A Classification Framework for Detecting Malicious Email Contents

Refilwe Moremi[1] and Johnson Dehinbo[2]

***Abstract*—**The advent of Internet has played a significant role in the development of business enterprises worldwide. However, this has led to the emergence of criminal activities that hamper the system infrastructure through malicious emails. These malicious activities are evolving through social engineering like phishing, spamming and others like Malware which is intentionally developed to perform malicious actions and malware has become a huge problem in our society [8]. The current tools are insufficient to deduce and filter this malicious information on the Internet, which are signature-based but it is often unseen and time-consuming. In this research project, we proposed the design of a novel framework that detects new and unseen behavior in sending and recipient email automatically. We extract the email body to understand the contextual feature and classify the malicious information using Naïve Bayesian method. We demonstrate our proposed framework using a database that was created manually to distinguish emails with malicious activities because everyone, every day is using internet for business, education and communication (emails) and many other purpose document.

***Keywords*—**Classification framework, detection system, Malicious email, Spamming protection.

I. Introduction

THE number of user affected by phishing website is increasing gradually. Within the growth several tools has been developed to reduce the problem, however the main focus of this study is to develop a novel framework that will be able to detect malicious email. Several efforts by International Establishments to detect malicious email has been done however the tools seems to be insufficient to filter this malicious web/email activities [3].

In this paper, Naïve Bayesian method is used to extract email body to understand contextual features and classify malicious information. Naïve Bayesian classifier computes or detects that the mail is malicious by given the features that are contained in the email, the output of Naïve Bayes algorithm labels each by features that it contained. For example, an email containing a many malicious features could be labelled as an email virus.

A. Problem Statement

First, the rise of email as medium communication raises several issues, because majority of an email sent are spam (email advertising for some product sent to a mailing list or newsgroup), however many emails received by many users are overwhelming, Email consumes significant time and attention in the workplace but they are affective and faster way of communication, hence they are several information security that were identified by use of emails, this study has come as a results of fraud behavior that is been done by means of email communication.it is very much important to keep an email as a quicker communication tool and the user must find it comfortable to use and unfortunately emails are the most popular and effective form of communication in our days [15].

B. Research Questions and Sub-questions

Email services drive opportunities for people to communicate or interact, they also create new opportunities for criminal, and therefore the research question for this paper is: How to Develop a Technique that will be able to Detect and Analyse Malicious Email?

Sub Questions

- ❖ How to extract features that can be malicious within content of an email?
- ❖ How to classify the malicious, harmful content that are involved in the email?
- ❖ How to conclude that the contents within the email body are malicious or not?

C. Objectives

The general objective is to design a novel framework that will be able to detect malicious content in an email on the Internet to achieve the sub-objective:

- ❖ To be able to develop a tool or a framework that will be able to extract malicious features within email content
- ❖ To develop a tool that will be able to classify harmful content that are involved in the email.
- ❖ To conclude that the contents within the email body are malicious or not.

.

Refilwe Moremi[1] is a B.Tech. student with the Department of Computer Science, Tshwane University of Technology, Soshanguve campus, Pretoria, South Africa. (phone: +27-12-3829261; e-mail: filwemare@gmail.com respectively).

Johnson Dehinbo[1] is with the Department of Computer Science, Tshwane University of Technology, Soshanguve campus, Pretoria, South Africa. (phone: +27-12-3829219; e-mail: DehinboOJ@tut.ac.za respectively).

https://doi.org/10.15242/ICEHM.E0717002

II. Literature Review

A. Introduction

Detecting malicious executable is not a new problem in security, several research effort has been done on how to classify malicious features within an email, however we are still going in researching for a better tool for this problem because most email based spam detector tools are unable to protect other web services [7]. The main purpose of an email is to share and distribute information [7]. Various researchers have suggested many tools and algorithms for spam filtering however most of these tools pays attention to individual parameters [10].

B. Spamcatter

Spamcatter is a tool to mine email in the current time, follow the embedded link structure and automatically cluster the destination web site by the use of image shingling to be able to capture graphical similarity between rendered web sites [9], was developed to characterise internet scam hosting infrastructure but still the problem was unresolved because hacker are however able to send anonymous emails to user day in and out. Email spam is sent by spammer directly to the victims or users therefore in order to deal with with spam based attacks, it is very important to characterize their infrastructure and how they are grouped with each other [17].

