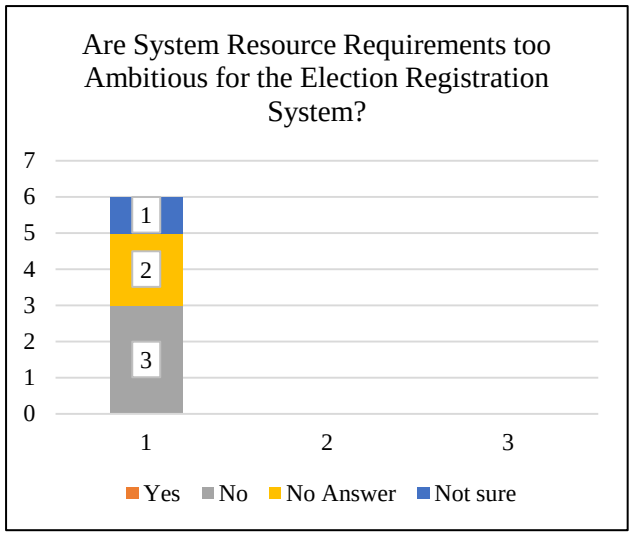
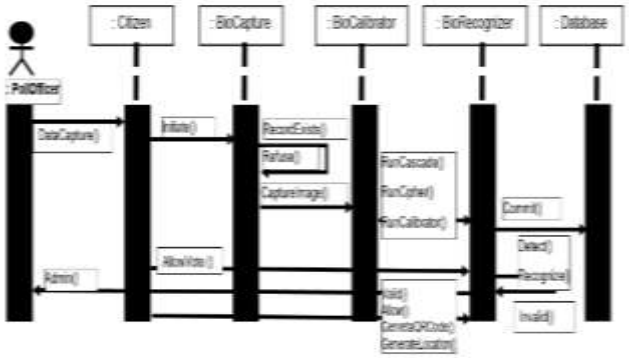



*Figure 17. Election Registration System Validation – System Resource Requirements*

Figure 18 below shows he UML interaction sequence diagram for the election system model. A citizen can be registered only once after that the recognizer would perform the authentication for every other function.



The biometric authenticator described in the paper was implemented on authenticating individuals at different times of the day. This image set collected used 3000 image faces. The system achieves a person detection rate of 66% with a 33% false acceptance error.

## VII. Discussions

The biometric model is able to yield a positive result of 66%, the false acceptance rate of 33% has been determined to be due to lighting conditions when the images are captured and the dark faces enrolled. Performance of the model has been

*Figure 18. UML interaction sequence for Election Authentication.*

observed to be higher or accurate when lighter faces are used. The researchers hold the view that that the darker regions around the eyes become fairly complex for the algorithm to determine on black faces. Improving lighting conditions has been observed to correct the recognition and detection process.

A web camera mounted on a laptop or computer is sufficient for this task. It must however be understood that sufficient research is needed into ensuring that false positives are dealt with as frontal face biometrics presents false positive errors. It is recommended that ISO 24745 is used to guide in the secure management and usage of biometric data.

## VIII. Conclusion and Summary

In this paper, we give the results of the implementation for an election authentication system based on frontal facial biometrics. The Test results shows the proposed system was able to give up to 66% accuracy level. For a developing country like Zambia, this would be a good starting point. The frontal facial biometrics uses OpenCV's boost algorithms which are open source and readily available for adaptation.

In this paper, we began by a review of the various forms of biometrics that can be used in authentication systems. We then presented the general security challenges in elections especially for developing countries such as Zambia. One of the solutions to these challenges is the integration of biometrics features in the authentication systems. A cheaper solution for most developing countries is the use of open source tools and cheaper devices. Our study was proposing the use of OpenCV for Biometric Facial recognition and simple cheaper Web Camera such as one that comes integrated in most mobile computing devices.

## IX. References

[1] K. Martin, Everyday Cryptography: Fundamental Principles & Applications, New York: Oxford University Press, 2012.

[2] I. J. Agbinya, N. Mastali, R. Islam and J. Phiri, "Design and Implementation of a Multimodal Digital Identity Management system using fingerprint matching and face recognition," *Broadband and Biomedical Communications (IB2Com)*, pp. 272-278, 21-24 Nov 2011.

[3] J. Phiri, T.-J. Zhao, H. C. Zhu and J. Mbale, "Using Artificial Intelligence Techniques to Implement a Multifactor Authentication System," *International Journal of Computational Intelligence Systems*, vol. 4, no. 4, pp. 420-430, 2011.

[4] J. Phiri and J. I. Agbinya, "Modelling and Information Fusion in Digital Identity Management Systems," in *International Conference on Systems and International Conference on Mobile Communications and Learning Technologies, 2006.*, Morne, Mauritius, 2006.

[5] B. E. &. B. Sankur, "Effects of Aging over Facial Feature Analysis and Face Recognition," *Bogaziçi Un. Electronics Eng. Dept.*, pp. 1-4, 2010.

[6] L. O'Gorman, "Comparing Passwords, Tokens, and Biometrics for User Authentication," *PROCEEDINGS OF THE IEEE*, vol. 91, no. 12, pp. 2021-2040, DECEMBER 2003.

[7] V. a. Tripathi, "A Comparative Study of Biometric Technologies with Reference to Human Interface," *International Journal of Computer Applications*, vol. 14, no. 5, pp. 1-6, 2011.

[8] GRZ, *The Constitution of Zambia*, GRZ, 2016.

[9] ECZ, "Elections," 12 February 2017. [Online]. Available: www.elections.org.zm/elections.php. [Accessed 12 February 2017].

[10] T. Kambilima, "History of elections in Zambia," *Zambia Daily Mail*, 2016..

[11] Sandi, "Identity Theft: When They Steal "You"," TechtTrends, 3 March 2014. [Online]. Available: www.techtrends.co.zm/identity-theft-steal/. [Accessed 25 November 2018].

[12] L. Times, "Zambian Pleads Guilty to Identity Fraud in the US," Lusaka Times, 27 July 2011. [Online]. Available: www.lusakatimes.com/2011/07/27/zambian-pleads-guilty-identity-theft. [Accessed 26 November 2018].

[13] A. G. Johansen, "4 Lasting Effects of Identity Theft," Symantec Corporation, 2018. [Online]. Available: www.lifelock.com/learn-identity-theft-resources-lasting-effects-of-identity-theft.html. [Accessed November 25 2018].

[14] ZPPA, "e-Procurement System," 12 February 2017. [Online]. Available: www.zppa.rg.zm/e-procurement-system. [Accessed 12 February 2017].

[15] T. Hall, "Internet Learning, Internet Voting: Using ICT in Estonia," *IPSA*, p. 31, 2012.

[16] M. Rose, "Biometrics," 21 February 2017. [Online]. Available: www.searchsecurity.techtarget.com/definition/biometrics. [Accessed 21 February 2017].

[17] A. S. A. D. R. S. W. TADAYOSHI KOHNO, "Analysis of an Electronic Voting System," *IEEE Symposium on Security and Privacy 2004*, p. 23, 2004.

[18] S. Yadav and A. K. Singh, "A Biometric Traits based Authentication System for Indian Voting System," *International Journal of Computer Applications*, vol. 65, no. 15, pp. 28-32, March 2013.

[19] D. Jagadiswary and D. Saraswady, "Biometric Authentication using Fused Multimodal Biometric," *Elsevier - International Conference on Computational Modeling and Security*, vol. 85, no. 2016, pp. 109-116, 2016.

[20] R. Saini and N. Rana, "COMPARISON OF VARIOUS BIOMETRIC METHODS," *International Journal of Advances in Science and Technology*, vol. Vol 2, no. I, pp. 1-7, 2014.

[21] N. Ferguson, B. Schneier and T. Kohno, Cryptography Engineering: Design Principles and Practical Applications, Indianapolis: Wiley, 2010.

[22] J. M. Stewart, E. Tittel and M. Chapple, Certified Information System Security Professional, Canada: Wiley, 2008.

[23] F. Alonso-Fernandez, J. Fierrez and J. Ortega-Garcia, "Quality Measures in Biometric Systems," in *IEEE*, 2011.

[24] A. Lanitis, "Facial Biometric Templates and Aging:Problems and Challenges for ArtificialProblems and Challenges for Artificial," in *AIAI-2009 Workshops Proceedings*, 2014.

[25] E. Bilgin and B. Sankur, "Effects of Aging over Facial Feature Analysis and Face Recognition," *Bogaziçi Un. Electronics Eng. Dept.*, pp. 1-4, 2010.

[26] S. MacDonald and N. Headlam, Research Methods Handbook, Manchester: Th Centre for Local Economic Strategies.

[27] P. Pandey and M. M. Pandey, RESEARCH METHODOLOGY: TOOLS AND TECHNIQUES, Romania: BRIDGE CENTER, 2015.

[28] Microsoft, "Get a better picture of your data," Microsoft, 13 April 2018. [Online]. Available: https://products.office.com/en-us/excel. [Accessed 13 April 13].

[29] A. Mohan , C. Papageorgiou and T. Poggio, "Example Based Object detection.," *IEEE Transactions on pattern Analysis and Machine Intelligence*, vol. 23, no. 4, pp. 349-361, 2001.

[30] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in *CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2001*, Cambridge, 2001.

[31] S.-K. Pavani, D. D. Delgado and A. F. Frangi, "Haar - like features with optimally weighted rectangles for rapid object detection," *Elsevier*, vol. 43, no. 160-172, pp. 160-172, 2010.

[32] R. Lienhart, A. Kuranov and V. Pisarevsky, "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection," *MRL Technical Report*, pp. 1-7, 2002.

[33] N. L. E. MUSEUM, "Bertillon System of Criminal Identification," NATIONAL LAW ENFORCEMENT MUSEUM, November 2011. [Online]. Available: http://www.nleomf.org/museum/news/newsletters/online-insider/november-2011/bertillon-system-criminal-identification.html. [Accessed 27 November 2018].

[34] M. Funga, "Voter apathy indictment on ECZ – HH," News Diggers, 27 July 2018. [Online]. Available: https://diggers.news/local/2018/07/27/voter-apathy-indictment-on-ecz-hh/. [Accessed 27 November 2018].

[35] I. Jacobson, M. Christerson, P. Jonson and G. Overgaard, Object-Oriented Software Engineering: A Use Case Driven Approach, Patparganj: Pearson Education, 2004.

[36] D. Avison and G. Fitzgerald, Information Systems Development: Methodologies, Techniques & Tools, Maidenhead, Berkshire: McGraw-Hill Education, 2002.

[37] M. S. Uddin and A. Y. Akhi, "Horse Detection Using Haar Like Features," *International Journal ofComputer Theory and Engineering*, vol. 8, no. 5, pp. 1- 4, October 2016.

[38] D. Yadav, R. Singh, M. Vatsa and A. Noore , "Recognizing Age-Separated Face Images:Humans and Machines," *Pone*, 2014.

[39] R. Rezaei, . H. . Z. Nafchi and S. Morales, "Global Haar-Like Features:A New Extension of Classic Haar Featuresfor Efficient Face Detection in Noisy Images," *R. Klette, M. Rivera, and S. Satoh (Eds.)*, pp. 302-313, 2014.

# A REVIEW OF MAJOR LOCAL AREA NETWORK SECURITY CHALLENGES

Jimmy Katambo[1], Mayumbo Nyirenda[2] and Jackson Phiri[3]
The University of Zambia
Department of Computer Science
Lusaka, Zambia.
[1]jimmy.katambo@cs.unza.zm, [2]mayumbo@gmail.com, [3]jackson.phiri@cs.unza.zm

***Abstract*—Based on the method of deployment, Intrusion Detection Systems (IDS) can be classified as Host Based IDS (HIDS) and Network Based IDS (NIDS). Network intrusion detection is a dynamic research area as intruders or attackers have increased attacks on all kinds of networking set-ups. Security policy is a main mechanism of information security management. The paper which is a review of the major Local Area Network Security challenges provides insight into major Local Area Network Security challenges. This paper reviews a number of articles in the areas of Local Area Network Security and discusses their major challenges. While there are a lot of security-related standards and guidelines which specify requirements for high-level security policies, implementation of network security policy still depends on interfaces provided by network security solutions. The need for tools to help network administrators in the network management process is increasing. Access Control Lists (ACLs) refer to security rules associated to network equipment, such as routers, switches and firewalls. Firewalls provide a mechanism for protecting enterprises from the less secure internet over which customers or partners transfer packets destined for the corporate network.**

***Index Terms*— Access Control, Authorization, Computer Security, Firewalls, Information Security Policy, Intrusion Detection Systems, Network Security Solutions, Packets, Policy Analysis, Policy Management.**

## I. INTRODUCTION

The paper which is a review of the major Local Area Network Security challenges provides insight into major Local Area Network Security challenges. It is very important to secure available resources on any corporate or academic data network because most of these networks connect to the Internet for a great deal of activities. Seeing that the network is under attack from hackers continually, network security technologies are ever evolving and playing catch-up with hackers. It is necessary to design security policies appropriate for both the servers and their clients. Network designs currently implement three levels of trust: most trusted, less trusted, and least trusted [1]. These support the objectives of network security which are Confidentiality, Integrity and Access. The responsibility for the design and implementation of network security is headed by the chief information officer (CIO) of the enterprise network [1]. In 1988 the U.S. Department of Defense established the Computer Emergency Response Team (CERT), whose mission is to work with the Internet community to prevent and respond to computer and network security breaches. Since the Internet is widely used for commercial activities by all kinds of businesses, the federal government has enacted stiffer penalties for hackers [1].

## II. IDENTIFY NETWORK THREATS

According to J. R. Vacca [1], Network security threats are divided into one of these two categories:
(1) disruptive type or (2) unauthorized access type. Firstly, most LANs are made as collapsed backbone networks using two switches namely; layer-2 or layer-3. The router or switch may fail to function due to power failure. An attack by viruses on the secondary storage may cause the network to fail thus leading to data loss.

Secondly unauthorized access can be internal (employee) or external (intruder). In both cases this is a person who would attempt to break into resources such as database, file, and email or web servers that they have no permission to access.

Although network security has increased over the years, the frequency of attacks on the networks has also increased. This is because the tools to breach the network security have become freely available on the Internet.

According to [1], the Computer Emergency Response Team (CERT), whose mission is to work with the Internet community to prevent and respond to computer and network security breaches, was introduced in 1988 by the U.S. Department of Defense established the Computer Emergency Response Team (CERT).

## III. ESTABLISH NETWORK ACCESS CONTROLS

These network controls are either software or hardware based and are implemented in a hierarchical structure to reflect the network organization [1]. Network control is in three categories namely; detect, prevent, and respond [1]. These categories correspond with the functions of the network control which are to detect an unauthorized access, to prevent network security from being breached, and finally, to respond to a breach.

In the role of prevention control a password can be used to allow the user to access the resource on the network. The role

of the detection control is to monitor the activities on the network and identify an event or a set of events that could violate network security. An example of such an event may be a virus, spyware, or adware attack. The detection control software must not only register the attack, but also generate or trigger an alarm to notify of an unusual event so that a corrective action can be implemented immediately, without compromising security. Corrective action is taken whenever network security is breached so that the same kind of breach is detected and any further damage is prevented in the role of response control [1].

## IV. RISK ASSESSMENT

During the initial design phase of a network, risk types and costs of recovering from attacks for all compromised resources are assessed. Risks could range from natural disaster to an attack by a hacker. There is need to design some sort of spreadsheet that lists risks versus threats as well as responses to those identified threats. According to [1], cost factors can be realized using well-established accounting procedures such as cost/benefit analysis, return on investment (ROI), and total cost of ownership (TCO).

## V.LISTING NETWORK RESOURCES

It is a necessity to identify the assets (resources) that are available on the corporate network. Table 10.1 [1] identifies mission-critical components of any enterprise network. These mission-critical components need to be prioritized because they do not all provide the same functions. According to [1], Some resources provide controlled access to a network; other resources carry sensitive corporate data. Hence the threats posed to these resources do not carry the same degree of vulnerabilities to the network. The network access control, therefore, has to be used and applied to each of the components listed, in varying degrees. For example, threats to DNS server pose a different set of problems from threats to the database servers [1].

**TABLE 10.1** Mission-Critical Components of Any Enterprise Network

| Threats / Resources | Fire, Flood, Earthquake | Power Failure | Spam | Virus | Spyware, Adware | Hijacking |
|---|---|---|---|---|---|---|
| Perimeter Router | | | | | | |
| DNS Server | | | | | | |
| WEB Server | | | | | | |
| Email Server | | | | | | |
| Core Switches | | | | | | |
| Databases | | | | | | |

## VI. THREATS

There is necessity to identify the threats posed to the network from internal users as opposed to those from external users because the internal users are easily traceable, compared to the external users. If a threat to the data network is successful, it could not only lead to loss or theft of data or denial of access to the services offered by the network but it would also lead to monetary loss for the corporation [1]. Once these threats have been identified by us, they can be ranked from the most probable to the least probable and design network security policy to reflect that ranking.

**TABLE 10.2** The Most Frequent Threats to the Network Are from Viruses

| Rank | Threat |
|---|---|
| 1 | Virus |
| 2 | Spam |
| 3 | Spyware, Adware |
| 4 | Hijacking |
| 5 | Power Failure |
| 6 | Fire, Flood, Earthquake |

**Table 10.2 The Most Frequent Threats to the Network [1].**

From Table 10.2 [1], it is observed that the most frequent threats to the network are from viruses, and we have seen a rapid explosion in antivirus, antispamware, and spyware and adware software. Hijacking of resources such as domain name services, Web services, and perimeter routers would lead to what's most famously known as Denial of Service (DoS) or Distributed Denial of Service (DDoS) [1]. Standby power supplies help keep the essential resources from crashing when there are power failures. Natural disasters such as fire, floods, or earthquakes can be most difficult to plan for; therefore, we see a tremendous growth in data protection and backup service provider businesses [1].

## VII. SECURITY POLICIES

According to [1], the fundamental goals of security policy are to allow uninterrupted access to the network resources for authenticated users and to deny access to unauthenticated users. The critical functions of a good security policy are:

- Appoint a security administrator who is conversant with users' demands and on a continual basis is prepared to accommodate the user community's needs.
- Set up a hierarchical security policy to reflect the corporate structure.
- Define ethical Internet access capabilities.
- Evolve the remote access policy.
- Provide a set of incident-handling procedures.

## VIII.THE INCIDENT-HANDLING PROCESS

The incident-handling process is the most important task of a security policy for the reason that you would not want to shut down the network in case of a network security breach [1]. The use of the network is to share the resources. So, an efficient procedure must be developed to respond to a breach. Set procedures must be developed jointly with the business operations manager and the chief information officer [1]. A modular design of the enterprise network is called for so that its segments can be shut down in an organized way, without causing panic [1].

There is need for a set of tools to monitor activities on the network. An intrusion detection and prevention system is needed. If an activity is detected and logged, response is activated. It is not merely enough to respond to an incident; the network administrator also has to activate tools to trace back to the source of this breach [1]. This is extremely important so that the network administrator can update the security procedures to ensure that this particular incident doesn't occur.

## IX. SECURE DESIGN THROUGH NETWORK ACCESS CONTROLS

A network is as secure as its weakest link in the overall design [1]. We need to identify the entry and exit points into the network in order to secure it. It is necessary to encrypt stored data so that if network security is breached, stolen data may still remain confidential unless the encryption is broken since most data networks have computational nodes to process data and storage nodes to store data. Encrypting data while it's being stored appears to be necessary to secure data [1].

Four things should follow our first network access control. Firstly, we define security policy on the perimeter router by configuring the appropriate parameters on the router. Secondly, in the line of defense is the external firewall that filters traffic based on the state of the network connection. Thirdly, following the firewall, we have the so-called demilitarized zone, or DMZ. In the zone we would place the following servers: Web, DNS, and email. Fourthly, the DMZ is placed between two firewalls, so our last line of defense is the next firewall that would inspect the traffic and possibly filter out the potential threat [1].

## X.IDS DEFINED

An intrusion detection system, or IDS, can be both software and hardware based [1]. IDSs listen to all the activities occurring on both the computer (node on a network) and the network itself. IDSs have the following variety of functions:

- Monitor and analyze user and system activities
- Verify the integrity of data files
- Audit system configuration files
- Recognize activity of patterns, reflecting known attacks
- Statistical analysis of any undefined activity pattern.

## XI. NIDS: SCOPE AND LIMITATIONS

According to [1], Network-based IDS (NIDS) sensors scan network packets at the router or host level, auditing data packets and logging any suspicious packets to a log file. Figure 10.2 [1] is an example of a NIDS. The data packets are captured by a sniffer program, which is a part of the IDS software package [1]. In promiscuous mode, the node on which the IDS software is enabled, runs. In the same mode, the NIDS node captures all the packets of data on the network as defined by the configuration script.

Here are some of the common malicious attacks on networks [1]:

- IP address spoofing
- MAC address spoofing
- ARP cache poisoning
- DNS name corruption

## XII. A PRACTICAL ILLUSTRATION OF NIDS

We show the use of Snort as an example of a NIDS. The signature files are kept in the directory signatures under the directory .doc [1]. To match defined signature against a pattern of bytes in the data packets thus identifying a potential attack, signature files are used. Files marked as rules in the rules directory are used to trigger an alarm and write to the file alert.ids [1]. With IP address 192.168.1.22, Snort is installed on a node. The security auditing software Nmap is installed on a node with IP address 192.168.1.20. Nmap software is able to generate ping sweeps, TCP SYN (half-open) scanning, TCP connect () scanning, and much more.

**UDP Attacks**

A UDP attack is generated from a node with IP address 192.168.1.20 to a node with IP address 192.168.1.22. Snort is used to detect a possible attack. Snort's detect engine uses one of the files in DOS under directory rules to generate the alert file alert.ids [1].

## XIII.FIREWALLS

A firewall is either a single node or a set of nodes that enforces an access policy between two networks. Firewall technology developed to protect the intranet from unauthorized users on the Internet. This was the experience in the earlier years of corporate networks [1]. It was later learnt by the network administrators that the network can also be attacked from trusted users as well as such users as the employee of a company. Because the corporate network consists of hundreds of nodes per department which thus aggregates to over a thousand or more, there is a need to protect data in each department from other departments.

A firewall is a combination of hardware and software technology, namely a sort of sentry, waiting at the points of entry and exit to look out for an unauthorized data packet trying to gain access to the network [1].

### *A. Firewall Security Policy*

The firewall empowers the network administrator to centralize access control to the campus wide network. A firewall adds an entry of every packet that enters and leaves the network. The network security policy done in the firewall provides many kinds of protection, including the following:

- Block unwanted traffic
- Direct incoming traffic to more trustworthy internal nodes
- Hide vulnerable nodes that cannot easily be secured from external threats
- Log traffic to and from the network

Whereas a firewall is transparent to authorized users (both internal and external), it is not transparent to unauthorized users. However, when the authorized user tries to access a service that is not granted to that user, a denial of that service will be echoed, and that attempt will be logged [1].

### *B. Types of Firewalls*

In [1], conceptually, there are three types of firewalls:

- *Packet filtering.* Permit packets to enter or leave the network through the interface on the router on the basis of protocol, IP address, and port numbers.
- *Application-layer firewall.* A proxy server that acts as an intermediate host between the source and the destination nodes.
- *Stateful-inspection layer.* Validates the packet on the basis of its content.

## XIV. NIDS COMPLEMENTS FIREWALLS

If so designed, a firewall acts as a barrier among various IP network segments. To protect resources, firewalls may be defined among IP intranet segments. There will always be more than one firewall in any corporate network because an intruder could be one of the authorized network users. The firewall can only monitor the traffic entering and leaving the interface on the firewall that connects to the network since the firewall sits at the boundary of the IP network segments. If the intruder is internal to the firewall, the firewall will not be able to detect the security breach [1].

The intruder would go undetected once an intruder has managed to transit through the interface of the firewall. This could possibly lead to stealing sensitive information, destroying information, leaving behind viruses, staging attacks on other networks, and most important, leaving spyware software to monitor the activities on the network for future attacks. Hence, an NIDS would play a critical role in monitoring activities on the network and continually looking for possible anomalous patterns of activities [1].

## XV. SIGNATURE ALGORITHMS

In [1] Signature analysis is based on these algorithms:

- Pattern matching
- Stateful pattern matching
- Protocol decode-based analysis
- Heuristic-based analysis
- Anomaly-based analysis

### *A. Pattern Matching*

Pattern matching has to do with searching for a fixed sequence of bytes in a single packet. If the suspect packet is associated with a particular service or, more precisely, destined to and from a particular port, then the pattern is matched against in most cases. This helps to bring down the number of packets that must get checked and thus accelerates the process of detection. However, it tends to make it more difficult for systems to deal with protocols that do not live on well-defined ports [1].

### *B. Stateful Pattern Matching*

This method is an addition to the pattern matching concept. It is because a network stream comprises more than a single atomic packet. In context within the state of the stream should matches be made. This means that systems that perform this type of signature analysis must consider arrival order of packets in a TCP stream and should handle matching patterns across packet boundaries [1]. This is quite similar to a stateful firewall.

Now, instead of looking for the pattern in every packet, the system has to begin to maintain state information on the TCP stream being monitored [1]. Consider the following scenario to understand the difference. Imagine that the attack you are looking for is launched from a client connecting to a server and you have the pattern-match method deployed on the IDS. If the attack is launched so that in any given single TCP packet bound for the target on port 3333 the string is present, this event triggers the alarm. If, however, the attacker sets off the offending string to be sent such that the fictitious *gp* is in the first packet sent to the server and *o* is in the second, the alarm doesn't get activated. If the stateful pattern-matching algorithm is deployed instead, the sensor has stored the *gp* portion of the string and is able to complete the match when the client forwards the fictitious *p* [1].

### *C. Protocol Decode-based Analysis*

Intelligent extensions to stateful pattern matches are protocol decode-based signatures in many ways. Implementation of this class of signature is done by decoding the various elements in the same manner as the client or server in the conversation would. To find violations, when the elements of the protocol are identified, the IDS applies rules defined by the request for comments (RFCs). In some instances, these violations are found with pattern matches within a specific protocol field, and some require more advanced techniques that account for such variables as the length of a field or the number of arguments [1].

In [1], the advantages of the protocol decode-based analysis are as follows:

- ❖ This method can allow for direct correlation of an exploit.
- ❖ This method can be broader and general to allow catching variations on a theme.
- ❖ This method reduces the chance for false positives if the protocol is well defined and enforced.
- ❖ This method reliably alerts on the violation of the protocol rules as defined in the rules script.

In [1], the disadvantages of this technique are as follows:

- ❖ This method can lead to high false-positive rates if the RFC is ambiguous and allows developers the discretion to interpret and implement as they see fit. These gray area protocol violations are very common.
- ❖ This method requires longer development times to properly implement the protocol parser.

### *D. Heuristic-Based Analysis*

A signature that would be used to detect a port sweep is a good example of this type of signature. This signature searches for the presence of a threshold number of unique ports being touched on a specific machine. The signature may further restrict itself through the specification of the types of packets that it is interested in (that is, SYN packets) [1].

In addition, it may be required that all the probes must originate from a single source. Such types of signatures need some threshold manipulations to make them conform to the utilization patterns on the network they are monitoring. This type of signature may be used to look for very complex relationships as well as the simple statistical example given [1].

The advantage for heuristic-based signature analysis is that some types of suspicious and/or malicious activity cannot be detected through any other means. The disadvantage is that algorithms may require tuning or modification to better conform to network traffic and limit false positives [1].

*E. Anomaly-Based Analysis*

Anomaly-based signatures are normally prepared to look for network traffic that deviates from normal observation. To first define what normal is the biggest problem with this methodology. Some systems could be considered heuristic-based systems since they have hardcoded definitions of normal. Others are built to learn normal. The challenge with such systems is in eliminating the possibility of improperly classifying abnormal behavior as normal. Also, if the traffic pattern being learned is assumed to be normal, the system must contend with how to differentiate between allowable deviations and those not allowed or representing attack-based traffic [1].

Although there are a few commercial products that claim to use anomaly-based detection methods, the work in this area has been mostly limited to academia. The profile-based detection methods are a subcategory of this type of detection. On changes in the way that users or systems interact on the network do these systems base their alerts. They incur many of the same limitations and problems that the overarching category has in inferring the intent of the change in behavior [1].

According to [1], the advantages for anomaly-based detection are as follows:

- ✓ If this method is implemented properly, it can detect unknown attacks.
- ✓ This method offers low overhead because new signatures do not have to be developed.

In [1], the disadvantages are:

- ✓ In general, these systems are not able to give you intrusion data with any granularity. It looks like something terrible may have happened, but the systems cannot say definitively.
- ✓ This method is highly dependent on the environment in which the systems learn what normal is.

The following are Freeware tools to monitor and analyze network activities:

- Network Scanner, Nmap, is available from www.insecure.org.

Being a free open-source utility, Nmap is used to monitor open ports on a network. A zip file by the name of nmap-3.75-win32.zip is the MS-Windows version. You are also required to download a packet capture library, WinPcap, under Windows. You can access it from http://winpcap.polito.it. In addition to these programs, you need a utility to unzip the zipped file, which you can download from various Internet sites [1].

- Use PortPeeker, a freeware utility, to capture network traffic for TCP, UDP, or ICMP protocols. You can easily and quickly see what traffic is being sent to a given port by using PortPeeker. According to [1], this utility is available from www.Linklogger.com
- Fport 2.0 and SuperScan 4.0 are Port-scanning tools that are easy to use and freely available from www.Foundstone.com.
- You can download Network Sniffer Ethereal from www.ethereal.com. Ethereal is a packet sniffer and analyzer for a variety of protocols [1].
- To capture and analyze the packets going through the network, a free network sniffer called EtherSnoop light was designed. It not only captures the data passing through your network Ethernet card but it also analyzes the data and represents it in a readable form. EtherSnoop light is a fully configurable network analyzer program for Win32 environments [1]. It is available from www.arechisoft.com.
- Snort is a fairly advanced tool, an open-source NIDS and available from www.snort.org.
- Available from www.Foundstone.com is a stress testing tool called UDPFlood that could be identified as a DoS agent.
- An application that allows you to generate a SYN attack with a spoofed address so that the remote host's CPU cycle's get tied up is called Attacker, and is available from www.komodia.com [1].

## XVI. LOCAL AREA NETWORK SECURITY REVIEW

J.R. Beulah and D.S. Punithavathani [2] in discussing Network Intrusion Detection in their paper state that Network intrusion detection is a dynamic research area as intruders or attackers have increased attacks on all kinds of networking set-ups. Outlier mining or detection has gained much attention of researchers in the field of data mining. For this reason, it finds application in many kinds of uncommon event detections like network intrusion detection and fraud detection. Several techniques for outlier detection are available in the literature but only a few have been applied for detecting network intrusions [2].

Their survey gives an introduction to intrusion detection and outlier detection. It also presents a generalized framework for outlier based intrusion detection systems which outlines the steps involved in processing the network traffic data to find possible intrusions and reports a listing of existing intrusion detection systems using outlier detection methods with a detailed tabulation categorized by the approach used. Their paper further discusses the commonly used performance measures used for evaluating intrusion detection systems, presents the inferences of the survey and outlines the issues and challenges which may help future researchers for designing new intrusion detection methods [2].

Based on the revelation of promising results given from all the methods analysed in their paper, outlier detection is well-suited for network intrusion detection. Researchers could concentrate on designing hybrid approaches that combine different outlier detection methods that may yield better performance. According to their analysis, outlier detection methods have verified their being effective in intrusion datasets but not in real time network traffic. Further research is necessary to test such methods on real network data [2].

According to M.N. Omer, A. Amphawan and R. Din [3], current computing trends such as cloud computing, file sharing and social networking promote collaboration and allow greater

mobility for users. However, it is also true that these computing trends increase the networks' being prone to security threats and challenge network resources. An ingenious technique employed by attackers for retaining anonymity is by exploiting intermediary host computers or stepping stones to instigate attacks on other computers [3].

M.N. Omer, A. Amphawan and R. Din [3], in their paper, explore novel application of the Stepping Stone Detection (SSD) concept in addressing network threats such as spams, backdoors, proxy server intrusions and denial of service attacks. The potential detection process is marked out when preliminary Stepping Stone Detection (SSD) models for each security threat are constructed. These preliminary concepts and models may prove useful for further optimization of network security in conjunction with other conventional detection techniques [3].

For the detection of series of host computers by attackers, SSD has untapped potential in many rising research fields, for instance, in spam, backdoor, proxy and Denial of Service (DoS) attack detection. To demonstrate the potential of SSD in addressing current issues in spam, backdoor and proxy detection, four novel SSD models are presented. They suggest undertaking, for future work, extensive SSD simulations and verification on real data such as wireless network, and mobile wireless network using Computational Intelligence (CI) methods for each emerging domain [3].

J.C. Alfaro and H. Debar [4] in their paper state that in a networking context, Access Control Lists (ACLs) refer to security rules associated to network equipment, such as routers, switches and firewalls. To verify if the corresponding ACLs are functionally equivalent, methods and tools to automate the management of ACLs distributed among several equipment should be employed. In [4], they address such a verification process. They present a formal method to verify when two ACLs have the same function and show their proposal over a practical example.

According to [4], the management of network Access Control Lists (ACLs) refers to the task of authoring and maintaining security rules associated to network equipment, such as routers, switches and firewalls. Methods and tools devised to assist during those processes validate if the ACLs distributed among different equipment are functionally equivalent. In concluding their paper, they suggest future works that aim to extend the concept of isofunctionality to general-purpose access control models [4].

Security policy is a main mechanism of information security management [5]. In spite of many security-related standards and guidelines which assign requirements for high-level security policies, implementation of network security policy still relies on interfaces provided by Network Security Solutions (NSS). Today there are numerous Network Security Solutions (NSSs) developed by different manufactures and all of them use different mechanisms for specification and implementation of IS policies [5].

The market Unified Threat Management (UTM) solutions combine a variety of different security functions and technologies that allow to embed various security modules into one physical appliance. As a result, concentration of all possible security functions in one solution becomes the main trend and borders between firewalls, IDS/IPS, DLP systems and other NSSs are blurring. Nevertheless, D. Chernyavskiy and N. Miloslavskaya have shown us that NSSs of different types and manufacturers are able to implement policies that are semantically equal [5].

Increase in efficiency by means of reducing useless diversity and redundancy can be brought about by usually using a type of standardization called Unification. In the model presented there is a demonstration of an effort to unify low-level policies for NSSs and simplify mechanisms of their implementation. It is possible to eliminate redundancy of policies for NSSs and, as a result, increase efficiency of policies development process by partitioning the set of NSSs to equivalence classes and applying a single policy to each class. By using a translator, unified policies become portable between different NSSs thereby simplifying processes of their implementation [5].

Finally, to add any new NSS to the model, split it to simple NSSs and assign them to some equivalence class. The new NSS forms its own class if there is no NSS in the classification to which the new one is equivalent. As a consequence, unified policy is to be created for this new class using the same principles as other unified policies (by the addition of new elements to an algebra or new grammar rules to the language). According to [5], the future challenge for the model is a development of a comprehensive classification of existing NSSs in terms of described equivalence, generation of unified policies for the equivalence classes and progress in construction of policy translation methods.

In their paper entitled, "An Operational Framework for Incident Handling," G. Bottazi and G.G. Rutigliano propose an operational framework for incident handling [6]. Their proposal intends to shift attention from the incident response stage to preparing first the rules, regulations, techniques, procedures, skills and tools to avoid having to approach the issue only after an incident. Increase in the time of the attack and decrease in the time of response thus bringing balance on the field can come as a result of the work done, or at least organized before. This can also be used as a "gym" for forming the awareness, developing the know-how and raising up the motivation of people involved in cyber defence [6].

The framework proposed in their conference paper requires efforts and investments at various levels that may appear excessive in normal conditions (peacetime), but assumes importance if scrutinized with the recovery difficulties of a cyber-incident, managed with general-purpose plans, with inevitable impacts on the exponential growth of time and costs [6].

According to S. Behal, M. Sachdeva and R. Mehto [7], today, the internet is the primary medium for communication which is used by a number of users across the Network. A denial-of-service (DoS) is one of the major security problems in the current Internet. It always makes an effort to stop the victim from serving legalized users. A Distributed Denial of Service (DDoS) attack is a DoS attack that uses many distributed attack sources. The network and transport layers are the targets of the majority of DDoS attacks.

In [7], they have measured the performance of Web services under DDoS attack using Real time test bed (GENI). GENI stands for Global Environment for Network Innovations. In their work, a GENI test bed was explored and a topology was created on which HTTP legitimate traffic and UDP attack

traffic was produced. One of the simplest Transport Layer communication protocol available of the TCP/IP protocol suite is the User Datagram Protocol (UDP). It takes into account a minimum amount of communication mechanism. Avg.Response Time, Avg.Round Trip Time (RTT) and Throughput in terms of good-put and bad-put is computed to measure impact of DDoS attacks on Web HTTP services [7].

F. Agbenyegah and M. Asante [8] observe that the requirement for firewalls has prompted their omnipresence. A firewall has been introduced by almost every organization connected with the Internet. This results in most organizations having some level of assurance against external dangers. Their study has discovered that network security and network performance are inversely proportional. Through distinctive scenarios, the connection between the security and execution proficiency is shown and the relationship between security and performance in firewalls is gauged. The simulation was done for 300 work stations and simulated in a such way that all the 300 work stations accessed an email and web application under three different scenarios. Different scenarios were assessed through simulations utilizing OPNET IT Guru Academic Edition 9.1 to demonstrate the impacts of firewalls on system performance [8].

The result of the simulation showed that network performance is adversely affected when firewall is implemented. When security policies of the organization are implemented, the performance suffers from degradation. Nevertheless, firewalls [8] don't just secure a system, but they additionally add to system performance by ceasing assaults, enhancing system accessibility and lessening superfluous preparing of illegitimate solicitations.

S. Manaseer and A.K. Hwaitat [9] in their paper propose a centralized web firewall system for web application security which will provide a new type of synchronized system, which has the ability to detect and prevent a variety of web application attacks for a wide range of hosts. At the same time, using a centralized command and control system, the attacked client then sends the information to a centralized command and control server which will distribute the attack information to all of the integrated clients connected to it [9].

All of the attack information including the type of attack, the IP address of the attacker, and the time of attack are contained in the distributed information. The question is how is the process of receiving the attacker's information and distributing it through the centralized web firewall done? It is done automatically and immediately at the time of the attack. Additionally, all of the receiving clients will act against the threat depending on the distributed information such as banning the IP address of the attacker. The main process [9] aims to protect multiple clients from any possible attack from the same attacker or the same type of attack.

To protect a real web application, this system has been implemented. Experiments demonstrated that the attacks were successfully prevented on many hosts at the time. The paper [9] provides a centralized web firewall system that connect different web firewalls in order to detect and prevent different types of web attacks and work as a fully integrated system with the different clients.

According to M. Turcanik [10], a very important problem of neural networks in real application is their structural complexity. The neural network structure gives you the structural complexity. As a result, optimisation of the neural network structure from point of actual application is very important and a problem that is deemed necessary. Because model and structure of neural network depends on real application, the analysis of neural network can be done from point of view of application criteria [10].

In complementing the existing approaches, the author proposes an artificial neural network for packet filtering. A new approach for packet filtering using neural networks was presented in his paper. To build a packet filter which could be trained on the base of set of rules was his main purpose. The core of the system was displayed. However, a complete system would need some additional modules for using as a real application. Artificial neural network topology and optimization criterions were designed to solve a problem of controlling of communication in the computer network [10].

Its speed is the main advantage of using ANN for packet filtering. The speed of a decision whether to allow or deny the packet could not be influenced in this approach by the size of Access Control Lists (ACL). To get a decision from ANN is employed always the same time. ANN could be learned to change the rules on what could be time consuming process [10].

A. Czubak and M. Szymanek [11] write that algorithmic complexity vulnerabilities are an opportunity for an adversary to conduct a sophisticated kind of attack i.e. on network infrastructure services. In their paper they describe the algorithmic complexity attacks as a possible vector which can be used for inflicting a DoS attack on computer systems. Their focus is on the hash table data structure. In [11], they conduct and describe in detail a successful DoS attack, performed on an industry-grade router using stateful firewall functionality provided by Cisco® Context-Based Access Control (CBAC).

They discussed an attack scenario, all necessary scripts, commands and tools to carry out such an attack and potentially replicate their results. The attack required as little bandwidth as 24kbit/s to execute. Their experiments demonstrated we can cause extremely bad network outage of a couple of minutes or even router’s crash and reboot, without leaving any trace whatsoever of the attack ever occurring. An industry grade equipment is vulnerable to algorithmic complexity attacks just as open source solutions are [11].

Conducted research proved that algorithmic complexity attacks are a serious threat even with corporate equipment, such as Cisco® routers. An exemplary device’s behavior was analyzed. They believe that in the presented case it is possible to use more uniformly distributed hash function and a different key to lower the threat or impact of complexity attack without any significant increase in computational complexity of firewalls session table algorithm [11].

They also believe that it is not only possible but it is also, in some cases, mandatory to use algorithms with slightly higher computational complexity which are invulnerable to algorithmic complexity attacks without crucially losing their performance. Unless properly secured or used for a very narrow spectrum of controlled input, software prone to such attacks should not be used for firewalls, IDS and IPS systems. As a follow-up they intend to develop a modified hash table data structure, customized for storing session objects representing

connections in firewalls, that would be invulnerable to algorithmic complexity attacks [11].

According to M.K. Chinyemaba [12], her study aimed at assessing the security GAPs using ISO 27001:2013 Information Security Management System (ISMS) standard. Quantitative and qualitative was the study approach used with survey questionnaires and interviews as assessment tools for empirical data collection. The study showed that Zambian public sector has related challenges in mitigation of insider attacks that calls for concerted efforts in developing measures for mitigation of these challenges in order to ensure national cyber security readiness and enhancing data privacy [12].

In the research it was evident that majority of the organizations assessed lack insider security deterring policies such as access control, non-disclosure agreements (NDA), pre-employment screening and unacceptable use. In addition, the findings indicated that the larger part of public organizations have made no efforts towards cyber security readiness, while only about 33% have adopted some security base practices. Further, using Actor Network Theory (ANT) and Theory of Planned Behavior (TPB), the study proposed an expedient insider mitigation model with an emphasis on user awareness and access control considering that it is difficult to model human behavior [12].

M. Nyirenda et al [13] observe that Web applications are usually installed on and accessed through a Web server. In general, Web servers provide very few privileges to Web applications, defaulting to executing them in the realm of a guest account because of security reasons. Additionally, performance is often a problem as Web applications may need to be initialized again with each access. Various solutions have been designed to address these security and performance issues, mostly independently of one another, but most have been language or system-specific [13].

The X-Switch system, as an alternative Web application execution environment, with more secure user-based resource management, persistent application interpreters and support for arbitrary languages/interpreters, is proposed. For this reason, it provides a general-purpose atmosphere for developing and deploying Web applications. With the X-Switch system, you can achieve a high level of performance as demonstrated by experiments carried out. Furthermore, it is evident that the X-Switch does not seem only to be able to provide functionality matching that of existing Web application servers but it also seems to be able to add the benefit of multi-user support. Finally, the X-Switch system showed that it is feasible to completely separate the deployment platform from the application code, thus ensuring that the developer does not need to modify his/her code to make it compatible with the deployment platform [13].

## XVII. CONCLUSION

Over the past 15 years, ecommerce-related activities such as online shopping, banking, stock trading, and social networking have permeated broadly, creating a dilemma for both service providers and their potential clients, as to who is a trusted service provider and a trusted client on the network. Securing available resources on any corporate or academic data network is of great importance because most of these networks connect to the Internet for commercial or research activities. The security policy must be a factor in clients' level of access to the resources. So, in whom do we place trust, and how much trust? The function of the detection control is to monitor the activities on the network and identify an event or a couple of events that could breach network security. The function of prevention control is to stop unauthorized access to any resource available on the network.

In future the authors intend to undertake a review of the challenges that affect System Security and how they can be dealt with.

## REFERENCES

[1] J. R. Vacca, "Computer and Information Security Handbook," Burlington, USA: Morgan Kaufmann Publishers, 2009, pp. 149-166.

[2] J.R. Beulah and D.S. Punithavathani, "Outlier detection methods for identifying network intrusions – A survey," pp. 1-10, Jan. 2015.

[3] M.N. Omer, A. Amphawan and R. Din, "A Stepping Stone Perspective to Detection of Network Threats," pp. 1-11, Dec. 2013.

[4] J.C. Alfaro and H. Debar, "On the Isofunctionality of Network Access Control Lists," pp. 1-7, Aug. 2015.

[5] D. Chernyavskiy and N. Miloslavskaya, "A Concept of Unification of Network Security Policies," pp. 1-6, Dec. 2013.

[6] G. Bottazi and G.G. Rutigliano, "An Operational Framework for Incident Handling," pp. 1-11, Mar. 2017.

[7] S. Behal, M. Sachdeva and R. Mehto, "Performance Measurement of Web Services under UDP Attack using GENI Testbed," pp. 1-10, Jan. 2018.

[8] F. Agbenyegah and M. Asante, "Impact of Firewall On Network Performance," pp. 1-8, Jul. 2018.

[9] S. Manaseer and A.K. Hwaitat, "Centralized Web Application Firewall Security System," pp. 1- 7, May 2018.

[10] M. Turcanik, "Packet filtering by artificial neural network," pp. 1-5, Jul. 2015.

[11] A. Czubak and M. Szymanek, "Algorithmic Complexity Vulnerability Analysis of a Stateful Firewall – Extended," pp. 1-25, Sep. 2016.

[12] M.K. Chinyemaba and J. Phiri, "An Investigation into Information Security Threats from Insiders and how to Mitigate them: A Case Study of Zambian Public Sector,'' pp. 1-13, Nov. 2018.

[13] M. Nyirenda, H. Suleman, A. Maunder and R. V. Rooyen, "X-Switch: An Efficient, Multi-User, Multi-Language Web Application Server," pp. 1-10, Dec. 2010.

# Assessing the Readiness of Students to use Mobile Applications in Collaborative Learning: Looking for Answers with UTAUT

Phillimon Mumba,
Copperbelt University, Zambia
phillimonmumba@gmail.com

Maybin Lengwe,
Copperbelt University, Zambia
maybin.lengwe@cbu.ac.zm

**ABSTRACT- To improve student performance and retention rates, higher institutions of learning are constantly researching on the approaches, tools and techniques to use. In recent times, concepts such as mobile learning, electronic learning, collaborative learning, flipped classroom and deep learning have emerged. These describe the different approaches that institutions are using to improve student performance and retention rates. However, the successful implementation of an approach largely depends on the willingness of the users (learners and educators) to use. In this paper we are using the Unified Theory of Acceptance and Use of Technology (UTAUT) to determine the willingness of students at Copperbelt University to use mobile application-aided collaborative learning in their studies.**

**Key Words – Collaborative Learning, Collaboration, Cooperative Learning, mobile learning (m-learning), Computer –aided learning, electronic learning (e-learning) INTRODUCTION**

Educators have researched on various tools and approaches that can be used to improve student performance and retention rates in tertiary education. These approaches and technologies have different emphasis, some of them emphasise on active learning, improvement of collaboration among learners and educators, mobile learning and adaptive learning, among many others [1, 2].

The New Media Consortium publishes the approaches and technologies that are likely to affect education in tertiary institution over a period of five years. In their 2017 report, they suggest that computer-aided collaborative learning is one of the education technologies that will be adopted in tertiary education in the next one to two years [1]. Collaborative learning or cooperative learning refers to a learning approach in which learners work in groups in a coordinated fashion to achieve a learning goal or complete a learning task [1, 3]. Collaborative learning aims to encourage student engagement and critical thinking in learning [4].

Computer-aided collaborative learning occurs when technology is introduced in collaborative learning environments [5, 6]. Odiakaosa, et al, [7] and Ledlow [8] identifies that the successful implementation of technology in education requires the involvement of learners and educators. Failure to involve these key stakeholders results in underutilisation of the technology or failure to achieve the intended goals for introducing the technology [9].

The authors of this article believe that it is important to assess the readiness of students to use mobile applications in collaborative learning to ensure that institutions can successfully incorporate this learning model in their institutions. The key objectives the authors are working to achieve are:

- To identify the key issues in order to successfully use mobile application-aided collaborative learning.
- To determine the key benefits of collaborative learning.
- Assess the willing of students to use mobile applications in collaborative learning.

The rest of the paper is structured as follows: Section II reviews literature on similar work, section III describes the methodology used in the research, Section IV gives the findings and analysis of the findings and section V gives the key implications of the research findings and section VI concludes the paper.

## I. LITERATURE REVIEW

As already defined, collaborative learning is a group-based learning approach in which learners work together to accomplish a specific goal [1]. Becker et al. [1] identifies placing the students at the centre of learning process, interaction and working in groups to develop solutions as the key principles on which collaborative learning is based. Chen and Denoyelles [10], identifies that collaborative learning activities typically occur outside the classroom where there is

very little guidance from the educator or instructor. This means that, instead of looking at students only as recipients of knowledge, they are considered as the originator of the knowledge. We will now consider several issues relating to mobile application-aided collaborative learning.

- **Mobile applications in education**

Researches done on the use of mobile applications in education have identified that mobile applications provide opportunities for educators to promote learning [10, 4]. The key characteristics that have made mobile applications very popular in education and other fields include ubiquity, privacy, interactivity, portability, immediacy and collaboration [11]. Ubiquity means the application can be accessed from anywhere. For this reason, mobile applications have been identified as key in collaboration [11]. The typical uses of mobile applications in education include collaboration, engagement and interactivity. Students are able to share lecture materials, discuss class activities and comment on class programs [12, 13]. Al-Hunaiyyan, et al. [14], identifies that mobile application and mobile learning offers us an opportunity to change the current learning strategies to those that are more flexible. Because the benefits offered by mobile application aided learning, many academic institutions are now incorporating mobile learning in their learning strategies.

Despite the many benefits that mobile application aided learning promises to offer, other researchers have cautioned against rushing to introduce mobile applications in education. Sung, et al. [15], argues that, for education software (mobile applications included) to be used most effectively, the educators should design the learning experience that matches the software with the learning objectives of the educational experience. In Sung. Et al's view, only those applications that are aligned with the leaning objectives of the educational experience will succeed. Al-Hunaiyyan, et al. [14] argues that there are five categories of mobile learning challenges. These are social and cultural challenges, management challenges, technology challenges, design challenges and evaluation challenges. Social and cultural challenges refer to how students in a particular social and cultural setup views the use of mobile learning. Management challenges refer challenges to deal with how mobile learning will be administered. Design challenges refer to how the education content for mobile learning should be designed. Mobile devices have limited resources, therefore one should pay attention to how the academic materials for mobile learning are designed to avoid depleting the resources of the device. Evaluation challenges are those challenges that arise when determining whether mobile learning was successfully carried out or not. Technology challenges refer to the challenges in the technologies to be used in mobile learning.

- **Benefits of collaborative learning**

A number of researches have been carried out to investigate the benefits of using collaborative learning in higher education. University of Arizona [16] argues that students' performance improves when active learning strategies as collaborative learning are used. University of Tennessee [17] argues that collaborative learning motivates the students to learn the material. This argument is substantiated by the fact that in collaborative learning students play the roles of being the educators to their peers [3].

Research also identifies offering of formative feedback, developing of social and group skills, better self-esteem, increased leadership and developing of analytical skills as other benefits of collaborative learning [1, 3, 17]. Collaborative learning also promotes positive interaction between members from different cultural and social backgrounds [17]. Konak and Bartolacci [18], concludes that, students that participate in collaborative learning tend to perform better than students who work in isolation.

- **Issues affecting collaborative learning**

To successfully implement computer-aided collaborative learning, there are some key issues that need to be resolved. Some of these issues include communication challenges and individual accountability [9, 19]. This is especially true in developing countries where there are still challenges with internet connectivity. To successfully conduct collaborative learning using computing devices, the participants should be able to have in-depth discussion and participate as openly as possible [20]. This would be hampered if the communication channels are limited, ineffective or very expensive to acquire.