C. Botimer

The Researcher have proposed few approaches [2], in which the Botminer was developed to be able to analyze the network traffic for protocol and structure autonomous botnet detection, botnets(is a network of bargained machines under the influence of malware code) are now key avenue for many internet misusers such as spam, identity theft and phishing, which I found very useful and malicious protective but the issue of anonymous email whereby the tool is unable to identify where the email come from is a serious weakness of this tool, because that anonymous email can come from anyone and it could be malicious.

D. N-grams

The Identification of new malicious code using N-grams they have developed an N-gram signature to be able to detect malicious code alphabet and character within email, I personally effective and substantial to use because it will be able to detect a specifically a malicious character and code within an email and reject find any malicious code it will reject that particular email, the too detect where the email comes from, subject, body and the email signature and validate this signature information which was quite interesting tool.

E. Summary of the Review

We are certainly not the first to research about malicious email detecting tool. Perhaps the work most closely to this is the work of Federico Maggie and Stefano Zanoro [15] which concentrated on developing a detection technique, which have both industrial applications such as network monitoring and protection as well as research application such as software behavioral analysis or malware classification.

Recently several papers proposed different approaches to detect malicious activities on the internet using email and we all know that detecting malicious URLs is essential task in network security intelligence [6], this author made new contributions beyond the state of the art method on malicious URL detection. Instead of using any pre-defined features, they propose to dynamically extract lexical patterns from URLs, which I find it to be a very good approach and unique to use it is also effective and efficiency approach.

In recent years targeted email attacks to enable computer network exploitation have become more prevalent, more insidious and more widely documented, this target email attacks are not singular unrelated events, methods that has been developed by current research and methods to detect malicious email seems to be limited on an issue of addressing the scale of internet email abuse, such as spam and they are not focusing on addressing malicious email [16], has discovered this particular problem however his approach was not good enough to solve the problem. The overflowing of spam consumes not only computer, storage and network resources but also to remove or dismiss unwanted email it consumes time and attention [12].

The growing and popularity of smartphone has made them a target of malicious activity, malicious activities targeting smartphones are continuously increasing and are projected to continue to increase as they become more lucrative for cybercriminal worldwide, a real time detection system is needed in order to detect a new and unknown malicious URLs using cloud computing to bypass the phone limited computing and battery resource.

In the past six years, a tremendous growth and popularity of social networking fundamental changed the way we use the internet, yet the security industry has been slow to respond in the act of providing adequate tools for protecting he user, the author sees a social approach alert people or user about the malicious activities on the network, let people not give away crucial information on the internet , people should also be cautious to give away their important information on the internet regardless of what the situation or circumstances user might be facing.

Traditional defenses focus on how to detect malicious traffic and separate them from legitimate one. Identifying anomalous traffic senders and predicting identities of possible futures attackers is more promising approach for early detection and prevention of network attack. Some author has augured that web security and filters services, such as Websense and Bright cloud, track and analyze website to classify content and to detect phishing sites and sites hosting other probable malicious content such as spyware and key logging. Protection of malware even on the smartphone is based on the signature [11].

Previously a way to detect malicious software was based on signature matching, however signature matching only detect known mischievous software, in order to detect unknown malevolent software, it necessary to analyze the software for

https://doi.org/10.15242/ICEHM.E0717002

its power on the system when the software is executed, in one method, the software code can be statically analyzed for any malicious patterns [13]. Email is mainly unsafe because nearly all organisation or institutions allow email to enter their networks [5]. This study is however unique in provision of Malicious Email Blocker (MEB), which extracts email body to understand contextual features and classify malicious information.

III. Research Methodology

A. Methodology

In this research study, we basically investigate the malicious content in a suspicious email on the internet. We consider the challenge of malicious email on the internet as a binary problem/classification using Bayes Theory to identify those email contents that are malicious. The strategy of identifying those malicious content is to extract features content, so is to find the representation of malicious email, then apply machine learning algorithm to discriminate malicious and non-malicious email. This research framework will be implemented using PHP and MySQL.

Decision for not doing plugin on other system was taken due to the fact that each and every system has its own security measure however developing new and fresh system made it easier. According to [4], the suitable methodology to use for this research is the Design Science Methodology, because for every project that involves system development design science is a suitable methodology. The Design Science Methodology Consist of five steps, below are 5 steps in to details.

i. Awareness of the Problem

Internet is a basic unabated phenomenal in our social life. Therefore, it will be very impossible to have digital device without internet infrastructure. However, criminals have also utilize the basic fundamental of internet for cybercrime such as phishing, spamming and others. In this research project, we investigate the technique that will enhance the detection and identification of malicious email content on the internet.

ii. Suggestion

It is very important to understand the problem because it will lead to a design with new functionality or a new tool and this could be used as a guideline of a design of prototype.

iii. Research Strategy

The research project follows the Research design model but it definitely involves some algorithms.

iv. Research Design

This will be through the design of these research project methodologies like prototyping for development phase and experiments used in the evaluation phase.

v. Development

The development of the system take place at this stage, it can involve some algorithms, some system development tools such as PHP, MySQL, java etc.

vi. Evaluation

Evaluation plays an important role for every system development project because every system that has been developed testing/evaluating is necessary.