The issue of individual accountability is very important to build trust in online collaborative learning [19]. Accountability entails that each participant in the collaboration must play their role in the group work. However, if, some participants are not faithful in their roles, the other participants tend to pull out of the collaboration.

Other researchers have pointed out that mobile application aided collaborative learning have several negative effects. For example, Heflin et al. [4], argues the use mobile application does not always encourage critical thinking. In some cases, the tools used may distract the student and their colleagues. This is true since mobile devices users are able to multitask. The user will therefore be switching between different applications.

- **Why mobile application-aided collaborative learning**

With the proliferation of mobile devices among learners in higher institutions of learning, educators are researching on how they can be used to support collaborative learning activities. Cheong [3] argues that, "there is little room for collaborative learning in the short time frame of a lecture." This makes it imperative that mobile devices-based collaborative learning a necessity in education. Ku et al [21] also argues that students favour collaboration in online or electronic learning environments.

There is a number of existing application software that can be used for aiding collaborative learning. These applications include wikis, google docs, social media and messaging apps [1]. These applications have been used in learning with varying degrees of success [22, 23, 24]. Deal [25] argues that educators can benefit from a more detailed and disaggregated study of available collaboration tools, and how to effectively use them. Morrison [26] outlines five steps that should be considered when introducing online collaboration tools. The first step is to determine whether a particular tool will motivate learners. The second step is to review the learning objective of the activity to be supported by the tool. The third step is to identify the contents that students need to learn. The fourth step is assess whether the tool will encourage learners to apply the content and learning the material, construct knowledge and promote critical thinking. The final step is to select and implement the best application.

## II. METHODOLOGY

To effectively investigate the students' willingness to use mobile applications-aided collaborative learning, we consider the problem from the perspective of technology acceptance. This section covers the theoretical framework used,research design, population sample, data collection instrument and method of data analysis.

- **Theoretical Framework**

In this research, we have used the Unified Theory of Acceptance and Use of Technology (UTAUT) framework. This framework was developed by Venkatesh et al. in 2003. This model identifies four direct determinants of behavioural intention and use behaviour. These determinants are performance expectancy (PE), effort expectancy (EE), social influence (SI) and facilitating conditions (FC). Performance expectancy refers to the expected functionality of the system. The students would be willing to use a technology if they can identify the benefits (functionalities) that the technology offers them. Effort expectancy refers to the amount of effort required in order for the user to get the expected results. Students would use a technology if it will not require too much effort from them. Social influence states that students would use a technology if influential people in their lives (for example, fellow students, lecturers, et. cetera) are using the technology. Facilitating conditions refers to the facilities required to successfully use a technology [27]. The conditions required to use mobile application-aided collaborative learning include internet availability and owning a smartphone. These four variables (PE, EE, SI and FC) directly affect intention to use a technology. The variables: age, gender, experience and being voluntary affect the usage of a technology. UTAUT variables were adopted as measures to evaluate the students' willingness to use mobile application-aided collaborative learning. Figure 1 shows the UTAUT framework.

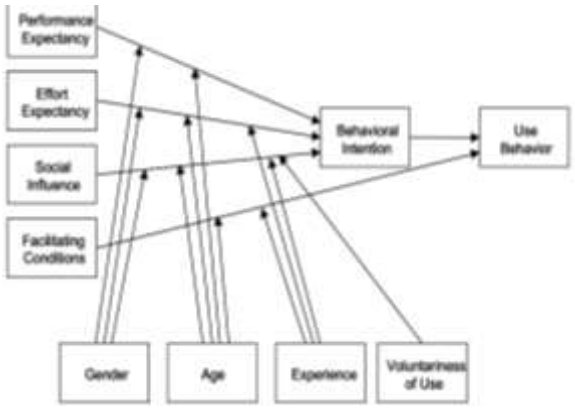



**Figure 1 UTAUT Framework [27]**

- **Design of the study**

A research design is a plan used by a scholar to obtain research participants and to collect information. The research design of this study is

exploratory in nature. An exploratory research is carried out when earlier studies to refer to are limited [7]. This design is useful for this study because it explores students' readiness to incorporate mobile apps in collaborative learning at Copperbelt University.

- **Participants**

Copperbelt University has over 10,000 students. Carrying out a study of the whole population was not realistic; a sample of the population was therefore taken. 709 students drawn from seven (7) schools and one (1) directorate were studied. These students' years of study ranged from first year to final year (that is, fourth or fifth year) of their undergraduate studies. Table 1 give a summary of the participants.

**Table 1 Cross tabulation of the respondents year of study and their gender**

| | | Gender of the respondent | | Total |
|---|---|---|---|---|
| | | Female | Male | |
| Respondent's year of Study | First Year | 112 | 167 | 279 |
| | Second Year | 59 | 197 | 256 |
| | Third Year | 24 | 73 | 97 |
| | 4th or 5th Year | 29 | 48 | 77 |
| | TOTAL | 224 | 485 | 709 |

- **data collection instrument and analysis**

A survey questionnaire was used in collecting the data for the research. There were different questions related to the research being conducted. Using the questionnaire, we were able to collect students' opinion of collaborative learning, their confidence in it and their willingness to use a collaborative learning mobile app, if it was developed.

The data that was obtained in the research was analysed using frequency analysis to get the general consensus of CBU students' views of mobile application-aided collaborative learning.

## III. RESULTS

900 questionnaires were distributed and 709 Students from seven (7) schools and from the Directorate of Distance Education and Open Learning responded. The results below show the responses from the respondents.

- **Amount of time spent online**

Figure 2 shows the amount of time students spend online. About 44% of students spend up to one hour of their day online while 56% of the students spent one or more hours online.

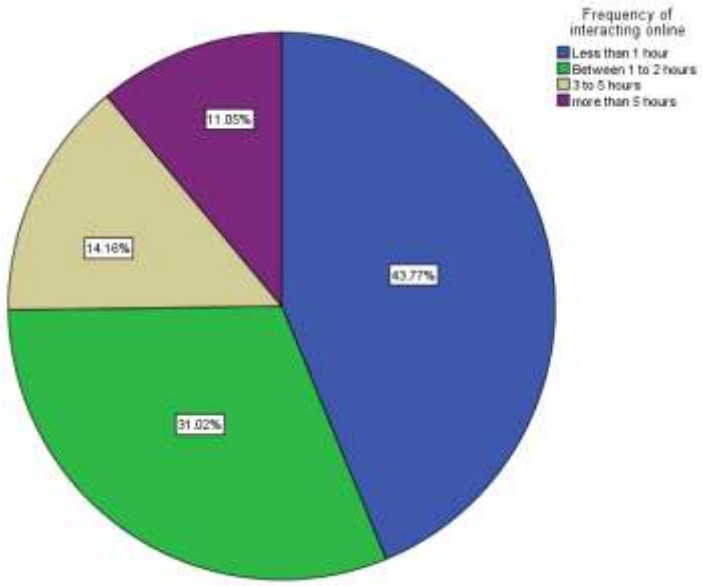



**Figure 2 amount of time spent online**

- **User confidence in collaborative learning**

Tables 2 and figure 2 show the confidence that students in different years of studies have in collaborative learning activities. As it can be seen from the table and figure, most undergraduate students have confidence in collaborative learning. In addition, we can also observe that the confidence in collaborative learning activities increases as students progress in their studies. Table 3 shows the confidence that students in different age groups have in collaborative learning. We can observe that students in all age groups have confidence in collaborative learning.

**Table 2Confidence in collaborative learning activities**

| | | Confidence in group academic activities | | Total |
|---|---|---|---|---|
| | | Yes | No | |
| Respondent's year of Study | First Year | 236 | 43 | 279 |
| | Second Year | 217 | 37 | 254 |
| | Third Year | 85 | 11 | 96 |
| | 4th or 5th Year | 69 | 8 | 77 |
| Total | | 607 | 99 | 706 |

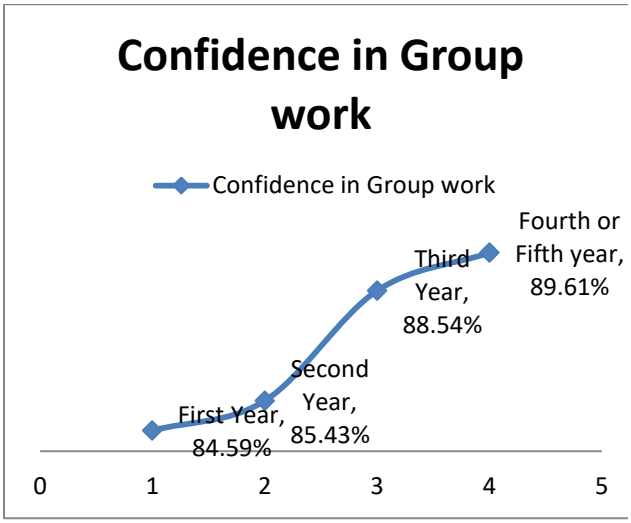



**Figure 3 Confidence in collaborative learning activities**

**Table 3 Age of the respondent * Confidence in group academic activities Crosstabulation**

| Age of the ondent | Confidence in group academic activities | | Total |
|---|---|---|---|
| | Yes | No | |
| Below 25 | 519 | 90 | 609 |
| 25-29 | 37 | 6 | 43 |
| 30-35 | 31 | 1 | 32 |
| Above 35 | 19 | 2 | 21 |
| Total | 606 | 99 | 705 |

- **Ways of conducting collaborative learning activities**

It is very clear that much of collaborative learning done face to face. That is, all the participants are in the same location (see figure 4). The reason why this is the most preferred way of collaboration is that the participants in the collaboration respond immediately a concern is raised.

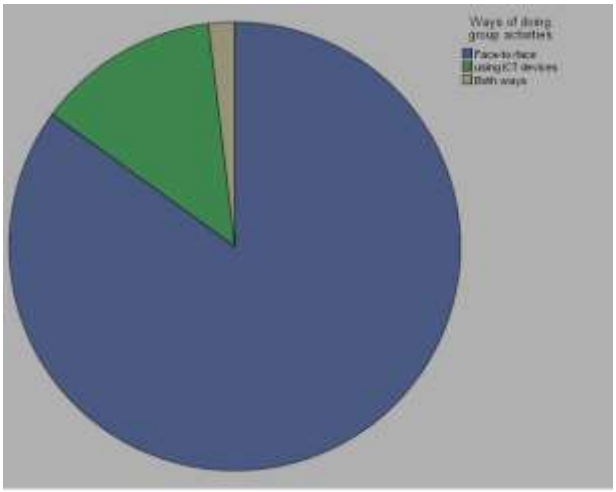


**Figure 4 Ways in which collaborative learning is currently being done**

- **Uses of mobile applications in education**

Figure 5 shows the academic activities that are currently being done using mobile application. The figure shows that most students (about 60%) are using mobile applications to conduct group discussions, share study materials and share announcements. Other students are using mobile applications to share class announcements, study materials or engage in class discussions.

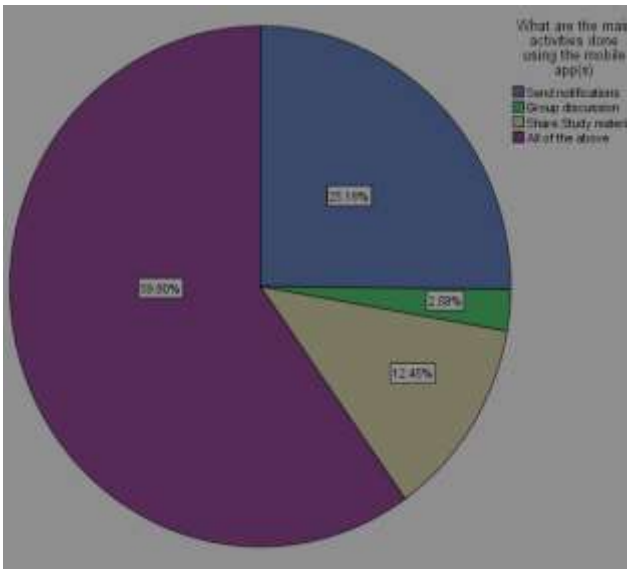



**Figure 5 Activities done using mobile applications**

- **Effectiveness of existing mobile application used in classes**

Figure 6 shows effectiveness of existing mobile apps being used in classes. Most respondents either feel the current applications are very effective (31.9%) or slightly effective (50.3%). The main weakness that was pointed out by students was that most applications (especially social interaction applications) were not designed to be used for educational purposes. Applications designed for educational purposes should allow students to share study material, allow them to research and engage in academic discussions. The students pointed out that they are unable to share study materials on most of the applications.

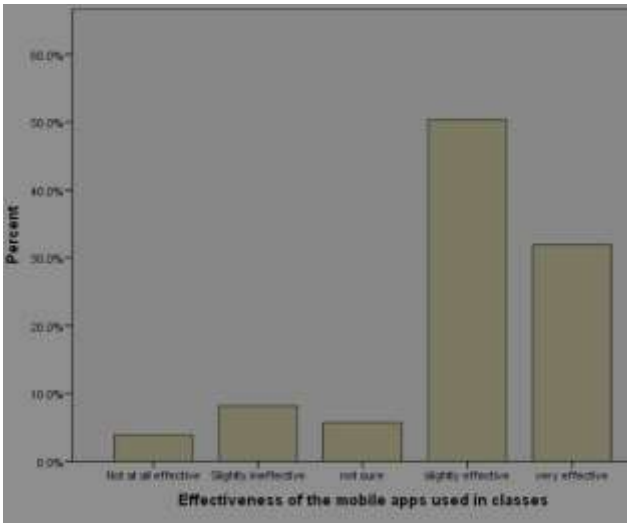



**Figure 6 Effectiveness of mobile applications used in classes**

- **Willingness to use collaborative learning app if developed**

Figure 7 shows that almost all the students would be willing to use a collaborative learning mobile app if it was to be developed. It can be observed that almost all the responded are willing to use a new collaboration app if developed. This is true even for those students who feel the apps currently being used are very effective as can be observed from table 4.

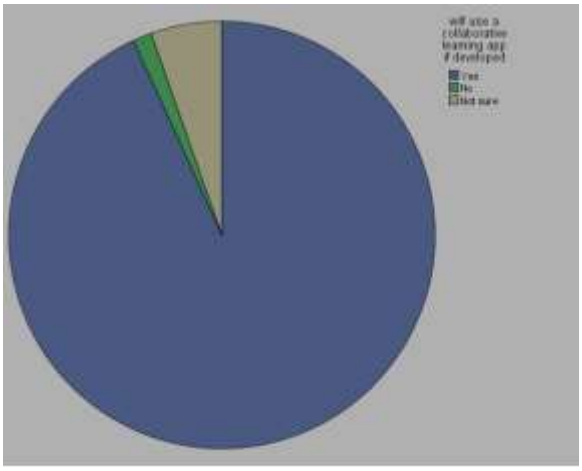



**Figure 7 willingness to use new collaboration apps if developed**

**Table 4 cross tabulation of the Effectiveness of apps currently being used in schools and willingness to use new app if developed**

| | | should a collaborative learning app be developed | | | Total |
|---|---|---|---|---|---|
| | | Yes | No | Not sure | |
| Effectiveness of the mobile apps used in classes | Not at all effective | 22 | 1 | 4 | 27 |
| | Slightly ineffective | 51 | 2 | 4 | 57 |
| | not sure | 30 | 0 | 3 | 33 |
| | slightly effective | 324 | 6 | 21 | 351 |
| | very effective | 217 | 1 | 4 | 222 |
| Total | | 650 | 10 | 36 | 696 |

- **Desired features in collaborative learning applications**

The respondents identified the following features that should be included in applications developed to assist in collaborative learning.

1. The application should allow the students to share study material.
2. The application should ensure that the discussions conducted and materials shared are strictly academic.
3. Allow students to view the study material offline.
4. The application should allow the students to chat with a tutor if they fail to grasp a concept.
5. The application should allow students to have voice conversations.
6. The cost of communication using the application should be affordable.

## IV. STUDY IMPLICATIONS

- The results indicate that most undergraduate students (around 87%) prefer to collaborate in their academic activities as opposed to working in isolation. Institutions should consider structure their education programs in a way that encourages student collaboration.
- Most students are using instant messaging applications such as WhatsApp to collaborate with their classmates. This is a clear indication that students have seen benefits in the use of mobile applications-aided collaborative learning. This is further supported by the fact that most of the students responded that they are willing to use mobile applications in collaborative learning (see figure 7).
- The existing telecommunication network is adequate for the successful implementation of mobile applications-aided collaborative learning, institution as seen from the fact that most students spend more than 1 hour of their day online.
- Most social networking application have enable students to interact. However, they lack features to effectively function as mobile learning applications. To effectively function as a mobile learning tool, an application should ensure that students are not distracted with non-academic matters at the time they are studying. In addition, the application should allow students to study note and take assessments. The application should also analyse the student performance and offer remedial suggestions [28, 29, 30].
- There is need to design collaborative learning applications that overcome the challenges of social networking applications identified in the previous point.
- Although collaborative learning is done in informal settings with no supervision from the educators, students would like the intervention of the educators in case they fail to grasp a concept.

## V. CONCLUSION

In this research, we used the UTAUT framework to determine the willingness of students at CBU to use mobile application-aided collaborative learning. According to the UTAUT framework, the acceptance and use of a technology will depend on four variables, namely, performance expectance, effort expectance, social influence and facilitating conditions. In this research, we have identified that students are aware of the services they can get from mobile application aided collaborative learning. We have also determined that the effort required to use mobile applications in collaboration is minimal. In addition, we have also established that most students and lecturers are using mobile applications to interact with students. This satisfies the third variable (that is, Social Influence) in UTAUT. Finally, we have also established that the facilitating conditions for the use of mobile applications in collaboration are available.

Since the necessary factors for the use of mobile applications in collaborative learn are available,

As it has been observed from the data collected in this research, there is growing interest in collaborative learning activities among students. Educators need to consider increasing the use of collaborative learning in higher education as it encourages students to improve their understanding of the material and to develop social skills that would help them excel in their education and their chosen careers. Developers of mobile learning applications should ensure that the tools they design contain all the features required for effective collaboration in learning.

## REFERENCES

[1] A. Becker, M. Cummins, A. Davis, A. Freeman, C. Hall Giesinger and V. Ananthanarayanan, "NMC Horizon Report: 2017 Higher Education Edition.,"

The New Media Consortium, Austin, Texas, 2017.

[2] N. Chaamwe and L. Shumba, "ICT Integrated Learning: Using Spreadsheets as Tools for e-Learning, A Case of Statistics in Microsoft Excel," International Journal of Information and Education Technology, vol. 6, no. 6, pp. 435-440, 2016.

[3] C. Cheong, V. Bruno and F. Cheong, "Designing a Mobile-app-based CollaborativeLearning System," Journal of Information Technology Education:Innovations in Practice, vol. 11, pp. 97-119, 2012.

[4] L. Lipponen, "Exploring Foundations for Computer-Supported Collaborative Learning," in *Fourth Computer Support for Collaborative Learning Conference*, 2002.

[5] I. Magnisalis, S. Demetriadis and A. Karakostas, "Adaptive and Intelligent Systems for Collaborative Learning Support:A Review of the Field," *IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES,* vol. 4, no. 1, 2011.

[6] Odiakaosa, D. N and J. N, "Teacher and Learner Perceptions on Mobile Learning Technology: A Case of Namibian High Schools from the Hardap Region," *HIGHER EDUCATOR-An International Journal,* vol. 1, no. 1, pp. 13-41, 2017.

[7] S. Ledlow, "Cooperative Learning in Higher Education.," Arizona State University, 1999.

[8] G. Smith, C. Sorensen, A. Gump, A. Heindel, M. Caris and C. Martinez, "Overcoming student resistance to group work: Online versus face-to-face," *Internet and Higher Education,* vol. 14, pp. 121-128, 2011.

[9] B. Chen and A. Denoyelles, "Exploring Students' Mobile Learning Practices in Higher Education," 7 October 2013. [Online]. Available: http://www.educause.edu/ero/article/exploring-students--mobile-learning-practices-higher-education. [Accessed 26 February 2018].

[10] The University of Arizona, "Collaborative Learning Spaces," 2018. [Online]. Available: http://academicaffairs.arizona.edu/cls.

[11] The University of Tennessee, "Cooperative Learning," 2017. [Online]. Available: https://www.utc.edu/walker-center-teaching-learning/teaching-resources/cooperative-learning.php. [Accessed 26 February 2018].

[12] A. Konak and M. Bartolacci, "Using a Virtual Computing Laboratory to Foster Collaborative Learning for Information Security and Information Technology Education," *Journal of Cybersecurity Education, Research and Practice,* 2016.

[13] H. Tseng and H. Yeh, "Team members' perceptions of online teamwork learning experiences and building teamwork trust: A qualitative study," Computers & Education, vol. 63, pp. 1-9, 2013.

[14] I. Oliveira, L. Tinoca and A. Pereira, "Online group work patterns: How to promote a successful collaboration," Computers & Education, vol. 57, pp. 1348-1357, 2011.

[15] C. Cheong, V. Bruno and F. Cheong, "Designing a Mobile-app-based CollaborativeLearning System," Journal of Information Technology Education:Innovations in Practice, vol. 11, pp. 97-119, 2012.

[16] H.-Y. Ku, H. Tseng and C. Akarasriworn, "Collaboration factors, teamwork satisfaction, and student attitudes toward online collaborative learning," Computers in Human Behavior , vol. 29, pp. 922-929, 2013.

[17] Ç. Güler, "Use of WhatsApp in Higher Education: What's Up With Assessing Peers Anonymously?," Journal of Educational Computing Research, vol. 55, no. 2, pp. 272-289, 2017.

[18] L. Cetinkaya, "The Impact of Whatsapp Use on Success in Education Process," International Review of Research in Open and Distributed Learning, vol. 18, no. 7, November 2017.

[19] C. Barhoumi, "The Effectiveness of WhatsApp Mobile Learning Activities Guided by Activity Theory on Students' Knowledge Management," CONTEMPORARY EDUCATIONAL TECHNOLOGY, vol. 6, no. 3, pp. 221-238, 2015.

[20] A. Deal, "A Teaching with Technology: Collaboration Tools White Paper," Carnegie Mellon University, 2009.

[21] D. Morrison, "How-to Integrate Collaboration Tools to Support Online Learning," 2016. [Online]. Available: https://onlinelearninginsights.wordpress.com/2016/07/02/how-to-integrate-online-collaboration-tools-to-support-learning/. [Accessed 26 February 2018].

[22] G. Burd, J. Pollard and J. Hunter, "Collaborative Learning Spaces Project (CLSP) Classroom Redesign for Active Learning Pedagogies," The University of Arizona, 2015.

[23] R. Basturk, "The effectiveness of computer-assisted instruction in teaching introductory statistics," Journal of Educational Technology & Society, vol. 8, no. 2, pp. 170-178, 2005.

[24] G. Stahl, T. Koschmann and D. Suthers, "Computer-Supported Collaborative Learning: A Historical Perspectiv," in *Cambridge Handbook of the Learning Sciences*, Cambridge University, 2006.

[25] E. Togatorop, "Teaching Writing with a Web Based Collaborative Learning," *International Journal of Economics and Financial Issues*, vol. 5, no. Special Issue, pp. 247-256, 2015.

[26] S. Chandrasekaran, G. Littlefair, M. Joordens and A. Stojcevski, "Cloud-Linked and Campus-Linked Students' Perceptions of Collaborative Learning and Design Based Learning in Engineering," *International Journal of Digital Information and Wireless Communications (*, vol. 4, no. 3, pp. 1-9, 2014.

[27] M. Laal, Z. Khattami-Kermanshahi and M. Laal, "Teaching and education; collaborative style," *Procedia - Social and Behavioral Sciences*, vol. 116, pp. 4057-4061, 2014.

[28] J. Lopez, A. Cerezo, J. Menendez and J. Ballesteros, "Usage of mobile devices as collaborative tools for education and preparation of official exams.," in *2015 I.E. International Symposium on Consumer Electronics (ISCE'15)*, 2015.

# THE MAJOR WIRELESS NETWORK SECURITY CHALLENGES - A REVIEW

Jimmy Katambo[1], Mayumbo Nyirenda[2] and Jackson Phiri[3]
The University of Zambia
Department of Computer Science
Lusaka, Zambia.
[1]jimmy.katambo@cs.unza.zm, [2]mayumbo@gmail.com, [3]jackson.phiri@cs.unza.zm

***Abstract*—The paper categorizes wireless networks into two major classes namely wireless ad hoc networks and cellular networks. Authors argue that the main difference between these two is whether a fixed infrastructure is present. They indicate that while cellular networks require fixed infrastructures to support the communication between mobile nodes and deployment of the fixed infrastructures is essential, Wireless ad hoc networks do not require a fixed infrastructure; thus it is relatively easy to set up and deploy a wireless ad hoc network. Security Protocols for Sensor Networks are a family of security protocols, which were specially designed for low end devices with severely limited resources, such as sensor nodes in sensor networks. This paper reviews a number of articles in the areas of Wireless Network Security and discusses the major Wireless Network Security challenges. To allow routers to automatically discover new routes and maintain their routing tables, routers exchange routing information periodically. Wireless Sensor Networks are not like Wired Sensor Networks or other types of wireless networks, and it is easier for the Wireless Sensor Networks to be attacked and more challenging to ensure the security of the Wireless Sensor Network. As a result, the security of Wireless Sensor Networks has been widely studied and many wonderful security policies have been proposed. The paper which is a review of the major Wireless Network Security challenges provides insight into major Wireless Network Security challenges.**

***Index Terms*— Attacks, Body Sensor Networks, Cellular Networks, Constraints, Key Establishment, Secure Network Encryption Protocol, Security Policies, Security Protocols.**

## I. INTRODUCTION

The paper provides insight into major Wireless Network Security challenges. With the rapid development of technology in wireless communication and microchips, wireless technology has been widely used in various application areas. The proliferation of wireless devices and wireless networks in the past decade shows the widespread use of wireless technology [1].

*Wireless networks* is a general term to refer to various types of networks that communicate without the need of wire lines. Wireless networks can be broadly categorized into two classes based on the structures of the networks: wireless ad hoc networks and cellular networks. The main difference between these two is whether a fixed infrastructure is present [1].

Three of the well-known cellular networks are the Global System for Mobile communication(GSM) network, the Code Division Multiple Access (CDMA) network, and the 802.11 wireless LAN (Local Area Network). The GSM network and the CDMA network are the main network technologies that support modern mobile communication, with most of the mobile phones and mobile networks that are built based on these two wireless networking technologies and their variants.

As cellular networks require fixed infrastructures to support the communication between mobile nodes, deployment of the fixed infrastructures is essential. Further, cellular networks require serious and careful topology design of the fixed infrastructures before deployment, because the network topologies of the fixed infrastructures are mostly static and will have a great impact on network performance and network coverage.

## II. CELLULAR NETWORKS

Cellular networks require fixed infrastructures to work (see Figure 11.2[1]). A cellular network comprises a fixed infrastructure and a number of mobile nodes. Mobile nodes connect to the fixed infrastructure through wireless links. They may move around from within the range of one base station to outside the range of the base station, and they can move into the ranges of other base stations. The fixed infrastructure is stationary, or mostly stationary, including base stations, links between base stations, and possibly other conventional network devices such as routers. The links between base stations can be either wired or wireless. The links should be more substantial than those links between base stations and mobile nodes in terms of reliability, transmission range, bandwidth, and so on [1].

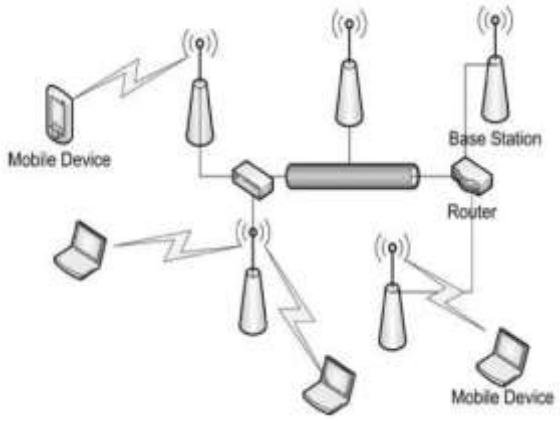



**Fig. 11.2 Cellular networking [1]**

The fixed infrastructure serves as the backbone of a cellular network, providing high speed and stable connection for the whole network, compared to the connectivity between a base station and a mobile node. In most cases, mobile nodes do not communicate with each other directly without going through a base station. A packet from a source mobile node to a destination mobile node is likely to be first transmitted to the base station to which the source mobile node is connected. The packet is then relayed within the fixed infrastructures until reaching the destination base station to which the destination mobile node is connected. The destination base station can then deliver the packet to the destination mobile node to complete the packet delivery [1].

### *A. Cellular Telephone Networks*

Cellular telephone networks offer mobile communication for most of us. With a cellular telephone network, base stations are distributed over a region, with each base station covering a small area. Each part of the small area is called a *cell* [1]. Cell phones within a cell connect to the base station of the cell for communication. When a cell phone moves from one cell to another, its connection will also be migrated from one base station to a new base station. The new base station is the base station of the cell into which the cell phone just moved. Two of the technologies are the mainstream for cellular telephone networks: The Global System for Mobile communication (GSM) and Code Division Multiple Access (CDMA) [1].

GSM is a wireless cellular network technology for mobile communication that has been widely deployed in most parts of the world. Each GSM mobile phone uses a pair of frequency channels, with one channel for sending data and another for receiving data. Time Division Multiplexing (TDM) is used to share frequency pairs by multiple mobiles [1].

CDMA is a technology developed by a company named Qualcomm and has been accepted as an international standard. CDMA assumes that multiple signals add linearly, instead of assuming that colliding frames are completely garbled and of no value. With coding theory and the new assumption, CDMA allows each mobile to transmit over the entire frequency spectrum at all times. The core algorithm of CDMA is how to extract data of interest from the mixed data [1].

### *B. 802.11 Wireless LANs*

Wireless LANs are specified by the IEEE 802.11 series standard [1], which describes various technologies and protocols for wireless LANs to achieve different targets, allowing the maximum bit rate from 2 Mbits per second to 248 Mbits per second. Wireless LANs can work in either Access Point (AP) mode or ad hoc mode, as shown in Figure 11.3 [1]. When a wireless LAN is working in AP mode, all communication passes through a base station, called an *access point*. The access point then passes the communication data to the destination node, if it is connected to the access point, or forwards the communication data to a router for further routing and relaying. When working in ad hoc mode, wireless LANs work in the absence of base stations. Nodes directly communicate with other nodes within their transmission range, without depending on a base station [1].

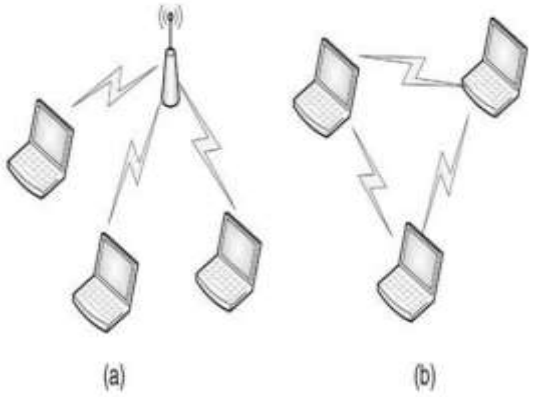



**Fig. 11.3 (a) A wireless network in AP mode; (b) a wireless network in ad hoc mode [1].**

One of the complications that 802.11 wireless LANs incur is medium access control in the data link layer. Medium access control in 802.11 wireless LANs can be either distributed or centralized control by a base station. The distributed medium access control relies on the Carrier Sense Multiple Access (CSMA) with Collision Avoidance (CSMA/CA) protocol. CSMA/CA allows network nodes to compete to transmit data when a channel is idle and uses the Ethernet binary exponential back off algorithm to decide a waiting time before retransmission when a collision occurs. CSMA/CA can also operate based on MACAW (Multiple Access with Collision Avoidance for Wireless) using virtual channel sensing. Request packets and clear-to-send (CTS) packets are broadcast before data transmission by the sender and the receiver, respectively. All stations within the range of the sender or the receiver will keep silent in the course of data transmission to avoid interference on the transmission [1].

The centralized medium access control is implemented by having the base station broadcast a beacon frame periodically and poll nodes to check whether they have data to send. The base station serves as a central control over the allocation of the bandwidth. It allocates bandwidth according to the polling results. All nodes connected to the base station must behave in accordance with the allocation decision made by the base station. With the centralized medium access control, it is possible to provide quality-of-service guarantees because the base station can control on the allocation of bandwidth to a specific node to meet the quality requirements [1].

## III. WIRELESS AD HOC NETWORKS

Wireless ad hoc networks are distributed networks that work without fixed infrastructures and in which each network node is willing to forward network packets for other network nodes. The main characteristics of wireless ad hoc networks are as follows:

- Wireless ad hoc networks are distributed networks that do not require fixed infrastructures to work. Network nodes in a wireless ad hoc network can be randomly deployed to form the wireless ad hoc network.
- Network nodes will forward network packets for other network nodes. Network nodes in a wireless ad hoc network directly communicate with other nodes within their ranges. When these networks communicate with network nodes outside

their ranges, network packets will be forwarded by the nearby network nodes and other nodes that are on the path from the source nodes to the destination nodes [1].

● Wireless ad hoc networks are self-organizing. Without fixed infrastructures and central administration, wireless ad hoc networks must be capable of establishing cooperation between nodes on their own. Network nodes must also be able to adapt to changes in the network, such as the network topology [1].

● Wireless ad hoc networks have dynamic network topologies. Network nodes of a wireless ad hoc network connect to other network nodes through wireless links [1].

The network nodes are mostly mobile. The topology of a wireless ad hoc network can change from time to time, since network nodes move around from within the range to the outside, and new network nodes may join the network, just as existing network nodes may leave the network [1].

### *A. Wireless Sensor Networks*

A wireless sensor network is an ad hoc network mainly comprising sensor nodes, which are normally used to monitor and observe a phenomenon or a scene. The sensor nodes are physically deployed within or close to the phenomenon or the scene. The collected data will be sent back to a base station from time to time through routes dynamically discovered and formed by sensor nodes [1].

Sensors in wireless sensor networks are normally small network nodes with very limited computation power, limited communication capacity, and limited power supply. Thus a sensor may perform only simple computation and can communicate with sensors and other nodes within a short range. The life spans of sensors are also limited by the power supply [1].

Wireless sensor networks can be self-organizing, since sensors can be randomly deployed in some inaccessible areas. The randomly deployed sensors can cooperate with other sensors within their range to implement the task of monitoring or observing the target scene or the target phenomenon and to communicate with the base station that collects data from all sensor nodes. The cooperation might involve finding a route to transmit data to a specific destination, relaying data from one neighbor to another neighbor when the two neighbors are not within reach of each other, and so on [1].

### *B. Mesh Networks*

One of the emerging technologies of wireless network is Wireless Mesh Networks (WMNs) [1]. Nodes in a WMN include mesh routers and mesh clients. Each node in a WMN works as a router as well as a host. When it's a router, each node needs to perform routing and to forward packets for other nodes when necessary, such as when two nodes are not within direct reach of each other and when a route to a specific destination for packet delivery is required to be discovered [1].

Mesh routers may be equipped with multiple wireless interfaces, built on either the same or different wireless technologies, and are capable of bridging different networks. Mesh routers can also be classified as access mesh routers, backbone mesh routers, or gateway mesh routers. Access mesh routers provide mesh clients with access to mesh networks; backbone mesh routers form the backbone of a mesh network; and a gateway mesh router connects the backbone to an external network.

Each mesh client normally has only one network interface that provides network connectivity with other nodes. Mesh clients are not usually capable of bridging different networks, which is different from mesh routers.

Similar to other ad hoc networks, a wireless mesh network can be self-organizing. Thus nodes can establish and maintain connectivity with other nodes automatically, without human intervention. Wireless mesh networks can be divided into backbone mesh networks and access mesh networks [1].

## IV. SECURITY PROTOCOLS

Wired Equivalent Privacy (WEP) was defined by the IEEE 802.11 standard [1]. WEP is designed to protect linkage-level data for wireless transmission by providing confidentiality, access control, data integrity and to provide secure communication between a mobile device and an access point in an 802.11 wireless LAN [1].

### *A. WEP*

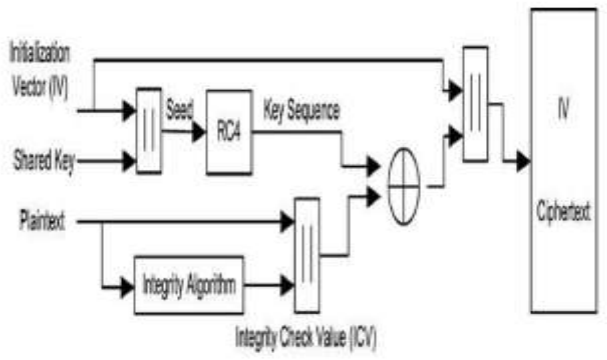
Implemented based on shared key secrets and the RC4 stream cipher, WEP's encryption of a frame includes two operations (see Figure 11.4 [1]). It first produces a checksum of the data, and then it encrypts the plain text and the checksum using RC4:

● *Checksumming*. Let *c* be an integrity checksum function. For a given message *M*, a checksum *c* (*M)* is calculated and then concatenated to the end of *M*, obtaining a plain text *P*<*M*, *c* (M)>. Note that the checksum *c* (*M)* does not depend on the shared key.

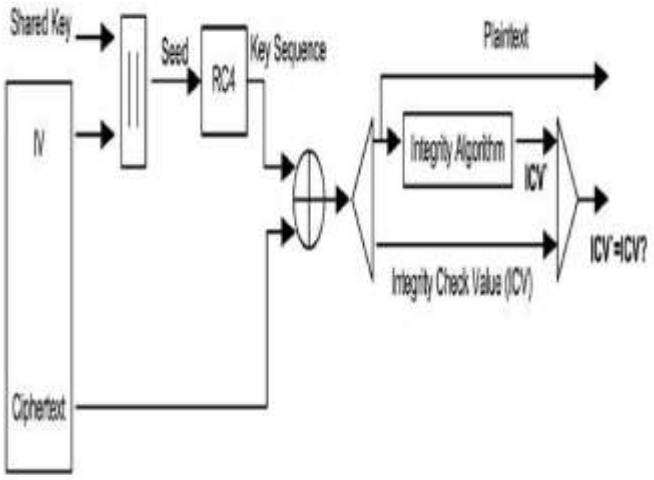
● *Encryption.* The shared key *k is* concatenated to the end of the initialization vector (IV) *v*, forming <*v*, *k*>. <*v*, *k*> is then used as the input to the RC4 algorithm to generate a keystream *RC* 4(v, *k)*. The plain text *P* is exclusive-or'ed (XOR, denoted by ⊕) with the keystream to obtain the cipher text:

$$C = P \oplus RC4(v, k) . \quad (1)$$

Using the shared key *k* and the (IV) *v*, WEP can greatly simplify the complexity of key distribution because it needs only to distribute *k* and *v* but can achieve a relatively very long key sequence [1]. IV changes from time to time, which will force the *RC* 4 algorithm to produce a new key sequence, avoiding the situation where the same key sequence is used to encrypt a large amount of data, which potentially leads to several types of attacks [2,3].




(a)WEP Encryption [1]




(b)WEP Decryption [1]

**Fig. 11.4 WEP encryption and decryption [1].**

WEP combines the shared key *k* and the IV *v* as inputs to seed the *RC* 4 function. 802.11B specifies that the seed shall be 64 bits long, with 24 bits from the (IV) *v* and 40 bits from the shared key *k*. Bits 0 through 23 of the seed contain bits 0 through 23 of the (IV) *v*, and bits 24 through 63 of the seed contain bits 0 through 39 of the shared key *k* [1].

When a receiver receives the cipher text *C*, it will XOR the cipher text *C* with the corresponding keystream to produce the plaintext *M* as follows:

$$M' = C \oplus RC(k,v) \; (P \oplus RC(k,v)) \oplus RC(k,v) = M \qquad (2)$$

*B. WPA and WPA2*

Wi-Fi Protected Access (WPA) is specified by the IEEE 802.11i standard, which is aimed at providing stronger security compared to WEP and is expected to tackle most of the weakness found in WEP [1].

(1) WPA

WPA has been designed to target both enterprise and consumers. Enterprise deployment of WPA is required to be used with IEEE 802.1x authentication, which is responsible for distributing different keys to each user. Personal deployment of WPA adopts a simpler mechanism, which allows all stations to use the same key. This mechanism is called the *Pre-Shared Key* (PSK) mode [1]. The WPA protocol works in a similar way to WEP. WPA mandates the use of the *RC* 4 stream cipher with a 128 – bit key and a 48 – bit initialization vector (IV), compared with the 40 – bit key and the 24 – bit (IV) in WEP [1].

WPA also has a few other improvements over WEP, including the Temporal Key Integrity Protocol (TKIP) and the Message Integrity Code (MIC). With TKIP, WPA will dynamically change keys used by the system periodically. With the much larger IV and the dynamically changing key, the stream cipher *RC* 4 is able to produce a much longer keystream. The longer keystream improved WPA's protection against the well-known key recovery attacks on WEP, since finding two packets encrypted using the same key sequences is literally impossible due to the extremely long keystream [1].

With MIC, WPA uses an algorithm named Michael to produce an authentication code for each message, which is termed the *message integrity code [1]*. The message integrity code also contains a frame counter to provide protection over replay attacks [1].

WPA uses the Extensible Authentication Protocol (EAP) framework to conduct authentication. When a user (supplicant) tries to connect to a network, an authenticator will send a request to the user asking the user to authenticate herself using a specific type of authentication mechanism. The user will respond with corresponding authentication information. The authenticator relies on an authentication server to make the decision regarding the user's authentication [1].

(2) WPA2

WPA2 is not much different from WPA. Though TKIP is required in WPA, Advanced Encryption Standard (AES) is optional. This is aimed to provide backward compatibility for WPA over hardware designed for WEP, as TKIP can be implemented on the same hardware as those for WEP, but AES cannot be implemented on this hardware. TKIP and AES are both mandatory in WPA2 to provide a higher level of protection over wireless connections. AES is a block cipher, which can only be applied to a fixed length of data block. AES accepts key sizes of 128 bits, 196 bits, and 256 bits [1].

*C. SPINS: SECURITY PROTOCOLS FOR SENSOR NETWORKS*

Sensor nodes in sensor networks are normally low-end devices with very limited resources, such as memory, computation power, battery, and network bandwidth.

Perrig et al. [4] proposed a family of security protocols named SPINS, which were specially designed for low end devices with severely limited resources, such as sensor nodes in sensor networks. SPINS consists of two building blocks: Secure Network Encryption Protocol (SNEP) and the "micro" version of the Timed, Efficient, Streaming, Loss-tolerant Authentication Protocol (μTESLA). SNEP uses symmetry encryption to provide data confidentiality, two-party data authentication, and data freshness. μTESLA provides authentication over broadcast streams. SPINS assumes that each sensor node shares a master key with the base station. The master key serves as the base of trust and is used to derive all other keys [1].

(1) SNEP

As illustrated in Figure 11.5, SNEP uses a block cipher to provide data confidentiality and message authentication code (MAC) to provide authentication. SNEP assumes a shared counter *C* between the sender and the receiver and two keys, the encryption key $K_{encr}$ and the authentication key $K_{mac}$. For an outgoing message *D*, SNEP processes it as follows:

- The message *D* is first encrypted using a block cipher in counter mode with the key $K_{encr}$ *and* the counter *C* , forming the encrypted text $E = \{D\}_{<Kencr,C>}$.
- A message authentication code is produced for the encrypted text *E* with the key $K_{mac}$ and the counter *C*, forming the MAC $M = MAC(K_{mac}, C \mid E)$ where *MAC ()* is a one-way function and $C \mid E$ stands for the concatenation of *C* and *E*.
- SNEP increments the counter *C*.

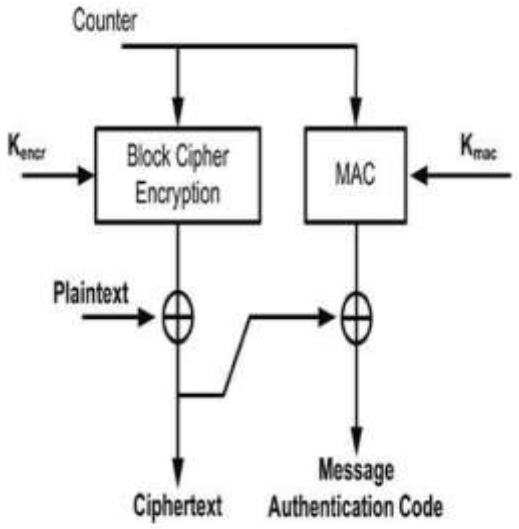



**Fig. 11.5 Sensor Network Encryption Protocol (SNEP) [1].**

To send the message *D* to the recipient, SNEP actually sends out *E* and *M*. In other words, SNEP encrypts *D* to *E* using the shared key $K_{encr}$ between the sender and the receiver to prevent unauthorized disclosure of the data, and it uses the shared key $K_{mac}$, known only to the sender and the receiver, to provide message authentication. Thus data confidentiality and message authentication can both be implemented [1].

The message *D* is encrypted with the counter *C*, which will be different in each message. The same message *D* will be encrypted differently even it is sent multiple times. Thus semantic security is implemented in SNEP. The MAC is also produced using the counter *C*; thus it enables SNEP to prevent replying to old messages [1].

(2) μTESLA

TESLA was proposed to provide message authentication for multicast. TESLA does not use any asymmetry cryptography, which makes it lightweight in terms of computation and overhead of bandwidth. μTESLA is a modified version of TESLA, aiming to provide message authentication for multicasting in sensor networks. The general idea of μTESLA is that the sender splits the sending time into intervals. Packets sent out in different intervals are authenticated with different keys. Keys to authenticate packets will be disclosed after a short delay, when the keys are no longer used to send out messages [1]. Thus packets can be authenticated when the authentication keys have been disclosed. Packets will not be tampered with while they are in transit since the keys have not been disclosed yet. The disclosed authentication keys can be verified using previous known keys to prevent malicious nodes from forging authentication keys [1].

μTESLA has four phases: sender setup, sending authenticated packets, bootstrapping new receivers, and authenticating packets. In the sender setup phase, a sender generates a chain of keys, $K_i$ $(0 \leq i \leq n)$. The keychain is a one-way chain such that $K_i$ can be derived from $K_j$ if $i \leq j$, such as a keychain $K_i$ (i =0,…,n), $K_i = F(K_{i+1})$, where *F* is a one-way function. The sender [1] also decides on the starting time $T_0$, the interval duration $T_{int}$, and the disclosure delay *d* (unit is interval), as shown in Figure 11.6 [1].

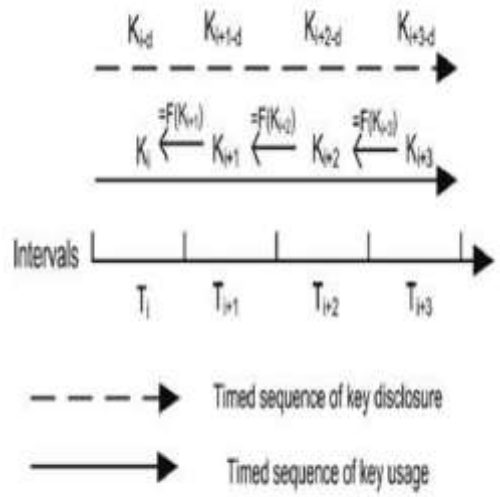



**Fig. 11.6 Sequences of intervals, key usages, and key disclosure [1].**

To send out authenticated packets, the sender attaches a MAC with each packet, where the MAC is produced using a key from the keychain and the data in the network packet. μTESLA has specific requirements on the use of keys for producing MACs. Keys are used in the same order as the key sequence of the keychain. Each of the keys [1] is used in one interval only. For the interval $T_i = T_0 + I \times T_{int}$, the key $K_i$ is used to produce the MACs for the messages sent out in the interval $T_i$ [1]. Keys [1] are disclosed with a fixed delay *d* such that the key $K_i$ used in interval $T_i$ will be disclosed in the interval $T_{i+d}$. The sequence of key usage and the sequence of key disclosure are demonstrated in Figure 11.6 [1].

To bootstrap a new receiver, the sender needs to synchronize the time with the receiver and needs to inform the new receiver of a key $K_j$ that is used in a past interval $T_j$, the interval duration $T_{int}$, and the disclosure delay *d* [1]. With a previous key $K_j$, the receiver will be able to verify any key $K_p$ where $j \leq p$ using the one-way keychain's property. After this, the new receiver will be able to receive and verify data in the same way as other receivers that join the communication prior to the new receiver [1].

To receive and authenticate messages, a receiver will check all incoming messages if they have been delayed for more than *d*. Messages with a delay greater than *d* will be discarded, since they are suspect as fake messages constructed after the key has been disclosed. The receiver will buffer the remaining messages for at least *d* intervals until the corresponding keys are disclosed. When a key $K_i$ is disclosed at the moment $T_{i+d}$, the receiver will verify $K_i$ using $K_{i-1}$ by checking if $K_{i-1} = F(K_i)$. Once the key $K_i$ is verified, $K_i$ will be used to authenticate those messages sent in the interval $T_i$ [1].

## V.SECURE ROUTING

Secure Efficient Ad hoc Distance (SEAD) vector routing is designed based on Destination-Sequenced Distance Vector (DSDV) routing. SEAD augments DSDV with authentication to provide security in the construction and exchange of routing information [1].

(1) SEAD

Distance vector routing works as follows. Each router maintains a routing table. Each entry of the table contains a specific destination, a metric (the shortest distance to the destination), and the next hop on the shortest path from the current router to the destination. For a packet [1] that needs to be sent to a certain destination, the router will look up the destination from the routing table to get the matching entry. Then the packet is sent to the next hop specified in the entry [1].

To allow routers to automatically discover new routes and maintain their routing tables, routers exchange routing information periodically. Each router advises its neighbors of its own routing information by broadcasting its routing table to all its neighbors. Each router will update its routing table according to the information it hears from its neighbors. If a new destination is found from the information advertised by a neighbor, a new entry is added to the routing table with the metric recalculated based on the advertised metric and the linking between the router and the neighbor [1].

If an existing destination is found, the corresponding entry is updated only when a new path that is shorter than the original one has been found. In this case, the metric and the next hope for the specified destination are modified based on the advertised information. Though distance vector routing is simple and effective, it suffers from possible routing loops, also known as the counting to infinity problem. DSDV [1] is one of the extensions to distance vector routing to tackle this issue. DSDV augments each routing update with a sequence number, which can be used to identify the sequence of routing updates, preventing routing updates being applied in an out-of-order manner. Newer routing updates are advised with sequence numbers greater than those of the previous routing updates [1].

SEAD provides authentication on metrics' lower bounds and senders' identities by using the one-way hash chain. Let $H$ be a hash function and $x$ be a given value [1]. A list of values [1] is computed as follows:

$$h_0, h_1, h_2, \ldots, h_n$$

where $h_0 = x$ and $h_{i+1}$ for $0 \leq i \leq n$. Given any value $h_k$ that has been confirmed to be in the list, to authenticate if a given value $d$ is on the list or not one can compute if $d$ can be derived from $h_k$ by applying $H$ a certain number of times, or if $h_k$ can be derived from $d$ by applying $H$ to $d$ a certain number of times [1]. If either $d$ can be derived from $h_k$ or $h_k$ can be derived from $d$ within a certain number of steps, it is said that $d$ can be authenticated by $h_k$.