Conclusion

This is the ending stage of the research project, results for evaluation has been populated and seemly they are good regardless of minor abnormalities, remember design science is formation of new knowledge, it was highly relevant to be used on this research project because the problem is an existing and to solve the problem one has to develop a system. A suitable system development methodology for the system is prototyping.

TABLE 1
DESIGN SCIENCE COMPONENTS

Problem: malicious content within an emailing system and how to detect them
Suggestion: To design a system that will be able to detect malicious content.
Artifact to be developed: System that detect malicious words within an email.
Evaluation: performance test of Developed System
Results: Developed system that detect malicious words in an email and alert the receiver.

B. System Design

We have presented in section I that various researcher took effort and dedications to develop a system or a tool that could detect malicious content within an email, which most of them the technology and methodology used is similar to this research project but in this section, we will explain in to details the methodology used to develop this system.

Email is an effective and efficient way to communicate however due to fraud behavior that is happening through emailing system, user are now uncomfortable to send their confidential information or data through an email. The aim of this research project is to develop a software that will be able detect malicious words within an email and alert the receiver if the email is malicious. Design Science methodology is employed in this research project hence, according to [4] every research project that involves system development it will eventually employ prototyping in their system development phase and on the evaluation phase. Quantitative methodology will be employed in the form of experiment to evaluate effectiveness or performance of the MEB system. The system was developed using PHP and MySQL.

IV. System Output and Usability Evaluation

A. System Output

Each and every user has to be registered on the system with their names, email address and passwords, there is a database created to store users and their passwords unregistered user will not be able to logon.

https://doi.org/10.15242/ICEHM.E0717002

i. Words Database

Words database will be used to store all bad words which will be stored manually**,** the database can store as many words as possible

ii. Extraction and Classification

At this stage word from the user are featured in the Words Database (stored malicious word only) then they will be extracted to be classified whether they are malicious or not, hence according to our system whether the email is malicious or not it will still go through hence it will alert the user that the email is malicious.

After extraction is done, we now have to classify the extracted words using Naive Bayes classifier, however Naïve Bayes algorithm is a classifier technique based on Bayesian theory and it is particularly ideal when the dimensionality of inputs is high, this classifier is very popular and it has been known as email filtering tool [14]. Bayesian analysis uses prior probability. Prior probability are based on the previous experience, in this case frequency of words or character on an email will be used to calculate which word appears most or frequently, however we ever we are not obsessed as to how many time the word is appearing we are mostly focusing on the features of that particular word.

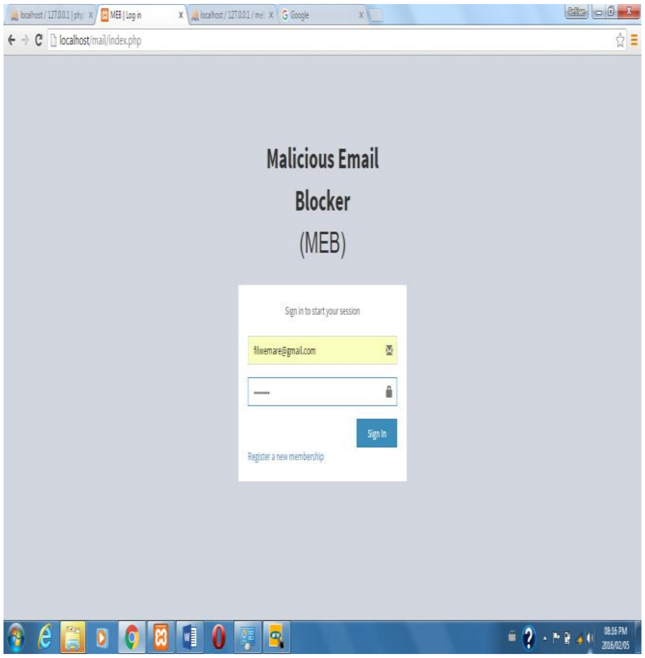
iii. Malicious and Non-Malicious Mails

The Receiver of an email will be able to receive the email whether is malicious or not however if the email is malicious the receiver will be alerted by a red star, but if the email does not contain malicious content it will not alert, a good email is indicated by a green star.

iv. System Output and Functional Testing.

System User have to register to be able to access the system, to register the user have to type their name, email address and password then the system will register them and enable them to login.

The user will see this page if they are not registered on the system or if they are registered then it means they have entered wrong password or their email address is invalid



Malicious Email

Blocker

(MEB)

Sign in to start your session

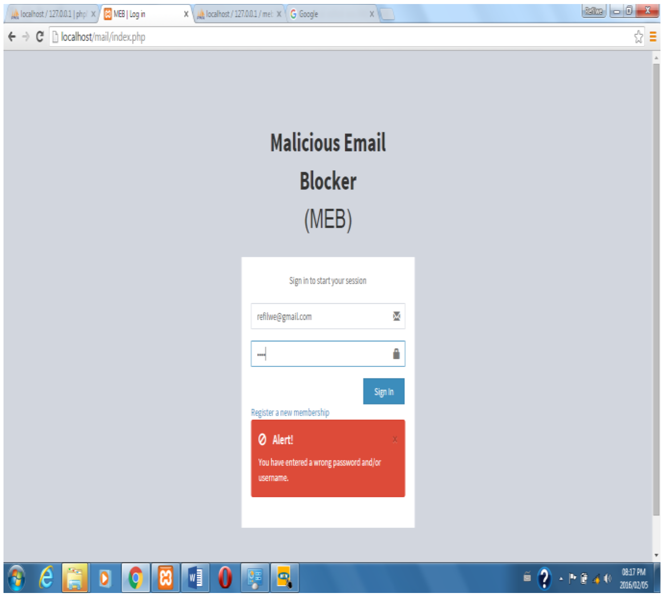
filwemare@gmail.com

••••••••

Sign In

Register a new membership

Fig. 1: Registration and Login Page



Malicious Email

Blocker

(MEB)

Sign in to start your session

refilwe@gmail.com

••••

Sign In

Register a new membership

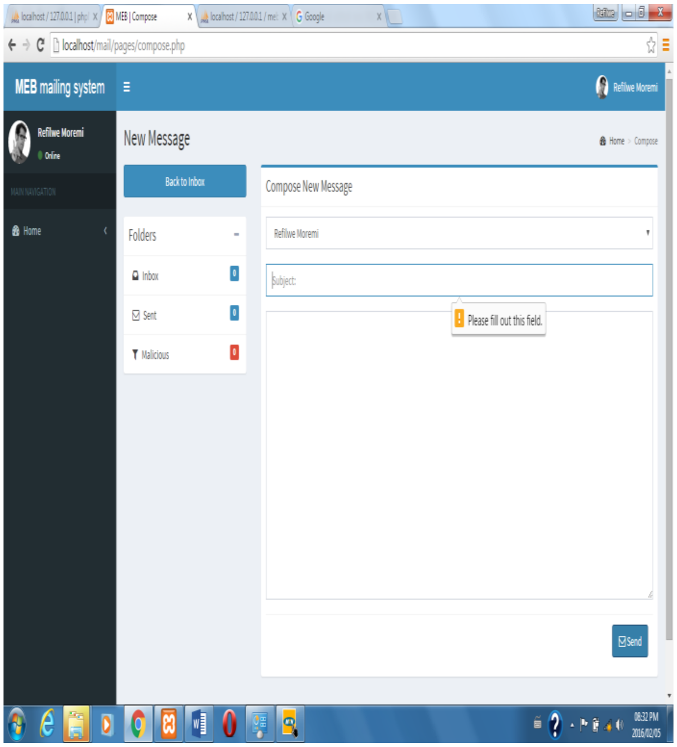
Alert!

You have entered a wrong password and/or username.