(2) ARIADNE

Ariadne is a secure on-demand routing protocol for ad hoc networks. Ariadne is built on the Dynamic Source Routing protocol (DSR) [1]. Routing in Ariadne is divided into two stages: the route discovery stage and the route maintenance stage. In the route discovery stage, a source node in the ad hoc network tries to find a path to a specific destination node [1]. The discovered path will be used by the source node as the path for all communication from the source node to the destination node until the discovered path becomes invalid [1].

In the route maintenance stage, network nodes identify broken paths that have been found. A node sends a packet along a specified route to some destination. Each node on the route forwards the packet to the next node on the specified route and tries to confirm the delivery of the packet to the next node. If a node fails to receive an acknowledgment from the next node, it will signal the source node using a ROUTE ERROR packet that a broken link has been found. The source node and other nodes on the path can then be advised of the broken link [1].

The key security features Ariadne adds onto the route discovery and route maintenance are node authentication and data verification for the routing relation packets. Node authentication is the process of verifying the identifiers of nodes that are involved in Ariadne's route discovery and route maintenance, to prevent forging routing packets [1].

In route discovery, a node sends out a ROUTE REQUEST packet to perform a route discovery. When the ROUTE REQUEST packet reaches the destination node, the destination node verifies the originator identity before responding. Similarly, when the source node receives a ROUTE REPLY packet, which is a response to the ROUTE REQUEST packet, the source node will also authenticate the identity of the sender. The authentication of node identities can be of one of the three methods: TELSA, digital signatures, and Message Authentication Code (MAC) [1].

(3) ARAN

Authenticated Routing for Ad hoc Networks (ARAN) [1] is a routing protocol for ad hoc networks with authentication enabled. It allows routing messages to be authenticated at each node between the source nodes and the destination nodes. The authentication that ARAN has implemented is based on cryptographic certificates. ARAN requires a trusted certificate server, the public key of which is known to all valid nodes. Keys are assumed to have been established between the trusted certificate server and nodes [1].

According to [5], the current authentication protocols are generally divided into three categories namely identity based authentication protocol, authentication protocol based on traditional public key cryptography, and certificate-less authentication protocol.

For each node to enter into a wireless ad hoc network, it needs to have a certificate issued by the trusted server. The certificate contains the IP address of the node, the public key of the node, a time stamp indicating the issue time of the certification, and the expiration time of the certificate. Because all nodes have the public key of the trusted server, a certificate can be verified by all nodes to check whether it is authentic [2].

(4) SLSP

Secure Link State Routing Protocol (SLSP) is a secure routing protocol for ad hoc network building based on link state protocols. SLSP assumes that each node has a public/private key pair and has the capability of signing and verifying digital signatures. Keys are bound with the Medium Access Code and the IP address, allowing neighbors within transmission range to uniquely verify nodes if public keys have been known prior to communication [1].

In SLSP, each node broadcasts its IP address and the MAC to its neighbor with its signature. Neighbors verify the signature and keep a record of the pairing IP address and the MAC. The Neighbor Lookup Protocol (NLP) of SLSP extracts and retains

the MAC and IP address of each network frame received by a node. The extracted information is used to maintain the mapping of MACs and IP addresses [1].

## VI. KEY ESTABLISHMENT

Because wireless communication is open and the signals are accessible by anyone within the vicinity, it is important for wireless networks to establish trust to guard the access to the networks [1]. Key establishment builds relations between nodes using keys; thus security services, such as authentication, confidentiality, and integrity can be achieved for the communication between these nodes with the help of the established keys. The dynamically changing topology of wireless networks, the lack of fixed infrastructure of wireless ad hoc and sensor networks, and the limited computation and energy resources of sensor networks have all added complication to the key establishment process in wireless networks [1].

### (1) BOOTSTRAPPING

Bootstrapping is the process by which nodes in a wireless network are made aware of the presence of others in the network. On bootstrapping, a node gets its identifying credentials that can be used in the network the node is trying to join. Upon completion of the bootstrapping, the wireless network should be ready to accept the node as a valid node to join the network.

To enter a network, a node needs to present its identifying credential to show its eligibility to access the network. This process is called *preauthentication*. Once the credentials are accepted, network security associations are established with other nodes [1].

### (2) KEY MANAGEMENT

Key management schemes can be classified according to the way keys are set up (see Figure 11.7). Either keys [1] are managed based on the contribution from all participating nodes in the network or they are managed based on a central node in the network. Thus key management Schemes [1] can be divided into contributory key management schemes, in which all nodes work equally together to manage the keys, and distributed key management schemes, in which only one central node is responsible for key management [1].

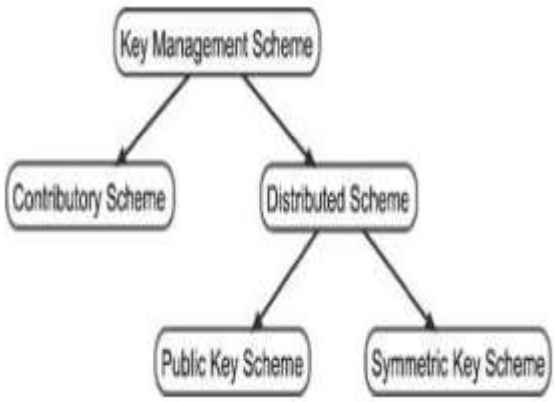



**Fig. 11.7 Key management schemes [1].**

*Classification*

The distributed key management scheme can be further divided into symmetric schemes and public key schemes. Symmetric key schemes are based on private key cryptography, whereby shared secrets are used to authenticate legitimate nodes and to provide secure communication between them. The underlying assumption is that the shared secrets are known only to legitimate nodes involved in the interaction. Thus proving the knowledge of the shared secrets is enough to authenticate legitimate nodes. Shared secrets are distributed via secure channels or out-of-band measures. Trust on a node is established if the node has knowledge of a shared secret [1].

Public key schemes are built on public key cryptography. Keys are constructed in pairs, with a private key and a public key in each pair. Private keys are kept secret by the owners. Public keys are distributed and used to authenticate nodes and to verify credentials. Keys are normally conveyed in certificates for distribution. Certificates are signed by trusted nodes for which the public keys have been known and validated. Trust on the certificates will be derived from the public keys that sign the certificates. Note that given $g^i \pmod p$ and $g^j \pmod p$, it is hard to compute $g^{i+j} \pmod p$ without the knowledge of $i$ and $j$ [1].

## VII. WIRELESS NETWORK SECURITY CHALLENGES

In [6], Cellular network forensics is a cross-discipline of digital forensics and cellular networks with the goal to investigate cellular network-facilitated crimes under a legally obtained warrant for the purpose of crime reconstruction. These criminal activities can be carried out with a direct network support (e.g., perpetrators communicate over a cellular network) or network is incidental to the crime (e.g., the network can provide historical data about calls or user locations). The investigations in cellular network can be in real time and non-real-time. The real-time investigations work with evidence transiting over the network at the time of the crime or the attack like ongoing calls, browsing sessions, or triangulated geolocation coordinates of a user [6].

The non-real-time investigations work with evidence in relation to past user activity such as charging data records or user's most visited cell. Prior to every investigation, operators and law enforcement agencies (LEAs) must establish forensic readiness to ensure secure identification, acquisition, and delivery of cellular network evidence. These operations are realized with two forensics mechanisms, Lawful Interception (LI) and Lawful Access Location Services (LALS) [6].

In his article, F. Sharevski [6] reviews the implementation of Lawful Interception (LI) and Lawful Access Location Services (LALS) in Long Term Evolution (LTE) and LTE-Advanced networks. Various types of LI and LALS evidence are also presented together with tools and techniques for cellular network forensic analysis. The challenges for continuous support of LI and LALS are discussed in the context of the key technologies for 5G evolution including Control and User Plane Separation (CUPS), Network Functional Virtualization (NFV), network slicing, and CIoT (Cellular Internet of Things) [6].

Several adaptations of the current LI and LALS operations for each 5G technology are proposed and elaborated to ensure the future cellular network forensic investigations are conducted as similarly as possible to the current practice. The

article concludes with a discussion of the legal and privacy aspects of the current and future cellular network forensics practice [6].

In their study, R. Beri and S. Singh [7] provide the information related to hidden node problem and also enlists some of the mechanism to avoid the hidden node problem. In a nutshell, the wireless ad hoc network is the infrastructure-less network that changes its location and gets connected with other nodes in the network on the fly. In this type of network, the number of nodes involved in the network is very large. The nodes in the network shared data with help of the node termed as access point. Each time, the node wants to send data to any other node, it will first send that data to access point, then that access point sends data to other receiving node [7].

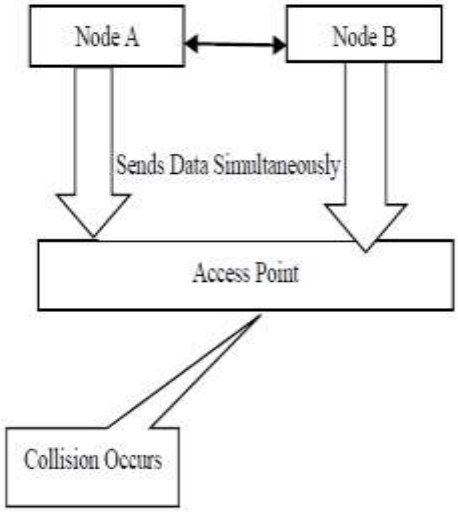



**Fig. 1: Hidden Node Problem [7].**

Each node is sensed by the access point, but may or may not be sensed by the other nodes in the network. If multiple hidden nodes send data to the access point at the same time, then collision occurs. So, it is necessary to avoid the hidden node problem. In their study [7], they found the methods to avoid hidden node problem occurring in the network. Some of the solutions that can be applied to solve the hidden node problem are as follows;

- ✓ Increase Transmitting Power from Nodes
- ✓ Use of Omni Directional Antennas
- ✓ Moving the Nodes
- ✓ Protocol Magnification Software

The future aspect of this study is the mechanisms involved in detecting the hidden node problem by simulating the behavior of the network [7].

According to M. Dahiya [8], Wi-Fi and Bluetooth are one of the most used radio technologies which supplement a new idea for pairing mobile devices via wireless connections, named as Mobile Ad-hoc Network (MANET) where mobile users come to a destination inside the area for communication. System of connections is built of individual setup of wireless links pairing mobile nodes. Nodes create a random type of topology in which nodes can shift as necessary for wireless communications. Even today security problems in MANET are not totally solved as MANET have active wavering topology when communicating between nodes [8].

Her paper [8] focusses on a few secure routing protocols. To secure routing protocols from attacks some solutions have been proposed that include the following:

- SAR (Security Aware Ad-hoc Routing)
- SAODV (Secure Ad-hoc on Demand Distance Vector)
- ARAN (Authenticated Routing for Ad-hoc Network)
- ARIADNE
- SRP (Secure Routing Protocol)
- SEAD (Secure Efficient Ad-hoc Distance Vector)
- SLSP (Secure Link State Routing Protocol)
- DSDV (Destination-Sequenced Distance Vector Protocol)
- DSR (Dynamic Secure Routing)

S. Gupta, M. Singh and S. Srivastava [9] in their survey paper, describe the brief overview regarding WSN, their applications and the factors influencing the sensor network design. Features such as flexibility, reusability, fast development of networks create endless new and thrilling areas for sensing. In future, this wide range where networks can be applied would take an integral place in our life. But still, sensor networks need to fulfill the constraints brought by dominating factors such as scalability issues, fault tolerance, cost of production, change in the layout of network and power consumption [9].

According to Y. Liu and K. Tong [10], Wireless Mesh Networks (WMNs) have all these great features used for communication in the IoT networks. It is still under-developed despite the industry advancing today. With the much more powerful MicroController Units (MCU) and processors today, the dynamic network topology can be achieved even in the tiny IoT devices. The mesh network topology has its unique advantage and disadvantage in the world of the IoT networks that can leverage the scale, distributed nature and lower requirement of data-rate of the IoT devices [10]. The advantage certainly outweighs the disadvantages of using the WMNs in such environment. Newer hybrid WMNs can be a solid choice when it comes to designing the structure of the network, especially in the remote areas with its robustness and scalability. Their paper [10] is simply discussing the possibility and the basic way of integrating such under-utilized network topology into the current and future IoT networks in the background of the advanced technology we have today. WMNs will certainly make a difference in the industry once being deployed on large scale in the IoT world and make the IoT more accessible to a wider audience [10].

M.A. Martin et al [11] says Wireless local area network (WLAN) can provide e-government services at all levels, from local to national as the WLAN enabled devices have the flexibility to move from one place to another within offices while maintaining connectivity with the network. However, government organizations are subject to strict security policies and other compliance requirements. Therefore, WLAN must ensure the safeguard of the privacy of individual data with the strictest levels of security [11].

The 802.11 MAC specifications describe an encryption protocol called Wired Equivalent Privacy (WEP) which is used to protect wireless communications from eavesdropping. It is also capable of preventing unauthorized access. However, the WEP protocol often fails to accomplish its security goal due to

the weakness in RC4 and the way it is applied in WEP protocol [11].

Their paper [11] focuses on the improvement of existing WEP protocol using the varying secret key for each transmission. This will remove the insecurities that currently make the RC4 unattractive for secured networking and this will add further cryptographic strength if applied to Rijndael algorithm. Their result shows that the proposed algorithm is more suitable for small and medium packets and Advanced Encryption Standard (AES) for large packets [11].

According to A. L. Alsahlany [12], WLANs being the most spread technology over the world are vulnerable to the threats of hacking. It is very important to protect a network from the hackers in order to prevent exploitation of confidential data [12]. Based on experimental results, the current implementation of WEP has proven to be flawed. Suggestions for improvement and tighten security of WLAN changing static WEP to the latest wireless encryption suit such as enterprise WPA2, or combination of enabling WEP key and MAC address filter security mechanism [12].

The experiment presented in their paper is used to emphasize that the WEP at any key size is an exposed encryption and shouldn't be the used to secure WLAN. In their paper, an active attack on the WEP protocol to recover a 64 and 128 bit WEP key using less than 50,000 distinctive IV packets is performed practically. An active attack [12] on the WEP protocol is able to recover a 64 and 128 bit WEP key using less than 50,000 frames with a success probability of 100% [12].

F. M. Sun et al [13] observe in their paper that the future of wireless sensor networks is promising; they are being deployed in many real-world applications, in the context of Ubiquitous Computing, Pervasive Computing, and Ambient Intelligence. In their paper, they concluded the unique characteristics of the wireless sensor networks and presented the requirements and the corresponding challenges of the WSNs security. Commonly seen WSNs attacks are introduced and classified according to different criteria and security approaches and key security techniques are presented in the following [13].

Finally, they summarized the security related issues and technologies in the area of body sensor networks as an illustrative example of the WSNs attacks and security mechanisms. Hopefully by reading their paper, the beginners can have a better view of attacks and countermeasures in wireless sensor networks and the researchers can be motivated to design smarter and more robust security mechanisms and make their networks safer [13].

According to S. Zadoo et al [14], the need for security in communication has existed in military communications for thousands of years. In their paper, their focus is on network protocols that provide security services. Wireless sensor network is an emerging technology that shows application both for public as well as military purpose. Monitoring is one of the main applications. Their paper compares all the protocols which are designed for security and proposed an improved protocol that reduces communication overhead by removing data redundancy from the network [14].

By using message sequence number, they can check whether it is old message or new message. If the message is old, then no need to send that message thereby reducing overhead. It also integrates security by data freshness in the protocol. This extra feature of freshness shows that this is superior to the existing protocols. This also improves the efficiency [14].

According to L. K. Musambo and J. Phiri [15], despite the control measures put to monitor traffic, there are additional trace back challenges beyond the reach of the Internet. Any IP-based trace back method assumes that the true source IP belongs to the computer being used by the attacker. However, in many scenarios this is not true e.g. Internet-connected mobile phone networks, open wireless (Wi-Fi) networks and public computers, such as those at libraries and Internet cafes. Most modern cell phones support text-messaging services such as Short Message Service (SMS), and many smart phones also have full featured IM software. Therefore, the botmaster can use a mobile device to control the botnet from any location with cell phone reception using a protocol translation service or a special IRC client for mobile phones [15].

M. Chibuye and J. Phiri [16] observe that in order to introduce modern warehousing, improve upon the storage of grain and grain marketing business processes for the Food Reserve Agency in Zambia, a prototype of a remote sensor network was developed and built as a proof of concept for a much wider deployment using cloud computing and the internet of things concept. It was determined that a wireless sensor network would aid the Food Reserve Agency in analytics, timely action and real-time reporting from all its food depots spread-out throughout Zambia. Google's Android Things Platform was used in order to achieve the objectives [16].

Advantages of Android Things over traditional platforms that have been used to develop wireless sensor networks were looked into and presented in their paper. The choice to use Google's Android Things platform for the development of a remote sensor network is mainly due to the fact that Google has created a platform with advantages that they believe are essential towards the further development and standardization of remote sensor network architecture at both the hardware and software level [16].

According to C. Chembe et al [17], Vehicular Ad Hoc Network (VANET) is envisaged to play an important role in the safety of drivers and passengers when moving on the roads. However, VANET still faces many challenges before it could be deployed. One such challenge is shortage of radio frequency spectrum channels. VANET has been allocated 7 channels for dedicated short range communication at 5.9 GHz band. The 7 channels are likely to get congested in high vehicle densities when many vehicles are contending for the same medium. Consequently, affecting the transmission of safety and emergency messages [17]. To alleviate the problem of scarcity of channels, dynamic spectrum access (DSA) through cognitive radio (CR) technology has been proposed. One of the core functions of a CR is to identify spectrum holes in licensed frequency bands that can be accessed by unlicensed users through spectrum sensing.

In VANET, spectrum sensing is challenging because of the mobility nature of vehicles, dynamic topological changes as well as other unique characteristics not found in other networks. However, these challenges have not been fully studied and how they affect spectrum sensing in cognitive vehicular network (CVN). In their paper [17], they discuss challenges associated with spectrum sensing in CVN. They describe the primary system activity model used by many

schemes proposed in literature. Furthermore, they present an in depth analysis of state-of-art cooperative spectrum sensing techniques for CVN from 2010 to May 2016. In addition, they present some of the open issues in spectrum sensing for CVN [17].

## VIII.CONCLUSION

Wireless technology has been widely used in various application areas. Wireless networks can be broadly categorized into two classes based on the structures of the networks: wireless ad hoc networks and cellular networks. Routing protocols such as Mobile Ad-hoc Network (MANET) need to be protected from various attacks. Sensor network is a promising and upcoming technology with usage in important applications.

Because wireless communication is open and the signals are accessible by anyone within the vicinity, it is important for wireless networks to establish trust to guard the access to the networks. Key establishment builds relations between nodes using keys; thus security services, such as authentication, confidentiality, and integrity can be achieved for the communication between these nodes with the help of the established keys.

In future we intend to undertake a review of the challenges of System Security and how they can be mitigated.

## REFERENCES

[1] J. R. Vacca, "Computer and Information Security Handbook," Burlington, USA: Morgan Kaufmann Publishers, 2009, pp. 169-181.

[2] International Standardization Organization: ISO/IEC 27035:2011 –Information security incident management, (Geneva, 2016).

[3] International Standardization Organization: ISO/IEC 27037:2012 guidelines for Identification, collection, acquisition and preservation of digital evidence, (Geneva, 2012).

[4] A. Perrig, R. Szewczyk, V. Wen, D. Culler, J.D. Tygar, SPINS: Security protocols for sensor networks, MobiCom' 01: Proceedings of the 7th annual international conference on Mobile computing and networking, 2001.

[5] D. Kunda and M. Chishimba, "A Survey of Android Mobile Phone Authentication Schemes. Mobile Networks and Applications, "pp. 1-9, Aug. 2018. Available. Accessed on: Nov 28, 2018. https://doi.org/10.1007/s11036-018-1099-7.

[6] F. Sharevski, "Towards 5G cellular network forensics," pp. 1-16, Jan. 2018.

[7] R. Beri and S. Singh, "Hidden Node Problem in Wireless Ad-Hoc Network," pp. 1-7, Jan. 2017.

[8] M. Dahiya [6], "MANET's: Security Attacks and Securing Routing Protocols," pp. 1-7, Apr. 2017.

[9] S. Gupta, M. Singh and S. Srivastava, "Wireless Sensor Network: A Survey," pp. 1-6, Oct. 2018.

[10] Y. Liu and K. Tong, "Wireless Mesh Networks in IoT networks," pp. 1-4, May 2017.

[11] M.A. Martin, M.A. Kabi, K.A. Sayeed, T. Mehenaz and M. Kamruzzaman, "Triple Layered Encryption Algorithm for IEEE 802.11 WLANs in E-Government Services," pp. 1-7, Sep. 2013.

[12] A. L. Alsahlany, "Experimental Analysis of WLAN Security Weakness by Cracking 64 & 128 bit WEP Key," pp. 1-13, Jan. 2014.

[13] F. M. Sun, Z. Zhao, Z. Fang, L. Du, Z. Xu and D. Chen, "A Review of Attacks and Security Protocols for Wireless Sensor Networks," pp. 1-12, May 2014.

[14] S. Zadoo and S. Sinha, "Security Protocol in Wireless Sensor Networks," pp. 1-7, Aug. 2014.

[15] L. K. Musambo and J. Phiri, "Identifying Botnets Intrusion & Prevention – A Review," pp. 1-6, Dec. 2017.

[16] M. Chibuye and J. Phiri, "A Remote Sensor Network using Android Things and Cloud Computing for the Food Reserve Agency in Zambia," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 11, pp. 1-8, 2017.

[17] C. Chembe, R. M. Noor, I. Ahmedy, M. Oche, D. Kunda and C. H. Liu, "Spectrum sensing in cognitive vehicular network: State-of-Art, challenges and open issues, Computer Communications," Vol. 97, pp. 15-30, Jan. 2017. Available. Accessed on: Nov 28, 2018. https://doi.org/10.1016/j.comcom.2016.09.002.

# Technology Paradigm Shift: A Case of Ethical and Unethical Hackers and their Subtle Tools

Raphael Banda[a], Jackson Phiri[b], Mayumbo Nyirenda[c], Monica M. Kabemba[d]
The University of Zambia
Department of Computer Science
Lusaka, Zambia
[a]raphael.banda@yahoo.co.uk, [b]jackson.phiri@cs.unza.zm, [c]mnyirenda@unza.zm, [d]monica.kalumbilo@cs.unza.zm

***Abstract* — Paradigm shift implies change of rules or change of the way we do same things. It is just the way one shifts camp and wear a different hat but still doing the same things to the advantage of the new camp. In doing so some old ways of doing things do not stop working but will just get into one's way. There are three major types of hackers that we can identify although it is not easy to draw a line between them. The three hackers discussed in this paper are black hat, grey hat and white hat hackers. The black hat hackers are the bad people, the grey hat hackers are the intermediate ones somehow on the fence and the white hat hackers are the good one and maybe the ones in charge of systems. The white hat hackers are referred to as the ethical hackers; the black hat hackers the bad people. The black hat hackers can make people's lives difficult and can add an extra bill to the company. We have adopted a systematic review of literature approach to discuss and analyse some of the common tools the black hat hackers have developed to hack into selected systems and commercial software. Why do they do it? For many its mostly to show off their computer skills and prowess or to gain some access to private data. Some of the hacked tools developed elsewhere have found their way to Zambia through the internet and other similar media like local area networks, flash, and CD media and the internet. In Zambia it is illegal to use illegal materials but due to the prices that are normally high for an ordinary Zambian it's difficult to prevent such dark business.**

**Keywords - White hat hacker, grey hacker, black hacker, Kali Linux, Android app, Malware SQLMAP**

## I. INTRODUCTION

People perceive Computer hacking with mixed perception. Our reliance on computer technologies and the critical information shared on networks, the art of computer hacking has been viewed with a lot of scepticism [1]. Some people think that due to expensive software that poor people cannot easily afford. Hackers think that they have a duty to make such software available to the poor at no cost at all. They feel that they have a duty to make such software available to the poor so that the poor can also benefit from the rich companies like Microsoft [2]. Having said that, there is also a "Robin Hood" mentality attached to the practice, where free programs or facilitated measures have been awarded to the average computer user [1].

The primary issue attached to computer hacking stems from an individual's ability to access crucial or personal information that is found on a computer network [3]. The ability to retrieve and subsequently tamper with such information will give way to the potential to commit heinous criminal acts [3].

Internet connectivity is everywhere and it has boosted most businesses. There are very few businesses that can do without the internet. Networks like the internet are the ones at the centre of most businesses worth discussing since stand-alone computer systems cannot do serious business on their own. Computer applications exist in many important sites that can pose a threat to anyone, such as banks, passports general directorate, universities, ministries, emails web hosts, social media sites and many other sensitive country sites [2]. Stand-alone computers have to be networked to other computers to facilitate communication with external world through the internet. However, there is a danger that comes with such connections. The use of computer and internet exposes them to the outside world and hacking [4].

## II. DEFINITIONS OF HACKING

According to New Hacker's Dictionary, a resource used to elucidate upon the art of computer hacking, has defined hacking through a number of definitions [5][6][7]:

Hackers are usually be defined as people who enjoys exploring the complexities of programmable systems and how to stretch their capabilities to some limits of their own interest.

Normally hackers are individuals who possess exceptional skills in computer programming and usage. Some have defined hackers especially the black hat hackers as malicious meddlers who attempt to discover and subsequently tamper with sensitive information through poking around computer based technologies. These individuals are commonly referred to as "network hackers" or "password hackers."

## III. TYPES OF HACKERS

Hackers are classified according to their intentions behind hacking a system. The terms for hacker types are: black hat and white hat hackers. These terms emerged from old spaghetti westerns, where bad guy wears black cowboy hat and good guy wears white.