Fig. 2: Declined Page

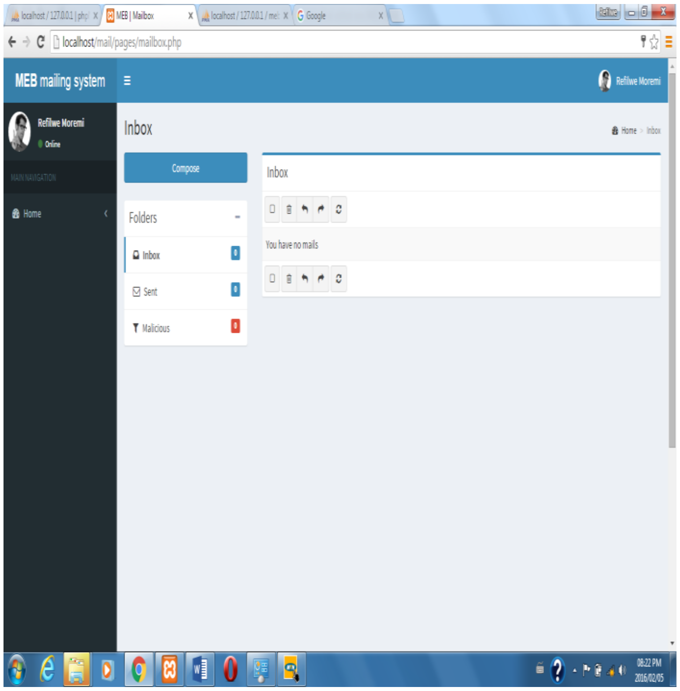
Figure 3 is the home page were users can be able to see their inbox, sent item and compose mail, by clicking compose the user is able to compose an email. Remember on this page every field must be field, if all fields are not filed the system will not send the message, when you click send it will pop a notification.

In figure 4, the user can now type a new email to a valid email address and click a send button, by clicking a send button an email will be sent to the dedicated email address.



Please fill out this field.

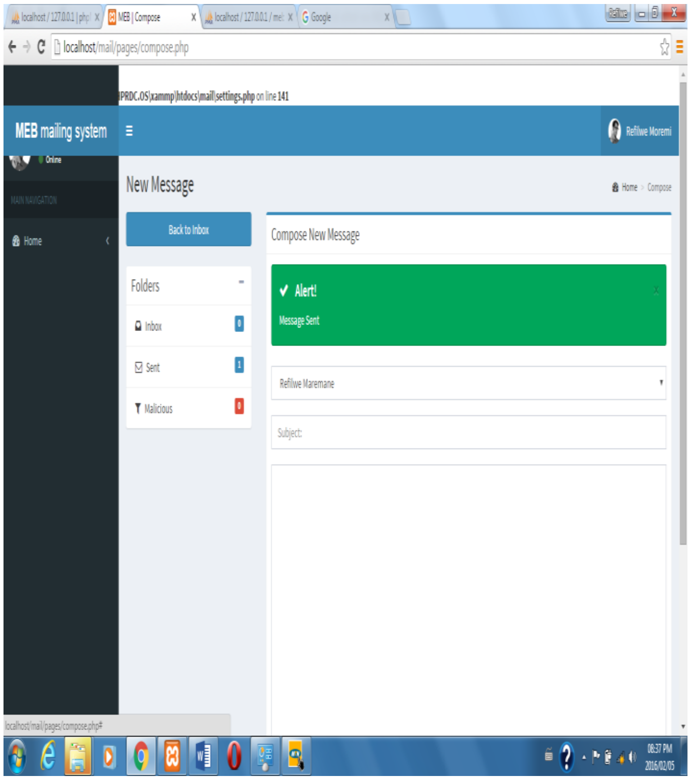
Fig. 3: Page for the inbox, sent items and compose mail



You have no mails

Fig. 4: Home page - Sender

User is done typing everything and clicked send button the page in figure 5 is alerting the sender that the message has been sent.



PRDC.OS\xampp\htdocs\mail\settings.php on line 141

✔ Alert!

Message Sent

Fig. 5: Sent mail page

On this page the new user is now logging, in this case we consider the user to be the receiver of the send email.