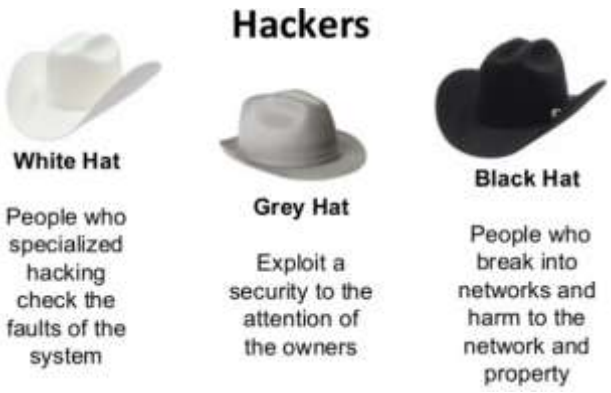



*Figure 1: Types of hackers*

Hackers are classified on the basis of their intentions:

*White hat hackers*
These are Ethical Hackers. They try to find out weaknesses of the computer system or the network with the help of penetration testing and vulnerability assessments. Their main intention of doing so is not to harm the system but to help. A white hat hacker breaks security for non-malicious reasons, perhaps to test their own security system or while working for a security company which makes security software *[3]*.

White hat hacker's job is one of the demanding jobs available in IT industry and its ethical hacking. Numerous companies hire ethical hackers for their system and network security via penetration testing and vulnerability assessments.

*Black hat hackers*
Black Hat hackers or crackers are people who hack the system illegally. When they gain unauthorized access to a system their intentions are to harm its operations or steal sensitive corporate data or secret information. They can also violate privacy block the system network communication, overload the system so that it becomes too slow, etc. A black hat hacker is a hacker who "violates computer security for little reason beyond maliciousness or for personal gain" (Moore, 2005) *[3]*.

*Grey hat hackers*
These are a combination or blend of both black hat and white hat hackers. They act without malicious intent but for their fun, they exploit a security weakness in a computer system or network without the owner's permission or knowledge. The intention behind their work is to bring the weakness to the attention of the owners and getting appreciation or a little bounty from the owners. A grey hat hacker is a combination of a black hat and a white hat hacker *[3]*. A grey hat hacker may surf the internet and hack into a computer system for the sole purpose of notifying the administrator that their system has a security defect, for example *[3]*.

*Red hat hackers*
Red hat hackers usually hack government agencies, top-secret information hubs, and generally anything that falls under the category of sensitive information.

*Blue hat hackers*
Blue hat hackers do not belong to the company but to the outside of computer security consulting firms. Companies and corporates use them to test bugs of the system prior to its launch. They look for loopholes or security holes that can be exploited and try to close these gaps.

A blue hat hacker is someone outside computer security consulting firms who is used to bug test a system prior to its launch, looking for exploits so they can be closed. Microsoft also uses the term Blue Hat to represent a series of security briefing events [3].

Green hat hackers | Newbie | Neophyte
Neophyte, "n00b", or "newbie" or "Green Hat Hacker" is one who is new to hacking or phreaking and has almost no knowledge or experience of the workings of technology and hacking. Newbie hackers desire to be a part of a hacking community even if their individual goals differ *[4]*. Newbie hackers face a constant battle to be accepted by hackers from the beginning of their quests to obtain the capital they need to hack *[4]*. For example, experienced hackers resent script kiddies and unskilled hackers who use hacking tools such as code and scripts developed by experienced hackers *[4]*.

Script Kiddie
Script kiddie is a non-expert who breaks into computer systems by using pre-packaged automated tools written by others, usually with little understanding of the underlying concept.

Hacktivist
A Hacktivist is one who utilizes technology to announce a social, ideological, religious, or political message. Most hacktivism involves website defacement or denial of service attacks.

A hacktivist is a hacker who utilizes technology to announce a social, ideological, religious, or political message. In general, most hacktivism involves website defacement or denial-of-service attacks [3].

## IV. HACKING FEATS

A hacking feat is an application that is programmed to take advantage of a known weakness in the system. Hackers develop tools that are used to perform malicious attacks on computer systems and are usually scripts that are designed to exploit weaknesses in software over a network, most commonly the Internet [3]. The following are some of the most popular techniques:

1. *Attacks*

A typical approach in an attack on Internet-connected system by Discovering information about the intended target, Identifying potential ways of attack and attempting to compromise the system by employing the vulnerabilities found through the vulnerability analysis through the usage of several recurring tools of the trade and techniques used by computer criminals and security experts [3].

*Security Exploits*

A security exploit is a prepared application that takes advantage of a known weakness through exploits of SQL substandard programming practice for example.

*Vulnerability Scanner*

A vulnerability scanner is a tool used to quickly check computers on a network for known weaknesses by checking to see which ports on a specified computer are "open" or available to access the computer, and sometimes will detect what program or service is listening on that port, and its version number.

*Password Scanner*

Password cracking is the process of recovering through illegal or legitimate means passwords for a computer system.

*E. Packet Sniffer*

A packet sniffer is an application that captures data packets, which can later be used to capture passwords and other data in transit over the LAN or Internet.

*Spoofing Attack*

A spoofing attack involves one program, masquerading as another. In doing so its able to falsify data and at the same time being treated as a trusted system by a user or another program. This program cheats or fools other programs or users into revealing confidential information, such as user names and passwords, to the attacker [3].

*Rootkit*

A rootkit is designed to conceal the compromise of a computer's security, and can represent any of a set of programs which work to subvert control of an operating system from its legitimate operators [3]. Usually, a rootkit will obscure its installation and attempt to prevent its removal through a subversion of standard system security [3]. Rootkits may include replacements for system binaries so that it becomes impossible for the legitimate user to detect the presence of the intruder on the system by looking at process tables [3].

*Social Engineering*

When a hacker, typically a black hat, is in the second stage of the targeting process, he or she will typically This is where the hacker or attacker uses some social engineering tactics to get enough information to access the network. For example an attacker will to contact the system administrator and play the role of a user who cannot get access to his or her system [3].

2. *Trojan Horses*

A Trojan horse is a program which enters a computer system as a welcome program but inside the program there is a hidden code that will exploit the computer system at an appropriate time programmed by the hacker. A Trojan can allow the ability to save their files on the user's computer or monitor the user's screen and control his computer [5].

## V. REASONS BEHIND HACKING

There are positive and negative ideas behind performing hacking activity. Reasons behind hacking may be positive or negative intentions. The following is a list of probable reasons why people involve themselves in hacking activities:

- To remove privacy and gain entry to the system
- To break policy compliance
- Just for fun
- To Show-off how powerful and knowledgeable they are in computing
- For money extortion especially in the banks
- To test System security
- To steal information either for sale or for themselves
- To damage or sabotage the system

Hackers hack for reasons such as: conflicts with authorities and revenge motives [2], beliefs that breaking into computer systems benefit society by showing how to increase computer security [3][4][5], achieving feelings of power due to low self-esteem, gaining entrance to a social group and to satisfy *[4]*.

## VI. HACKING WIRELESS TOOLS

Kali Linux operating system is one of the best operating systems hackers normally use to hack computer systems. Most ethical and unethical hackers use Kali Linux to do their respective jobs [6]. Linux is a popular operating system for hackers. There are two main reasons why Kali is preferred by hackers [6]. Linux's source code is freely available because it is an open source operating system and everyone can access it at no cost at all [6]. This means that Linux is very easy to modify or customize. Second, there are countless Linux security distributions (distros) available that can double as Linux hacking software [6].

Wherever you go there is a possibility on one trying to gain access to a wireless hot spot. Wi-Fi are very popular and it's the very popularity and vulnerability that makes it a soft spot for hackers. Wi-Fi is can be seen from any corner of someone's secret place and that is the very reason why hackers find it easy to hack a wireless fidelity or Wi-Fi. hacking wireless networks and how to prevent it from being hacked.

The authors on the Website Hacking Tutorials have discussed a very popular subject of hacking wireless networks and how to prevent it from being hacked. They say that Wi-Fi signals can be picked from anywhere and anytime and this is the very reason why they are vulnerable. Most routers contain vulnerabilities which can be easily exploited with the right equipment and software such as the tools included with Kali Linux. A lot of router manufacturers and ISPs still turn on WPS by default on their routers which makes wireless security and penetration testing even more important.

## VII. HACKING UPC WIRELESS NETWORKS AND OTHER WLAN

The following steps show how to hack UPC wireless networks with the default password which is a common thing for many UPC customers [6].

The first step is to create a password list which contains all possible combinations of 8 capital letters using Maskprocessor in Kali Linux to create the password list. We will then be capturing a 4 way handshake using Airodump-ng by deauthentication of a connected client with Aireplay-ng. The last step is to brute force the password using Aicrack-ng [6].

**Step 1: Creating the password list with Maskprocessor** [6]

Maskprocessor is used to generate the password lists piping each letter to a file so we could use multiple computers to speed up brute forcing the password.

maskprocessor A?u?u?u?u?u?u -o /usr/A.txt
maskprocessor B?u?u?u?u?u?u -o /usr/B.txt
maskprocessor C?u?u?u?u?u?u -o /usr/C.txt
etc…. Repeat for every letter in the alphabet.

The filesize for each document will be approximately 60 GB. You can use the following command to see how many different combinations each file will contain:

maskprocessor A?u?u?u?u?u?u –combinations
8.031.810.176 combinations…
* 26 letters
208.827.064.576 possible combinations

**Step 2: Capturing the handshake with Airodump-ng**

The next thing to do is capture the handshake with Airodump-ng. Airodump-ng was used first to select the target and retrieve it's BSSID and channel the WiFi access point is broadcasting on [6]. Aireplay-ng was used used to de-authenticate a connected client to force a reconnect, which will give us the four way handshake we need [6].

Airodump-ng is started to find the target by using the following command [6]:

airodump-ng mon0

Our BSSID target is picked and channel and restart Airodump-ng with the following command and look for a connected client [6]:

airodump-ng –bssid [BSSID] -c [channel]-w [filepath to store .cap]wlan0mon

New terminal is now opened and a de-authentication command is issued for the connected client using Aireplay-ng as shown in figure 2.

aireplay-ng -0 2 -a [BSSID] -c [Client MAC] mon0


Figure 2: 4 way handshake is captured after deaunthetication *[6]*

Aircrack-ng aireplay-ng

**Step 3: Brute forcing the password with Aircrack-ng**
The following mathematics was used to calculate the time required to crack the password.

The computer specifications used were the 1x AMD hd7970 1000mhz core clock with oclHashcat v1.35 can do 142.000 combinations per second.
26^8 = 208,827,064,576 combinations
26^8 / 142,000 keys per second = 1470613 seconds
2,610,338 / 60 seconds = 24510 minutes
43,505 / 60 minutes = 408,5 hours
725 hours / 24 hours = 17 Days
50% chance of cracking the password in 8.5 days.

It takes 17 days to brute force a standard UPC password and hack UPC wireless networks with a single average videocard using oclHashcat.
It is a network software suite consisting of a detector, packet sniffer, WEP and WPA/WPA2-PSK cracker and analysis tool for wireless LANs.

Aircrack-ng is a network software suite consisting of a detector, packet sniffer, WEP and WPA/WPA2-PSK cracker and analysis tool for wireless LANs [2].

Use the following command to bruteforce the password with Aircrack-ng:
aircrack-ng -a 2 -b [Router BSSID] -w [Filepath to password list] [Filepath to .cap file]
Eventually password will be cracked.

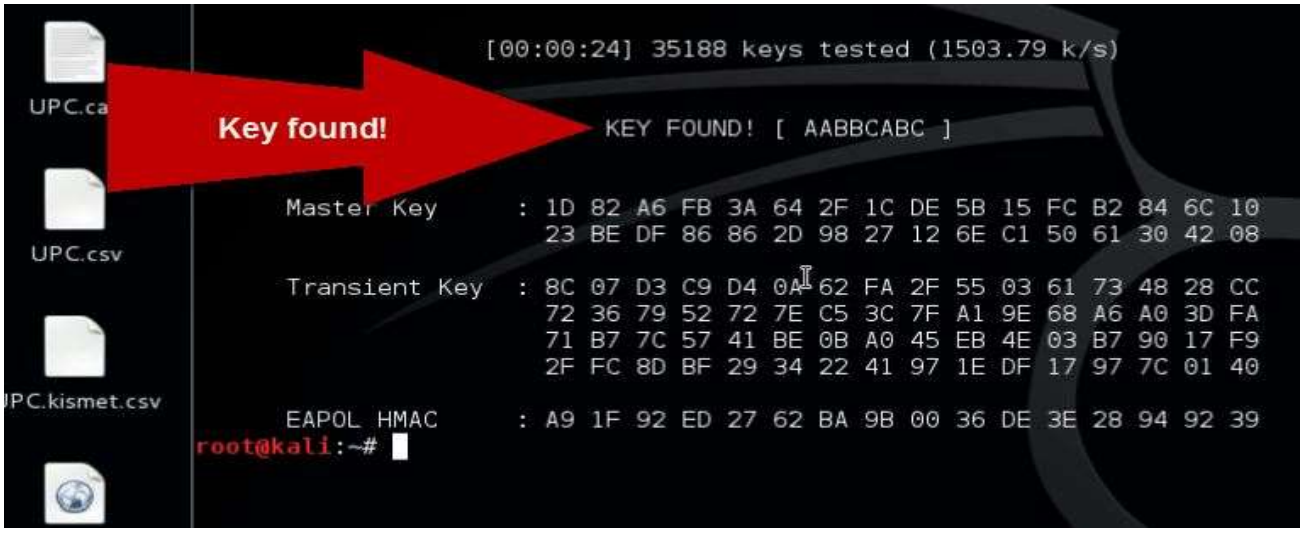



Figure 3: Password cracked as shown *[6]*

## VIII. ANALYSIS

Most people buy fast Fast GPU's at affordable prices for home use. Using such powerful CPU's and GPU's the average home user has the power to crack passwords which are considered strong and safe by many end users [6]. Although 17 days is too long for most to crack a Wifi password it is accessible if you really want to [6]. If you add 3 more letters, or even better, numbers or special characters like a ! or a $-sign it will be close to impossible to crack for an average home user [6].

## IX. SQL Injection

Many companies use databases for keeping data about the company and its employees. It is with little wonder that hackers also can target such databases to crack and gain access to the databases.

SQL injection is a set of SQL commands placed in URL string or in data structures to retrieve a response from the databases connected with the web applications *[6]*. This type of attacks generally takes place on webpages developed using PHP or ASP.NET *[6]*.

There are many reasons why hackers target databases. However the intentions behind SQL injection attack can be as follows *[6]*:

- To dump the whole database of a system,
- To modify the content of the databases,
- To perform different queries that are not allowed by the application.

SQL Injection works when the applications don't validate the inputs properly before passing them to an SQL statement *[6]*. SQL Injections are normally placed in address bars, search fields, or data fields *[6]*.

There are several way of finding whether a web application is vulnerable to an SQL injection attack out but the the easiest way is to use the " ' " character in a string and see if you get any error *[6]*.

## X. The SQLMAP

Data bases normally created using SQL language. It is therefore important that a tool like SQLMAP be used to detect SQL injections. SQLMAP can be downloaded from http://sqlmap.org/

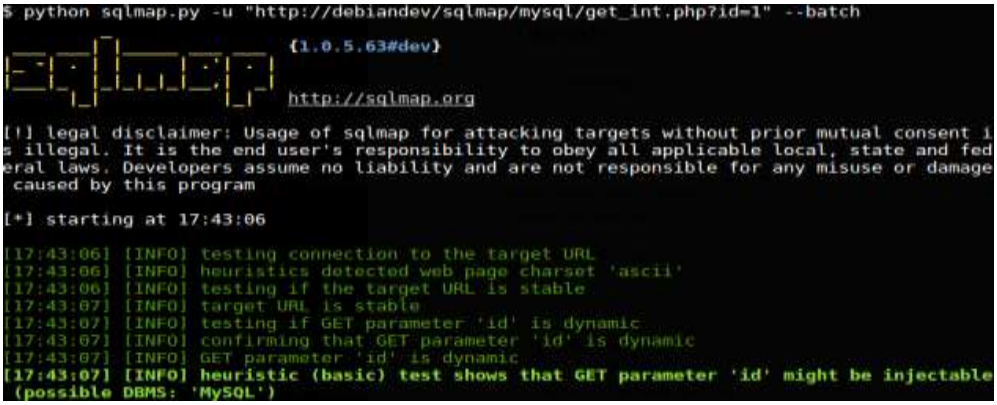



Figure 4: sqlmap is sponsored by Netsparker Web Application Security Scanne *[7]*

It comes pre-compiled in the Kali distro (distribution). It can be located at − Applications → Database Assessment → Sqlmap.

After opening SQLMAP, we go to the page that we have the SQL injection and then get the header request. From the header, we run the following command in SQL −

./sqlmap.py --headers="User-Agent: Mozilla/5.0 (X11; Ubuntu; Linux i686; rv:25.0)
Gecko/20100101 Firefox/25.0" --cookie="security=low; PHPSESSID=oikbs8qcic2omf5gnd09kihsm7" -u '
http://localhost/dvwa/vulnerabilities/sqli_blind/?id=1&Submit=Submit#' -
level=5 risk=3 -p id --suffix="-BR" -v3
The SQLMAP will test all the variables and the result will show that the parameter "id" is vulnerable.

Here are few tips to prevent your web application from SQL injection attacks *[6]*:

Unchecked user-input to database should not be allowed to pass through the application GUI.
Every variable that passes into the application should be sanitized and validated.
The user input which is passed into the database should be quoted.

Penetration Testing
Penetration testing is a method of reducing the risk of security breaches in a system. Most of the companies hire ethical hackers for penetration testing *[6]*. This is the way to find out security breaches and loopholes of a system so that it can be fixed *[6]*.

White hat hackers test penetration in the system. In fact it is legal to test system penetration because it is done with the permission of the owner of the system. Penetration testing is conducted by professional ethical hackers. They mainly use commercial, open-source tools, automate tools and manual checks are normally used to test penetration. There are no restrictions for their work. The only objective here is to reveal as many security flaws as possible and create necessary interventions to make the system stronger and better *[6]*.

Penetration testing can also cause problems such as system malfunctioning, system crashing, or data loss. Therefore, a company should take calculated risks before going ahead with penetration testing. The risk is calculated as follows and it is a management risk.

RISK = Threat × Vulnerability

## XI. WORKING WITH JOHN THE RIPPER

John the Ripper is probably the world's best known password cracking tool. John the Ripper is one of the most popular password cracking tools available that can run on Windows, Linux and Mac OS X [8]. Its lack of a GUI (Graphical User Interface) makes it a bit more challenging to use.

**How to Crack Windows Password with John the Ripper**

There are many reasons why one would want to hack a Windows password. For example if you have forgotten the password to your Windows admin account you definitely need to recover it to have access to your account. The following steps show how to use John the Ripper to crack Windows 10, 8 and 7 password on your own PC. Emphasis "Your own PC".

Step 1: Extract Hashes from Windows

Security Account Manager (SAM) is a database file in Windows 10/8/7/XP that stores user passwords in encrypted form, which could be located in the following directory:

*C:\Windows\system32\config*

Grab the password hashes from the SAM file as the first thing you need to do. Just download the freeware **PwDump7** and unzip it on your local PC.

Open a Command Prompt. Navigate to the folder where you extract the PwDump7 app, and then type the following on the command line:

*PwDump7.exe > d:\hash.txt*

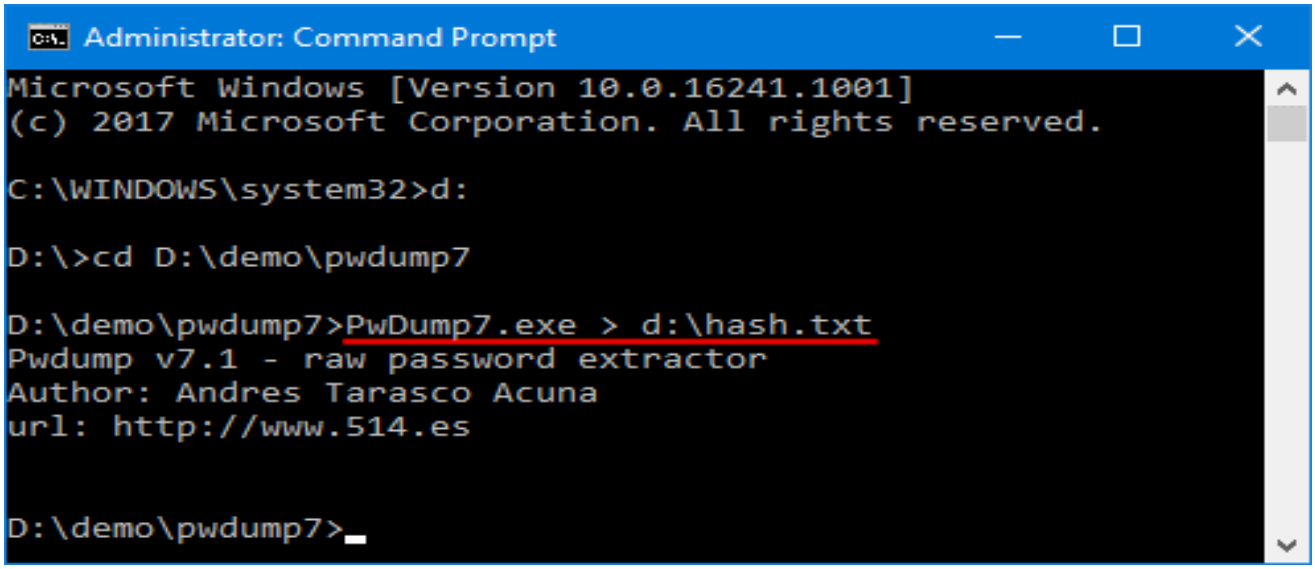



*Figure 5: PwDump7.exe > d:\hash.ext [8]*

*extract-windows-password-hashes*

Hit Enter, PwDump7 will grab the password hashes from your current system and save it into the file d:\hash.txt.

Step 2: Cracking Passwords with John the Ripper

In this stage the password hashes are still unreadable. This is the time to crack them using John the Ripper.

Just download the Windows binaries of John the Ripper, and unzip it.

Open a Command Prompt and change into the directory where John the Ripper is located, then type:

*john --format=LM d:\hash.txt*

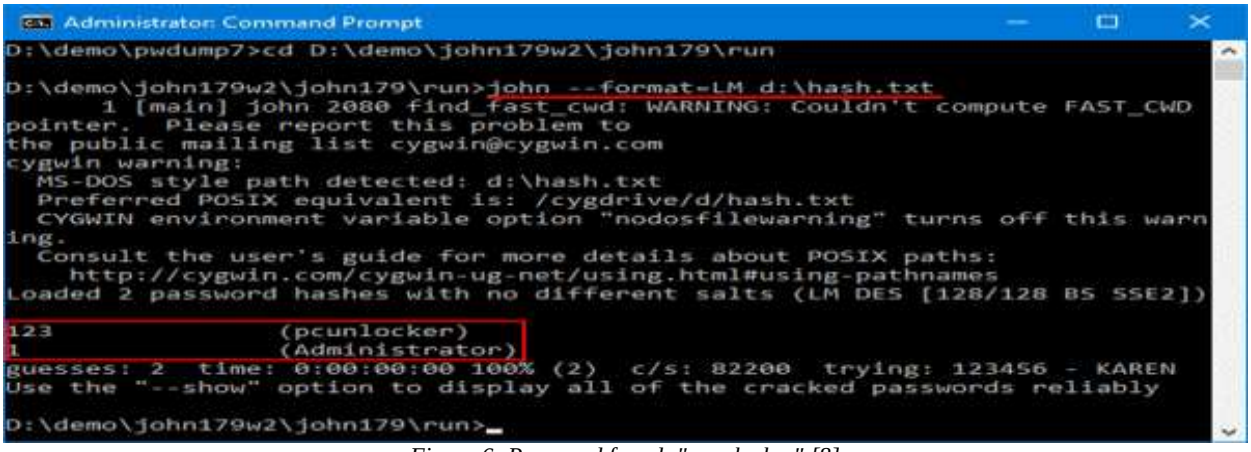



*Figure 6: Password found: "pcunlocker" [8]*

It will start cracking your Windows password. In this example John the Ripper has cracked the password within matter of seconds as time indicates 0:00:00:00 [8].

## XII. HACKED TOOLS

**Crack for Office 2013/ 2016**

Microsoft Office 2013/ 2016 is one of the most hacked software in the World. The hacking is due to the fact that its one of the most used software on a PC. About ¾ of the world population use the software for word processing. The following are some of the hacked tools that I download ed from using UTorrent. They are illegal but can be used to do the job.

*Figure 7: UTorrent downloading*

Steps on how the crack of Office 2013/2016 can be installed are as follows:

Step 1; Install the software as shown below:


*Figure 8: Starting to install MS Office 2013/2016*

Figure 2: Starting the Office installation
Step: You may need to choose to install the ones you need from the selection. Ticked ones are the ones to be installed.

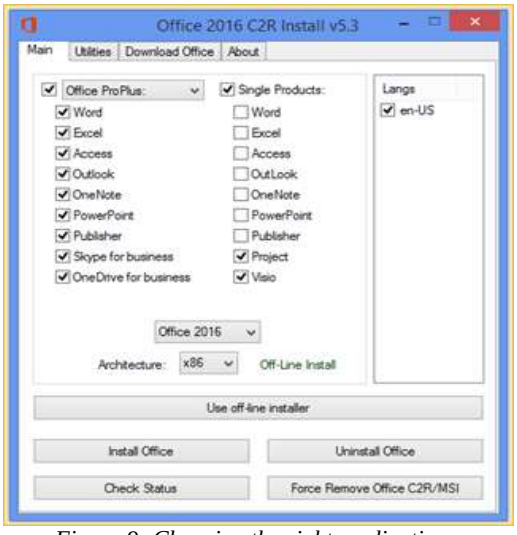



*Figure 9: Choosing the right applications*

Activating the Office once it has been installed is easy. Just follow the GUI shown below and you have a fully working version of Office 2016 or Office 2013 depending on the installed choice.


*Figure 10: Smart Utilities includes activation button*

The tool can also assist to download Microsoft 2013 or Microsoft 2016 office software that can be cracked using the same tool.


*Figure 11: GUI for downloads for MS Office 2013/ 2016*

## XIII. Analysis

The crack that could have been used in creating the crack for MS Office above is the root kit.

A rootkit is designed in such a way as to conceal the compromise of software's security. Usually, a rootkit will obscure its installation and attempt to prevent its removal through a subversion of standard system security [3]. Rootkits may include replacements for system binaries so that it becomes impossible for the

owner to detect the presence of the intruder on their own software by analysing process tables [3]. The hacked software can work normally for years without being detected by the owner even though the internet. In the past software owners used to detect their cracked software and disabled them from running on the machines nowadays hackers have become more complicated and the hacked software are not that easy to detect even by the developers. The advancement of the hacking that is going on now leaves a lot as to wonder who is behind the scourge. Other forms in which this type of intrusion or threat may be created if an authorised individual who has sufficient rights and privileges to access organisational resources, decides to deliberately destabilize the organisational networks or computing resources by disabling security features of the network so as to allow harmful applications such as botnets into the organisation [6].

## XIV. CONCLUSION

For years, IT professionals have built barriers to prevent hackers' entry that could compromise the organisations networks and systems. [9]. However no matter how good these barriers are the hackers have developed even more advanced tools to access or attack the systems. Systems and networks will be compromised, almost always regardless of what network engineers do [5]. Different institutions like banks have also fallen victims of hacking no matter how high they have set the bar of security. Security protocols development would help in safe guarding customer information and should be made sophisticated for the hackers and personalized to each user of the e-banking facilities and tough penalties should be imposed on the wrong does who would want to misuse technology to still customer information as argued by Hamidi et al., 2013 [10]. Another institution thathas tried to keep its security for its patience is the health. Health researches in Zambia have also tried to raise the level of confidentiality of its patient to the maximum. They have suggested the addition of pin numbers for smart cards and staff access cards with passwords have improved security of the smart-care program in Zambia although some scholars have advocated for the inclusion of encryption as a key security feature to prevent hackers [11].

Software development can be very expensive but in most cases, the returns are also good. That is why Microsoft and Apples are part of the most expensive and rich companies in the world now. Their software are also expensive and there are poor people who cannot afford them. Hackers especially black hat hackers think they are duty bound to make such expensive available to all at no price at all. They feel they should help the poor by making such expensive software available to them at not cost. They want to play the Robin Hood style of the old days where Robin Hood stole from the rich to help the poor. Robin himself was not poor but he thought he should steal to help the poor because corporates are stingy and do not want to share.

Some hackers hack the system just for fun and to show off that they control computers. They make the systems vulnerable and for them that all and feel good that someone is in trouble because of their own actions.

In Zambia there are a lot of pirated or illegal software on peoples PCs but the numbers are just too small to raise the alarm siren. People install and uninstall illegal software created by hackers at will.

Illegal software will continue to find market so long the owners of such software continue making such software expensive.

Companies and corporates need to spend more to protect their data from hackers but computer scientists are still not their favourite candidates to get a better pay. Its for such reasons that sometimes IT personnel deliberately hack the system just to show that they are the owners of such companies and without them they are nothing. Professionals and corporate should understand that business is directly related to software and hardware security.

Sometimes it is difficult to draw a line between different types of hackers. Companies for not taking care of them could have created black hat hackers. Today one is a white hacker tomorrow he is a black hacker for the system that he knows better. Sometimes its all about taking good care of each other.

There are a lot of initiatives nowadays that aim to enhance security of data. One such method of enhancing the security of data is that of the cloud. Some scholars and experts in the field have contended that this is also not secure. The owners of the cloud service system have literal control over the data they hold on behalf of their customers and this may imply that they can modify data to their liking without the data owner's consent [12]. This act amounts to insider threats [12].

The paradigm shifts for hackers is very difficult to predict. Sometimes outsiders or insiders instigate their roles and their movement from one discipline to another is always unpredictable.

## REFERENCES

[1] Laws.com, "Hacking," Laws.com, 2017. [Online]. Available: https://cyber.laws.com/hacking. [Accessed 03 11 2017].

[2] M. A. Ghanem, "BackTrack System: Security against Hacking,"

International Journal of Scientific and Research Publications, vol. 5, no. 2, p. 4, 2015.

[3] S. Goel, K. Gupta, M. Garg and M. A. K, "Ethical Hacking and Its Countermeasures," International Journal of Advance Research and Innovation, vol. 2, no. 3, pp. 624-629, 2014.

[4] M. Nycyk, "The New Computer Hacker's Quest and Contest with the Experienced Hackers: A Qualitative Study applying Pierre Bourdieu's Field Theory," International Journal of Cyber Criminology, vol. 10, no. 2, p. 93, 2016.

[5] Jackson Phiri, Tiejun Zhao, Hua. C. Zhu and J. Mbale, "Using Artificial Intelligence Techniques to Implement a Multifactor Authentication System,"International Journal of Computational Intelligence Systems, vol. 4, no. 4, pp. 420-430, 2011.

[6] Hacking Tutorials, "Hacking-tutorials," Hacking Tutorials, 24 May 2015. [Online]. Available: https://www.hackingtutorials.org/wifi-hacking-tutorials/how-to-hack-upc-wireless-networks/. [Accessed 03 11 2018].

[7] G. Kambourakis, F. G. Marmol and G. Wang, "Security and Privacy inWireless and," Future Internet, vol. 10, no. 3390, p. 1, 2018.

[8] Admin, "Password Recovery," Top Password.com, 7 August 2017. [Online]. Available: https://www.top-password.com/blog/tag/how-to-use-john-the-ripper/. [Accessed 4 11 2018].

[9] K. I-Lung, "Securing mobile devices in the business environment," IBM Global Technology Services; Thought Leadership White Paper, pp. 2-10, October 2011.

[10] A. Nuwagaba and B. Ngoma, "Analysis of E-Banking as a Tool to Improve Banking Services in Zambia," International Journal of Business and Management Invention, vol. 3, no. 11, pp. 62-66, 2014.

[11] K. Mwebo, "Security of electronic health records in a resource limited setting: The case of smart-care electronic health record in Zambia," in Australian eHealth Informatics and Security, Perth, 2014.

[12] M. K. P. J. Chinyemba, "Gaps in the Management and Use ofBiometric Data: A Case of Zambian Public andPrivate Institutions," ZAMBIA INFORMATION COMMUNICATION TECHNOLOGY (ICT) JOUNAL, vol. 2, no. 1, p. 37, 2018.

[13] A. Parkar, S. Sharma and S. Yadav, "Introduction to Deep Web," International Research Journal of Engineering and Technology (IRJET) e, vol. 4, no. 6, pp. 1-4, 2017.

[14] K. Christian, B. Katja, T. Markus, H. Stephan and R. Kai, "How to Enhance Privacy and Identity Management for Mobile Communities: Approach and User Driven Concepts of the PICOS Project," Mobile Business & Multilateral Security, 2010.

[15] R. Bhasker and B. Kapoor, "Information Technology Security Management," in Computer and Information Security Handbook, Burlingtone, Morgan Kaufmann Publishers is an imprint of Elsevier, 2009, pp. 259 -267.

[16] M. Ahmad and J. Parvez, "A Novel Strategy to Enhance the Android Security Framework," International Journal of Computer Applications (0975 – 8887), vol. 91, no. 8, pp. 1-5, 2014.

[17] R. Bhasker and B. Kapoor, "Computer and Information Security Hand Book," in Computer and Information and System Security, Burlington, Morgan Kaufmann Publishers is an imprint of Elsevier, 2009, pp. 259-257.

[18] K. Kathirvel, "Credit Card Frauds and Measures to Detect and Prevent Them," International Journal of Marketing, Financial Services & Management Research, vol. 2, no. 3, pp. 1-8, 2013.

[19] M. S. Gaigole and M. A. Kalyankar, "The Study of Network Security with Its Penetrating Attacks and Possible Security Mechanisms," International Journal of Computer Science and Mobile Computing, vol. 4, no. 5, p. 729, 2015.

[20] M. B. Mollaha, M. A. K. Azada and A. Vasilakosb, "Security and privacy challenges in mobile cloud computing: Survey and way ahead," Journal of Network and Computer Applications, vol. 84, p. 38, 2017.

[21] MSA Technosoft, "Hacking | Types | Purpose | Hackers | SQL | Injection | SQLMAP | Penetration Testing," MSA Technosoft, 30 May 2018. [Online]. Available: https://msatechnosoft.in/blog/tech-blogs/hacking-types-purpose-hackers-sql-injection-sqlmap-penetration-testing?fbclid=IwAR0lfQ51CGPDOPLiWY5Z0eNHP8SkSC-c65qQxlQrkrKvqWPDWy6_Dt_qrKc. [Accessed 3 November 2018].

[22] D. Bernardo and G. Assumpcao, "Automatic SQL injection and database takeover tool," Sqlmap, 2018. [Online]. Available: http://sqlmap.org/. [Accessed 3 11 2018].

# A Systematic Literature Review of Big Data Analytics Implementation in Health Care

Richard C. Chellah
*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
richard.chellah@cbu.ac.zm

Prof. Douglas Kunda
*School of Science, Engineering and Technology*
*Mulungushi University*
*Kabwe, Zambia*
dkunda@mu.edu.zm

***Abstract*-Big Data analytics implementation and use especially in industries like healthcare has huge potential for improving the quality of care, reducing waste and error, and reducing the cost of care. This systematic review of Big Data Analytics literature aims to determine the opportunity of Big Data analytics in healthcare including its challenges in relation to adoption and implementation and ways to overcome any such challenges in healthcare. Articles were systematically searched using ScienceDirect, PubMed, IEEEXplore and ResearchGate databases. To be considered for review, articles shortlisted were on Big Data analytics in healthcare published from January 2008 to June 2018. The review selected descriptive articles and usability studies of Big Data analytics in healthcare and medicine.**

**The analyses of these articles found that: researchers lack consensus about the operational definition of Big Data in healthcare; Big Data in healthcare comes from the internal sources within the hospitals or clinics as well external sources including government, laboratories, data aggregators, medical journals etc.; Big Data analytics finds its application for clinical decision support; optimization of clinical operations and reduction of cost of care. The major challenge in adoption of Big Data analytics is non-availability of evidence of its practical benefits in healthcare. This review study unveils that there is a little of information on evidence of real-world use of Big Data analytics in healthcare. This is can be owed to the fact that, the usability studies have considered only qualitative approach which describes potential benefits but does not take into account the quantitative study. It should also be noted that, the majority of the studies were from developed countries which brings out the need for promotion of research on Healthcare Big Data analytics in developing countries.**

***Key words: Big data, big data analytics, Health care, Implementation***

## I. INTRODUCTION

There has been a rapid digitalization across the industries over the few years. The global health sector has also undergone this digital transformation with an increase in use of Electronic Medical Records (EMRs); Healthcare Information Systems (HIS); and handheld, wearable and smart devices. This has resulted in massive amount and variety of health-related data today which is in digital form. This high-quality healthcare data offers potential value for improving care delivery, but it is still not being adopted in most health sectors around the world because most decision markers do not see it as an asset that can give then competitive advantages. Since most electronic health data remains largely underutilized and hence wasted, there is a need for converting the raw data in healthcare into meaningful and actionable information [1]–[3].

Most of the highly valuable data in healthcare is in unstructured or were possible it is semi-structured form. Added to it, the complex, dynamic and heterogeneous characteristics of the data [4]–[6] makes it difficult for health care information specialists to extract useful information using traditional data analytical tools and techniques [7]. This has created the need for integration of Big Data analytics into healthcare in order to gain its benefits. Big Data analytics has the ability to analyze a wide variety of complex data and generate valuable insights which otherwise would not have been possible. Big data in healthcare has the potential to identify patterns and lead to improved healthcare quality and reduced costs and enable timely decision-making [5], [7]–[9].

With the promising value of Big Data technology in other industries such as banking and education it created an increasing interest of academic and industry research for Big Data in health. There have been only a few health care literature reviews and the literature remains largely fragmented. This review's purpose is to gain a comprehensive understanding of current position on this technology. It aims at answering the research question on: How "Big Data analytics" is or can be used in healthcare to enhance its value? It also reviews the big data challenges in healthcare and some of the ways they can be solved.

## II. REVIEW METHOD

A systematic review was carried out in order to capture relevant literature from different sources, with the focus being on the following objectives:

- To determine different perspectives to definition and concepts of Big Data in general and also in healthcare.
- To explore and review the sources of Big Data in health.
- To identify the techniques and technologies for big data analytics in healthcare.
- To show the potential benefits and applications of Big Data within healthcare.
- To explore the implementation or application challenges of big data in health.
- To present strategies for tackling the challenges of Big Data application within healthcare

By reviewing these objectives, this review will make a significant contribution in understanding the overall context and the future application of Big Data techniques in the healthcare domain.

### *A. Information sources*

A search for articles was made on following databases: ScienceDirect, PubMed, IEEEXplore and Research Gate. The references included in these articles were also scanned for a thorough review

### *B. Selection criteria*

To select the literature for inclusion in the literature review, following inclusion criteria were used:

- Inclusion Criteria 1: Articles that are on Big Data analytics in healthcare
- Inclusion Criteria 2: Articles published between 2011 and 2018
- Inclusion Criteria 3: Only articles published in English

To capture the literature relevant to the research interest, the articles with their primary emphasis on the traditional analytics in healthcare were excluded.

### *C. Study selection*

The procedure for search and selection of research material was carried out in the following phases:

1. The search for publications on electronic databases containing the keywords “big data” or “big data analytics”, and “Big data analytics in healthcare”
2. Scrutiny of articles that are published in English.
3. Scrutiny of the title, abstract and keywords of identified articles and selection of the significant articles on the basis of selection criteria
4. Perusal of articles that were not eliminated in the previous phase for the review
5. Scanning of cross-reference articles for detailed study

### *D. Quality assessment*

Activities ensuring the quality of the search process were undertaken during the review. All the web searches for the review were made in incognito mode to avoid any external influence of historical searches. From initial searches, the authors manually extracted relevant papers and articles. The analysis and evaluation of abstracts were carried out and this helped with articles were to be included or excluded from the study

## III. RESULTS

The study selection heighted in this review was followed for each of the databases. The literature included in this review comprises mainly of descriptive articles and usability studies. With the basis on the main research objectives, the content from these articles was extracted and the articles were organized into different groups. The following section summarizes the findings in each of these categories.

### *A. Concept of big data and its definitional perspectives*

Big Data Analytics (BDA) is has been increasingly becoming a trending practice that is generating an enormous amount of data and provides a new opportunity that is helpful in relevant decision making. The recent developments in Big Data Analytics has provided a new paradigm and solutions for big data sources, storage, and advanced analytics. The BDA provides a new view of big data development, and insights on how it can truly create value for firm and customer. The BDA helps in acquiring a deep understanding and useful insights of various sectors such as: agriculture, healthcare, cyber physical system, smart cities and social media analytics etc. The enormous amount of information is needed to analyze it in an iterative way and time sensitive manner [8].

The concept of Big Data came up in late 1990s when Michael Cox and David Ellsworth [10] considered visualization as a problem of Big Data. One of the early Big Data definitions was given by Francis X. Diebold [11] in 2000 when he referred to Big Data as “explosion in the quantity (and sometimes, quality) of available and potentially relevant data”. Later the key dimensions of Big Data – volume, velocity and variety: the 3Vs – were derived from the study of Doug Laney in 2001 [12]. Manyika et al. [13], highlighted in their report, that value was another important element of Big Data. In terms of Big Data in healthcare, Feldman et al. [14] introduced veracity as yet another critical feature. Several definitions of Big Data currently exist based on the characteristics of Big Data.

In today’s world, data is considered a powerful raw material that can impact multidisciplinary research endeavors as well as government performance. This is an era of big data, data sets that are characterized by high volume, velocity, variety, resolution and indexicality, rationality and flexibility. This data becomes the very important source for valuable insights and ultimately helps to make more precise decisions [15]. Avoiding the mention of Big Data anywhere we turn today is a bit hard. There is broad recognition of the value of data, and products obtained through analyzing it [16]. Big data is defined in different ways by different researchers; this is usually based on its characteristics of velocity, volume, value, variety, and veracity. To sum it up

Big Data can be defined as massive data sets having large, more varied and complex structure with the difficulties of storing, analyzing and visualizing for further processes or results. The process of research into massive amounts of data to reveal hidden patterns and secret corrections is called Big Data Analytics [15], [17]–[24] . Big Data is one of the current and future research frontiers in healthcare. It is right to say that Big Data will revolutionize many fields, including business, the scientific research, public administration, and so on [21].

In healthcare, "Big Data includes mixed, incomplete and imprecise observations (e.g., diagnosis, demographics, treatment, prevention of disease, illness, injury, and physical and mental impairments) derived from different sources using incongruent sampling" [25]. Some of these data are structured and they focus on genotype, phenotype, genomics data, ICD codes [2], [6], [26]; but the unstructured data includes memos, clinical notes, prescriptions, medical imaging, EHRs, lifestyle, environmental, and health economics data [2], [6], [26], [27]. The challenge for Big Data analytics is to deal with this mixed data in order to generate insights for improved healthcare outcomes.

By analyzing the literature, it is evident that although the significance of big data in strengthening healthcare is recognized and understood, there is still a lack of consensus on the operational definition of big data in healthcare. Therefore, the review of definitions from previous studies allows discernment of the common elements.

*B. Sources of healthcare big data*

The health care sector has grown tremendously in last few decades. It has generated huge amounts of data that has huge volume, enormous velocity and vast variety. Also, it comes from a variety of new sources as some hospitals now tend to implement electronic health record (EHR) systems. These sources have strained the existing capabilities of existing conventional relational database management systems. In such scenario, big data solutions offer to harness these massive, heterogeneous and complex data sets to obtain more meaningful and knowledgeable information [24], [28].

Data in healthcare are disorganized and distributed, coming from various sources and having different structures and forms [29]. Healthcare Big Data includes data on physiological, behavioral, molecular, clinical, environmental exposure, medical imaging, disease management, medication prescription history, nutrition, or exercise parameters [26]. Some of the primary sources of Big Data in healthcare are administrative databases (insurance claims and pharmaceuticals), clinical databases, electronic health record data [9], and laboratory information system data [30]. The other sources of data [9] are biometric data (wearable or sensor generated [31], patient-reported data (standardized health surveys), data from social media [31], medical imaging data, and biomarker data.

According to Belle et al. [1], healthcare data is spread among different healthcare systems, health insurers, researchers, government entities. Huang et al. [32] recognizes that Big Data in precision medicine comes from four different stakeholders. Government and large companies, Smaller stakeholders (such as academic groups and technology, biotech, and device startups), Health care providers and payers, and Not-for-profit foundations and patient advocacy groups. Clinical data such as past medical history, vital signs, medications, immunizations and medical imaging can be derived from electronic health records, CPOE, clinical decision support systems medication administration records, laboratory and pharmaceutical records [8], [33], cohort studies, government surveys & clinical trials [10]. Administrative data, on the other hand, contains patient demographic data and visit information, admit date, discharge date, ICD diagnosis & procedure codes, admit source, discharge disposition and claims data such as charges for the visit, payer and reimbursement [33]. The table below summarizes some of the sources of big data in the health care.

| Type | Description | More Details | Source |
|---|---|---|---|
| Clinical | Electronic Medical Records (EMRs) | Detailed patient-related information (physician prescriptions, medications, medical history | Hospitals and clinics |
| | Diagnostic | Diagnostic results (imaging results, laboratory reports) | Laboratory<br>Radiology Departments |
| | Biomarkers | Molecular data (genomic, proteomic, transcriptomic, metabolomic) | Diagnostic companies |
| | Ancillary | Administrative data (admission, discharge, transfer) and financial data (claims) | Hospitals & Clinics<br>Data Aggregators |
| Claims | Medical Claims | Medical reimbursement data (procedures, hospital stay, insurance policy details) | Payers<br>Data Aggregators |
| | Prescription Claims | Prescription reimbursement data (drugs, dose, duration) | Payers<br>Data Aggregators |
| Clinical Research | Clinical Trails | Design parameters (compound, size, end points) | Pharma/Companies<br>Medical Journals |
| Patient-generated Data | Social Media | Community discussions | Web Health Portals<br>Social Media Websites |
| | Wearable & Sensors | Wellness & lifestyle data (smartphones, fitness monitors) | Device Data Systems |

Table 1: Sources of healthcare data (adapted from [59])

*C. Big data analytical techniques and technologies in healthcare*

The multi-dimensional healthcare data – medical images, biomedical signals, audio transcripts, handwritten prescriptions and structured data from EMRs [34] −and its dynamicity and complexity makes it difficult to analyze them. There is paucity of analytical strategies that can handle such heterogeneous data and facilitate decision-making [35]. The literature mentions some of the analytical approaches which can apply to healthcare and medicine.

As described by Asante-Korang & Jacobs [2] and Groves et al. [36], by incorporating descriptive and comparative analytics, healthcare organizations have seen improved quality of care. Nevertheless, they have stated that the long-term tangible benefits can be accrued with utilization of predictive analytics. As highlighted by the literature, predictive analytics can be used for prediction of high-cost patients, readmissions, triage, decompensation (when a patient's condition worsens), adverse events, and treatment optimization for diseases affecting multiple organ system [27], [37]–[39].

MapReduce framework has been used by Markonis et al. [40] for finding optimal parameters for lung texture

classification and to increase the speed of medical image processing. Peek et al. [41], in their study discussed about some of the Hadoop based Big Data processing tools which can be used for batch processing; and non-Hadoop processing tools like GraphLab which can be used for streaming data analysis. Regardless of these potential applications, there is a need for analytical tools to offer parallelization, in order to enable the timely processing of data [4].

*D. Implementation of big data analytics in healthcare*

Big Data analytics has the potential to transform business and clinical models for smart and efficient delivery of care [42]. It enables integration of de-identified health information to allow secondary uses of data [43]. Also, by recognizing patterns and deciphering associations it can facilitate autonomous-decision making [2]. In clinical practice, Big Data analytics can help early detection of disease, accurate prediction of disease trajectory, and identification of deviation from healthy state, changed disease trajectories and detection of fraud. By providing this information, it helps the healthcare organizations in personalization of predictions, targeted-treatment and cost-effectiveness of care, and reduction in waste of resources; and by giving actionable recommendations to individuals it encourages them maintain themselves in good health [8], [26], [37]. Big Data in healthcare presents an opportunity to detect relatively low-frequency events that nonetheless can have significant clinical impact. Apart from that, clinical data integration and its effective usage support a vast range of applications, such as disease surveillance, clinical decision support systems, and individual healthcare management; improvement of health-process efficiency; enhancement of healthcare quality and reduction of healthcare cost [7], [44]. Sukumar et al. [42].

Belle et al. [1] identified three major areas for the application of Big Data analytics in Healthcare: Image Processing, Signal Processing and Genomics. Rumsfeld et al. [9] highlights eight areas of application of Big data analytics to improve healthcare and these include: predictive modelling for risk and resource use; population management; drug and medical device safety surveillance; disease and treatment heterogeneity; public health; precision medicine and clinical decision support; quality of care and performance measurement; and research applications. Raghupathi & Raghupathi [8] state that, 'the areas in which advanced analytical techniques yield the greatest results include: pinpointing patients who are the greatest consumers of health resources and or at the risk for opposing outcomes;; identifying programs, treatments and processes that do not deliver demonstrable benefits or cost too much; reducing readmissions by identifying environmental or lifestyle factors that increase risk and or trigger adverse events and adjusting treatment plans accordingly; improving outcomes by examining vitals from at-home health monitors; managing population health by detecting vulnerabilities within patient populations during disease outbreaks or disasters; providing individuals with the information they need to make informed decisions and more effectively manage their own health as well as more easily embrace and track healthier behaviors and bringing clinical, financial and operational data together to analyze resource utilization productively and in real time'. Electronic phenotyping is another area which can successfully exploit Big Data technologies for ascertaining a clinical condition or characteristic (phenotype) [45], [46]. These studies show that there is a vast potential of Big Data analytics in Healthcare.

Apart from these clinical benefits and applications mentioned above, the literature also presents operational and financial benefits of Big Data analytics. From among all the articles examined, some other articles [5], [27], [47] unveil the business value in healthcare. Findings from the study of Wang & Hajli [5], exhibits that benefits of Big Data analytics are improved IT effectiveness and efficiency, and optimization of clinical operations.

*E. Challenges in big data analytics in healthcare*

In as much as there are huge and potential benefits, the healthcare industry is in its nascent stage for adoption of Big Data analytics. With the huge amount of data available, there is a lack of knowledge about which data to use and for what purpose [48]. Another major challenge that healthcare faces is the lack of appropriate IT infrastructure [7], [27], [49] and transition from use of paper-based records to use of distributed data processing [41], [50]. The resistance for redesigning processes and approving technology that influences the health care system [30], [47] and need for huge initial investment [48], [51], makes it more difficult to utilize Big Data technology. Studies show that because of the lack of knowledge about the best algorithm and tool for analysis [52] and unavailability of trained clinical scientists and Big Data managers for interpretation of Big Data outcomes [2], [26], [49], [51], healthcare remains far from realizing the potential of Big Data analytics. A major concern with the use of Big Data analytics in Healthcare is the processing of information without human supervision which might lead to erroneous conclusions [53], [54]. According to Raghupathi and Raghupathi [8], there is a need for a simple, convenient and transparent Big Data analytics system which can be applied for real-time cases.

Technical challenges include integration of structured, semi-structured and unstructured data from a variety of resources [26]. Studies show that the main technical issues in Big Data analytics include siloed/fragmented data [8], [34], [48], limitations of observational data [9], [32], validation [9], data structure issues, data standardization issues [8], [32], [55]–[57], data inaccuracy and inconsistency (veracity) [10], [42], [48], [57], data reliability [58], semantic interoperability [25], [41], [58], network bandwidth, scalability, and cost [1]. The problems such as missing data and the risk of false-positive associations [9], [59] also add to it. Security issues such as Big Data breaches can be significant threat in healthcare [2], [31].

It should be noted that patient privacy and confidentiality are of utmost importance in healthcare. But data sharing between various stakeholders for deriving insights, can deepen the concern for privacy [7], [30], [32], [47], [51], [54]. According to Mittelstadt et al. [51] informed consent and privacy are the key areas of concern. Lack of data protocols and standards are some of the governance issues faced by Big Data analytics in healthcare [1].

*F. Challenges specific to developing countries*

In as much as they are many benefits expected with the implementation of Big Data projects in all areas, there are difficulties common to all of them, regardless of whether they are conducted in developing or developed countries. Developing countries and their healthcare industries in particular, have unique characteristics that merit special analysis on the challenges faced by the application of big data and the ways they can be surmounted. Six broad challenges faced in big data implementation are presented below as review in different papers and these are; data capture, Infrastructure, organizational changes, integration and interoperability, privacy and security, and adoption challenges [15], [60], [61].

Given the growing importance of big data and the potential benefits of its use in the healthcare; there is great need for developing countries to embrace it so as to solve most of the healthcare challenges they are facing. Despite increasing awareness of the benefits of big data and the related methodological and technological advances that are being made, many countries appear to be slow in adopting approaches based on such data. The reasons may include gaps in funding, leadership and technical expertise and competing priorities within the health system. Many governments are still considering appropriate policy options [23], [62]. The table below summarizes challenges specific to developing countries.

| Challenge | More specific details | Studies by |
|---|---|---|
| Data | Because of huge data, it is hard for organizations to capture, store, manage and analyze data in a timely manner | [15], [60], [61] |
| Infrastructure | Lack of infrastructure, software, trained workforce | [15], [60], [61] |
| Organizational changes. | Scarce human resource in the field. | [15], [60], [61] |
| Integration and interoperability | Lack of standards to achieve Integration and interoperability. | [15], [60], [61] |
| Privacy and security | Lack of safety measures in the software used in the big data implementations | [15], [60], [61] |
| Adoption | | [15], [60], [61] |

Table 2: Challenges specific to developing countries

*G. Strategies to overcome Big Data Challenges*

In order to overcome the highlighted big data challenges, various strategies were found in the literature. The strategies for curbing the aforementioned issues include:

- Implementing (big) data governance: Due to poor governance, healthcare organizations incur huge financial costs in IT investment [14]. With appropriate data governance, the enterprise-wide data resources can be leveraged effectively to create business value [27], [51].
- To develop an information sharing culture: Information sharing and aggregation of data can address the issue of interoperability and enable effective utilization of the Big Data analytical and predictive capabilities [27], [34].
- Employing security measures: Strong encryption of data, validation of source of data, access control and authentication [63] and de-identification [64] are some of the measures for securing the data and maintaining confidentiality.
- To train key personnel to use Big Data analytics: In order to extract meaningful insights and valuable information from Big Data, healthcare professionals should be trained with Big Data analytics competencies. This is critical for healthcare, because incorrect interpretation of the reports generated could lead to unanticipated consequences [27].
- Incorporating cloud computing into the organization's Big Data analytics: The challenge of storage of voluminous data can be tackled by making use of cloud computing. Doing so would enable small and medium sized hospitals and care organizations to eliminate cost and data storage issues [27].

According to Wang et al. [27], a shift of focus from technology tools to the managerial, economic, and strategic impacts of Big Data analytics and exploration of effective path for acquiring healthcare business value would enable realizing the benefits of Big Data analytics.

## V. Discussion

*A. Main findings*

The systematic review assessed the new trend of Big Data analytics for healthcare. Specifically, it identified the best available literature about the concept of Big Data analytics, sources of Big Health Data, Big Data analytical techniques for clinical data, implementation of Big Data analytics in healthcare, the causes for underutilization of Big Data analytics in healthcare and strategies to mitigate them.

The concept of Big Data covers a wide range of definitions extending from the data that is difficult to manage using traditional analytical tools, to the characterization of big data in terms of the volume, high velocity, huge variety and varied veracity. Though most of the studies define Big Data in terms of the aforementioned characteristics, one of the studies state that Big Data in healthcare is also characterized by having energy and life-span [25], but the literature lacks detailed description of these characteristics especially with regards to healthcare. Studies have demonstrated that Big Data analytics differ from traditional analytical approach in terms that instead of tracking care quality and outcomes in retrospective view by using deductive reasoning, it uses inductive reasoning for prospective analysis of data [65]. These techniques of data analysis are hypotheses-generating rather than hypotheses-testing since they focus on finding association and correlation in the observational data and not on the casual relationship between variables. But it needs for the testing of hypotheses before applying results into a clinical practice [66]. Few of the studies state that there is a need for the application of human judgment and supervision on the insights obtained from application of Big Data analytics in healthcare [53], as it would prevent the occurrence of adverse events which result from relying solely on Big Data analytics. Big Data analytics can, thus,

lead to a shift from experience-based to evidence-based decision making in healthcare [8].

Most studies in literature have shown that there are a large number of sources of healthcare Big Data. Clinical and administrative data in healthcare comes from various sources including healthcare providers, laboratories, diagnostic companies, insurance companies, pharmaceutical firms, not-for-profit organizations, government and web-health portals [29], [30], [32], [33], [67]–[69]. The literature on Big Data technologies and techniques that are used to clinical data is largely fragmented. The majority of this literature highlights the use of natural language processing for clinical as well as operational applications. As identified from the review, other Big Data techniques which find application in healthcare are cluster analysis, data mining, graph analytics, machine learning, neural networks, pattern recognition and spatial analysis. Studies showed that in most of the cases, Hadoop and tools which run on top of Hadoop are used for processing of patient-data [40], [41], [70]–[72], but since they are batch-processing tools, newer tools like Storm, Spark and GraphLab have started finding their application for streaming and real-time data [41].

Most of the studies that were reviewed, concerned the application of Big Data analytics in different areas of healthcare. According to them, Big Data analytics finds application in clinical decision support; personalized medicine; and optimization of clinical operations and cost-effectiveness of care. Thus, it should be noted that the integration of Big Data technology into healthcare can not only improve the quality of care, but enable early identification of high-risk patients by making use of real-time analytics and hence can benefit nations by saving lives. Studies on the use of Big Data analytics for cardiovascular diseases, diabetes, oncology, elderly care, gynecology, and clinical research have shown that it can enable delivery of timely care and cost-saving by eliminating inefficiencies [47], [52], [55], [55], [73]–[79].

Despite the tremendous value-addition with use of Big Data analytical tools, healthcare industry still lags in adoption of this technology due to many challenges. The non-availability of appropriate IT infrastructure, huge investment costs associated with implementing analytical tools, data privacy and security issues, fragmented data ownership and technical challenges such as data quality and multi-dimensionality of data are some of the issues. One of the studies identified lack of evidence of practical benefits as a major cause behind the reluctance for using Big Data analytics in healthcare [10]. Few studies highlighted that there is a shortage of skilled Big Data analysts with knowledge of healthcare, who have the ability to identify right data and right tools to use for analysis of health related data and interpret insights obtained after analysis, which makes the use of technology difficult. The complexity of Big Data analytical systems, is also one of the factors for limited use of technology for healthcare applications [8]. There is need of strategies for mitigating these challenges in order to realize its full potential. Some of these strategies include change in organizational culture; health-information exchange; training of key healthcare personnel; development of simple-transparent Big Data systems; use of cloud storage and distributed data processing; and strengthening of IT security.

### *B. Gaps and implications for future research*

- None of the usability studies on Big Data analytics included in this review discussed about quantitative results by the usage of technology. Most of these studies used qualitative approach to explain the benefits and challenges of using Big Data technology for healthcare applications, although the use of the quantitative approach will provide evidence for the practical benefits and will help greatly in the adoption of technology.
- Most of the studies included in the review, were from the developed countries. It is essential to promote the research on Big Data analytics in healthcare in the developing countries, since that will enable delivery of better quality care.

### *C. Limitations*

While the literature covers information about Big Data analytics and its role in healthcare and medicine, current research has few limitations. The contents of this study consists of a systematic review of the current status of Big Data technology in healthcare, but it does not take into consideration the technical details regarding the implementation and results obtained in each of the study reviewed. Despite the use of a systematic approach for review, the inclusion of studies on 'big data analytics' in 'healthcare' for this review was based on subjective judgment, hence the cross-reference articles were also considered for this review.

## V. Conclusion

The Big Data analytics concept has emerged as a new frontier for enhancing healthcare delivery. With the opportunities created by digital and information revolution, healthcare industry can exploit the potential benefits of leveraging Big Data technology. Big Data analytics increasingly provides value to healthcare by improving healthcare quality and outcomes and providing cost-effective care. The predictive nature and pattern-recognition aspect of Big Data analytics enable the shift from experience-based medicine to evidence-based medicine. Through its systematic review, the study presents a useful starting point for the application of Big Data analytics in future healthcare research. In addition, the study reflects that once the scope of Big Data analytics is defined; its characteristics and features are understood; and challenges are properly tackled, its application will help maximize the healthcare value through promoting the extensive usage of insights and informed decisions.

## References

[1] A. Belle, R. Thiagarajan, S. M. R. Soroushmehr, F. Navidi, D. A. Beard, and K. Najarian, "Big Data Analytics in Healthcare," *BioMed Research International*, 2015. [Online]. Available: https://www.hindawi.com/journals/bmri/2015/370194/abs/. [Accessed: 05-Jul-2018].

[2] A. Asante-Korang and J. P. Jacobs, "Big Data and paediatric cardiovascular disease in the era of transparency in healthcare," *Cardiol. Young*, vol. 26, no. 08, pp. 1597–1602, Dec. 2016.

[3] Z. Gon-Mohammadi, M. Sargolzaei Javan, and M. K. Akbari, "An effective model for store and retrieve big health data in cloud computing," *Comput. Methods Programs Biomed.*, vol. 132, pp. 75–82, Aug. 2016.

[4] F. J. Martin-Sanchez, V. Aguiar-Pulido, G. H. Lopez-Campos, N. Peek, and L. Sacchi, "Secondary Use and Analysis of Big Data Collected for Patient Care: Contribution from the IMIA Working Group on Data Mining and Big Data Analytics," *Yearb. Med. Inform.*, vol. 26, no. 01, pp. 28–37, Aug. 2017.

[5] Y. Wang and N. Hajli, "Exploring the path to big data analytics success in healthcare," *J. Bus. Res.*, vol. 70, pp. 287–299, Jan. 2017.

[6] B. Cyganek *et al.*, "A Survey of Big Data Issues in Electronic Health Record Analysis," *Appl. Artif. Intell.*, vol. 30, no. 6, pp. 497–520, Jul. 2016.

[7] F. F. Costa, "Big data in biomedicine," *Drug Discov. Today*, vol. 19, no. 4, pp. 433–440, Apr. 2014.

[8] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: promise and potential," 2014.

[9] J. S. Rumsfeld, K. E. Joynt, and T. M. Maddox, "Big data analytics to improve cardiovascular care: promise and challenges," *Nat. Rev. Cardiol.*, vol. 13, no. 6, pp. 350–359, Jun. 2016.

[10] M. Cox and D. Ellsworth, "Application-controlled demand paging for out-of-core visualization," in *Proceedings. Visualization '97 (Cat. No. 97CB36155)*, Phoenix, AZ, USA, 1997, pp. 235-244,.

[11] F. Diebold, and F. X. Diebold, and F. X. Diebold, *Big Data" Dynamic Factor Models for Macroeconomic Measurement and Forecasting*. 2000.

[12] E. Guerra, J. de Lara, A. Malizia, and P. Díaz, "Supporting user-oriented analysis for multi-view domain-specific visual languages," *Inf. Softw. Technol.*, vol. 51, no. 4, pp. 769–784, Apr. 2009.

[13] J. Manyika, M. Chui, B. Brown, and A. Byers, "Big Data: The Next Frontier for Innovation, Comptetition, and Productivity | Request PDF," *ResearchGate*, 2011. [Online]. Available: https://www.researchgate.net/publication/260480165_Big_Data_The_Next_Frontier_for_Innovation_Comptetition_and_Productivity. [Accessed: 05-Oct-2018].

[14] B. Feldman and E. Martin, "Big Data in Healthcare Hype and Hope | Electronic Health Record | Big Data," *Scribd*, 2012. [Online]. Available: https://www.scribd.com/document/107279699/Big-Data-in-Healthcare-Hype-and-Hope. [Accessed: 29-Sep-2018].

[15] Khushboo Wadhwani and Dr. Yun Wang, "Big Data Challenges and Solutions," 2017.

[16] A. Labrinidis and H. V. Jagadish, "Challenges and Opportunities with Big Data," *Proc VLDB Endow*, vol. 5, no. 12, pp. 2032–2033, Aug. 2012.

[17] V. Ganjir, B. K. Sarkar, and R. Kumar, "BIG DATA ANALYTICS FOR HEALTHCARE," *IJRETS*, 09-Dec-2016. .

[18] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big Data: Issues and Challenges Moving Forward," in *2013 46th Hawaii International Conference on System Sciences*, 2013, pp. 995–1004.

[19] A. Katal, M. Wazid, and R. H. Goudar, "Big data: Issues, challenges, tools and Good practices," in *2013 Sixth International Conference on Contemporary Computing (IC3)*, 2013, pp. 404–409.

[20] R. Nambiar, R. Bhardwaj, A. Sethi, and R. Vargheese, "A look at challenges and opportunities of Big Data analytics in healthcare," in *2013 IEEE International Conference on Big Data*, 2013, pp. 17–22.

[21] C. L. Philip Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," *Inf. Sci.*, vol. 275, pp. 314–347, Aug. 2014.

[22] S. Sagiroglu and D. Sinanc, "Big data: A review," in *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 2013, pp. 42–47.

[23] S. Selamat, "Survey on the Emergence of Big Data," *ResearchGate*, 2018. [Online]. Available: https://www.researchgate.net/publication/323116268_Survey_on_the_Emergence_of_Big_Data. [Accessed: 13-Aug-2018].

[24] J. Sun and C. K. Reddy, "Big Data Analytics for Healthcare," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, 2013, pp. 1525–1525.

[25] I. D. Dinov, "Volume and value of big healthcare data," *J. Med. Stat. Inform.*, vol. 4, no. 1, p. 3, 2016.

[26] C. Auffray *et al.*, "Making sense of big data in health research: Towards an EU action plan," *Genome Med.*, vol. 8, no. 1, Dec. 2016.

[27] Y. Wang, L. Kung, and T. A. Byrd, "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations," *Technol. Forecast. Soc. Change*, vol. 126, pp. 3–13, Jan. 2018.

[28] S. Gagana and K. Palamesha, "Big Data Revolution in Health Care Sector," 2018, pp. 265–279.

[29] W. B. Rouse and N. Serban, *Understanding and managing the complexity of healthcare*. Cambridge, Massachusetts: The MIT Press, 2014.

[30] E. A. Mohammed, B. H. Far, and C. Naugler, "Applications of the MapReduce programming framework to clinical big data analysis: current landscape and future trends," *BioData Min.*, vol. 7, no. 1, Dec. 2014.

[31] C. Weng and M. G. Kahn, "Clinical Research Informatics for Big Data and Precision Medicine:," *IMIA Yearb.*, no. 1, pp. 211–218, 2016.

[32] B. E. Huang, W. Mulyasasmita, and G. Rajagopal, "The path from big data to precision medicine," *Expert Rev. Precis. Med. Drug Dev.*, vol. 1, no. 2, pp. 129–143, Mar. 2016.

[33] P. S. Bradley, "Implications of Big Data Analytics on Population Health Management," *Big Data*, vol. 1, no. 3, pp. 152–159, Sep. 2013.

[34] D. V. Dimitrov, "Medical Internet of Things and Big Data in Healthcare," *Healthc. Inform. Res.*, vol. 22, no. 3, p. 156, 2016.

[35] G. Taglang and D. B. Jackson, "Use of 'big data' in drug discovery and clinical trials," *Gynecol. Oncol.*, vol. 141, no. 1, pp. 17–23, Apr. 2016.

[36] P. Groves, B. Kayyali, D. Knott, and S. V. kuiken, "The 'big data' revolution in healthcare," p. 22, 2013.

[37] D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar, "Big Data In Health Care: Using Analytics To Identify And Manage High-Risk And High-Cost Patients," *Health Aff. (Millwood)*, vol. 33, no. 7, pp. 1123–1131, Jul. 2014.

[38] K. R. Ghani, K. Zheng, J. T. Wei, and C. P. Friedman, "Harnessing Big Data for Health Care and Research: Are Urologists Ready?," *Eur. Urol.*, vol. 66, no. 6, pp. 975–977, Dec. 2014.

[39] N. M. S. kumar, T. Eswari, P. Sampath, and S. Lavanya, "Predictive Methodology for Diabetic Data Analysis in Big Data," *Procedia Comput. Sci.*, vol. 50, pp. 203–208, 2015.

[40] D. Markonis, R. Schaer, I. Eggel, H. Muller, and A. Depeursinge, "Using MapReduce for Large-Scale Medical Image Analysis," in *2012 IEEE Second International Conference on Healthcare Informatics, Imaging and Systems Biology*, La Jolla, CA, USA, 2012, pp. 1–1.

[41] N. Peek, J. H. Holmes, and J. Sun, "Technical Challenges for Big Data in Biomedicine and Health: Data Sources, Infrastructure, and Analytics:," *IMIA Yearb.*, vol. 9, no. 1, pp. 42–47, 2014.

[42] R. K. Ferrell, S. R. Sukumar, and R. Natarajan, "Quality of Big Data in health care," *Int. J. Health Care Qual. Assur.*, vol. 28, no. 6, pp. 621–634, Jul. 2015.

[43] I. Cano, A. Tenyi, E. Vela, F. Miralles, and J. Roca, "Perspectives on Big Data applications of health information," *Curr. Opin. Syst. Biol.*, vol. 3, pp. 36–42, Jun. 2017.

[44] T. Schultz, "Turning healthcare challenges into big data opportunities: A use-case review across the pharmaceutical development lifecycle," *Bull. Am. Soc. Inf. Sci. Technol.*, vol. 39, no. 5, pp. 34–40, Jun. 2013.

[45] R. Bellazzi, A. Dagliati, L. Sacchi, and D. Segagni, "Big Data Technologies: New Opportunities for Diabetes Management," *J. Diabetes Sci. Technol.*, vol. 9, no. 5, pp. 1119–1125, Sep. 2015.

[46] L. J. Frey, L. Lenert, and G. Lopez-Campos, "EHR Big Data Deep Phenotyping: Contribution of the IMIA Genomic Medicine Working Group," *IMIA Yearb.*, vol. 9, no. 1, pp. 206–211, 2014.

[47] Y. Wang, L. Kung, W. Y. C. Wang, and C. G. Cegielski, "An integrated big data analytics-enabled transformation model: Application to health care," *Inf. Manage.*, vol. 55, no. 1, pp. 64–79, Jan. 2018.

[48] N. Szlezák, M. Evers, J. Wang, and L. Pérez, "The Role of Big Data and Advanced Analytics in Drug Discovery, Development, and Commercialization," *Clin. Pharmacol. Ther.*, vol. 95, no. 5, pp. 492–495, May 2014.

[49] S. Fodeh and Q. Zeng, "Mining Big Data in biomedicine and health care," *J. Biomed. Inform.*, vol. 63, pp. 400–403, Oct. 2016.

[50] F. Zhang, J. Cao, S. U. Khan, K. Li, and K. Hwang, "A task-level adaptive MapReduce framework for real-time streaming data in healthcare applications," *Future Gener. Comput. Syst.*, vol. 43–44, pp. 149–160, Feb. 2015.

[51] J. Wu, H. Li, S. Cheng, and Z. Lin, "The promising future of healthcare services: When big data analytics meets wearable technology," 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378720616300775. [Accessed: 09-Jul-2018].

[52] A.-T. Maia, S.-J. Sammut, A. Jacinta-Fernandes, and S.-F. Chin, "Big data in cancer genomics," *Curr. Opin. Syst. Biol.*, vol. 4, pp. 78–84, Aug. 2017.

[53] G. V. Asokan and V. Asokan, "Leveraging 'big data' to enhance the effectiveness of 'one health' in an era of health informatics," *J. Epidemiol. Glob. Health*, vol. 5, no. 4, pp. 311–314, Dec. 2015.

[54] M. Grossglauser and H. Saner, "Data-driven healthcare: from patterns to actions," *Eur. J. Prev. Cardiol.*, vol. 21, no. 2_suppl, pp. 14–17, Nov. 2014.

[55] H. Geerts *et al.*, "Big data to smart data in Alzheimer's disease: The brain health modeling initiative to foster actionable knowledge," *Alzheimers Dement.*, vol. 12, no. 9, pp. 1014–1021, Sep. 2016.

[56] R. Budhiraja, R. Thomas, M. Kim, and S. Redline, "The Role of Big Data in the Management of Sleep-Disordered Breathing," *Sleep Med. Clin.*, vol. 11, no. 2, pp. 241–255, Jun. 2016.

[57] C. S. Kruse, R. Goswamy, Y. Raval, and S. Marawi, "Challenges and Opportunities of Big Data in Health Care: A Systematic Review," *JMIR Med. Inform.*, vol. 4, no. 4, p. e38, Nov. 2016.

[58] S. Salas-Vega, A. Haimann, and E. Mossialos, "Big Data and Health Care: Challenges and Opportunities for Coordinated Policy Development in the EU," *Health Syst. Reform*, vol. 1, no. 4, pp. 285–300, May 2015.

[59] T. R. McNutt, K. L. Moore, and H. Quon, "Needs and Challenges for Big Data in Radiation Oncology," *Int. J. Radiat. Oncol.*, vol. 95, no. 3, pp. 909–915, Jul. 2016.

[60] A. Boilson, A. Staines, J. Connolly, P. Davis, and R. Connolly, "Transformation of Big Data for Health in the European Union.," *Eur. Union*, p. 8, 2018.

[61] D. R. Luna, J. . Mayan, M. J. García, A. A. Almerares, and M. Househ, "Challenges and Potential Solutions for Big Data Implementations in Developing Countries," *Yearb. Med. Inform.*, vol. 9, no. 1, pp. 36–41, Aug. 2014.

[62] E. Vayena, J. Dzenowagis, J. S. Brownstein, and A. Sheikh, "Policy implications of big data in the health sector," *Bull. World Health Organ.*, vol. 96, no. 1, pp. 66–68, Jan. 2018.

[63] R. Naydenov and D. Liveri, "Good Practices and Recommendations on the Security and Resilience of Big Data Services | Big Data | Analytics," *Scribd*, 2015. [Online]. Available: https://www.scribd.com/document/335796630/Good-Practices-and-Recommendations-on-the-Security-and-Resilience-of-Big-Data-Services. [Accessed: 29-Sep-2018].

[64] J. Adler-Milstein and A. K. Jha, "Healthcare's 'big data' challenge," *Am. J. Manag. Care*, vol. 19, no. 7, pp. 537–538, Jul. 2013.

[65] C. H. Lee and H.-J. Yoon, "Medical big data: promise and challenges," *Kidney Res. Clin. Pract.*, vol. 36, no. 1, pp. 3–11, Mar. 2017.

[66] H. Care and B. Outcomes, "Leveraging big data and analytics in healthcare and life sciences," p. 4, 2012.

[67] K. Miller, "Big Data Analytics In Biomedical Research | Biomedical Computation Review," 2012. [Online]. Available: http://biomedicalcomputationreview.org/content/big-data-analytics-biomedical-research. [Accessed: 28-Sep-2018].

[68] M. Swan, "The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery," *Big Data*, vol. 1, no. 2, pp. 85–99, Jun. 2013.

[69] M. J. Ward, K. A. Marsolo, and C. M. Froehle, "Applications of business analytics in healthcare," *Bus. Horiz.*, vol. 57, no. 5, pp. 571–582, Sep. 2014.

[70] C.-L. Hung and Y.-L. Lin, "Implementation of a Parallel Protein Structure Alignment Service on Cloud," *Int. J. Genomics*, vol. 2013, pp. 1–8, 2013.

[71] L. Wang, D. Chen, R. Ranjan, S. U. Khan, J. KolOdziej, and J. Wang, "Parallel Processing of Massive EEG Data with MapReduce," in *2012 IEEE 18th International Conference on Parallel and Distributed Systems*, Singapore, Singapore, 2012, pp. 164–171.

[72] B. Meng, G. Pratx, and L. Xing, "Ultrafast and scalable cone-beam CT reconstruction using MapReduce in a cloud computing environment: Ultrafast and scalable CBCT reconstruction using MapReduce," *Med. Phys.*, vol. 38, no. 12, pp. 6603–6609, Nov. 2011.

[73] H. T. Wong, Q. Yin, Y. Q. Guo, K. Murray, D. H. Zhou, and D. Slade, "Big data as a new approach in emergency medicine research," *J. Acute Dis.*, vol. 4, no. 3, pp. 178–179, Aug. 2015.

[74] M. Viceconti, P. Hunter, and R. Hose, "Big Data, Big Knowledge: Big Data for Personalized Healthcare," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 4, pp. 1209–1215, Jul. 2015.

[75] A. Mandawat, A. E. Williams, and S. A. Francis, "Cardio-oncology," *Heart Fail. Clin.*, vol. 13, no. 2, pp. 403–408, Apr. 2017.

[76] J. Kim, "Big Data, Health Informatics, and the Future of Cardiovascular Medicine," *J. Am. Coll. Cardiol.*, vol. 69, no. 7, pp. 899–902, Feb. 2017.

[77] E. A. Erekson and C. B. Iglesia, "Improving Patient Outcomes in Gynecology: The Role of Large Data Registries and Big Data Analytics," *J. Minim. Invasive Gynecol.*, vol. 22, no. 7, pp. 1124–1129, Nov. 2015.

[78] G. N. Nadkarni, S. G. Coca, and C. M. Wyatt, "Big data in nephrology: promises and pitfalls," *Kidney Int.*, vol. 90, no. 2, pp. 240–241, Aug. 2016.

[79] I. El Naqa, "Perspectives on making big data analytics work for oncology," *Methods*, vol. 111, pp. 32–44, Dec. 2016.

**The University of Zambia, Copperbelt University and Mulungushi University, The three major public universities are the organisers of the International Conference in Information and Communication Technologies (ICICT2018)**

**The main sponsor of the ICICT2018 is the Information and Communications Technology Society of Zambia (ICTSZ). The other key major sponsors were ZAMREAN and Zambia National Airport.**


ISBN 978-9982-70-787-9