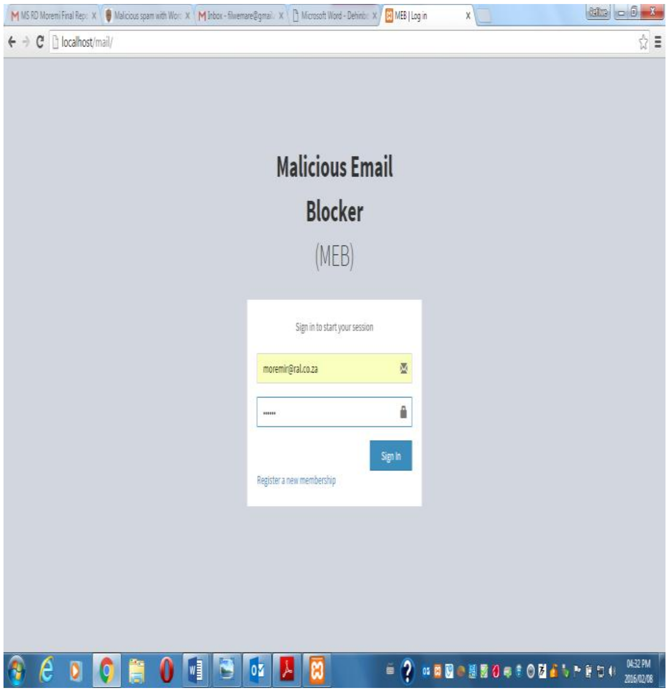
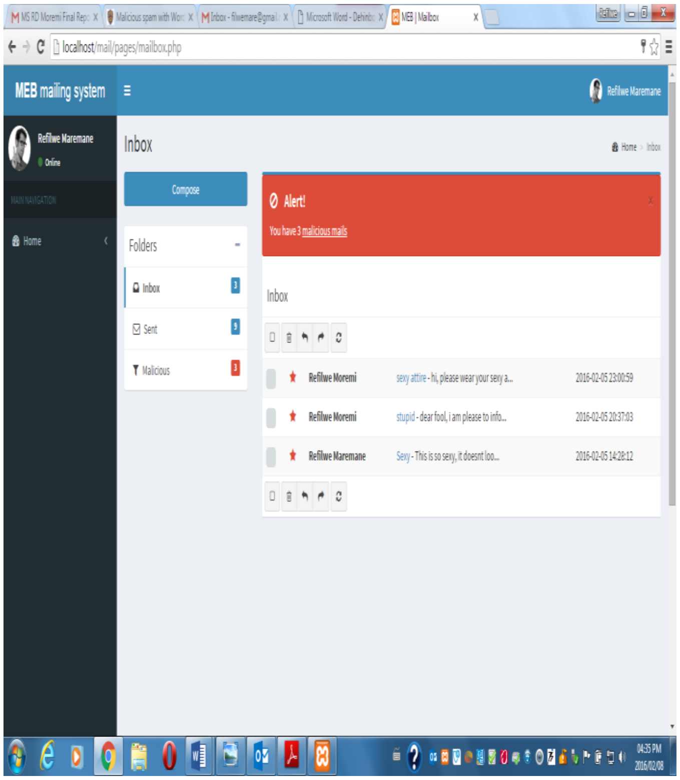


Fig. 6: Receiver Login Page

The receiver has opened the inbox as you can see there is a red notification that alerting the user that you have a malicious email in your inbox.



Alert!

You have 3 malicious mails

Refilwe Moremi	sexy attire - hi, please wear your sexy a...	2016-02-05 23:00:59
Refilwe Moremi	stupid - dear fool, i am please to info...	2016-02-05 20:37:03
Refilwe Maremane	Sexy - This is so sexy, it doesnt loo...	2016-02-05 14:28:12

Fig. 7: Receiver Inbox Page

B. Usability Testing and Evaluation

After the development of the MEB system was done, we tested the system then we took the system to the users (newly Developed Media Company in Limpopo Province, South Africa), the user had positive attitude towards the system hence MEB was their very first emailing system, in this case they could not evaluate MEB based on any system, MEB is new and fresh system for their environment, however the organization consists of mostly young and vibrant employee as we all know that young people of our days are more into electronic communication than any other type of communication, so the MEB system met their needs.

i. Usability Evaluation

Usability evaluation plays a significant role on every system project as developer get to understand the views of the user so that we can generally analyze their need base on their comments or feedback, the total number of 10 Feedbacks from questionnaires that were sent to user/ employee was, the one that populated most was the following:

- ❖ That the system recognizes some words as malicious even though they feel they are not e.g. Fool and Fooled in the MEB system are recognized as malicious. Stupid and stupidity are both registered as malicious simply because the word "stupid" is been stored in the database so when the system checks the minute it finds the word "Stupid" it automatically address it as a malicious hence the user has only typed Stupidity.

The Following issues were raised and discussed:

i. Overall Design

Most User have valued the system being simple and user friendly and overall they also find the system effective

ii. Notifications/ Popup messages

Users felt that notification or popup Messages should have set to expire after an hour or two because the malicious notification stays there even if you open you cannot delete until administrator attend to it.

iii. Communication

The user found the system impressing and effective with the fact that they did not have any emailing system in their organization, now they are able to communicate without one user/employee walking or calling the other employee/user for information and clarity and mostly but not lastly, emails will also assist them in terms of record keeping because if something is in writing you can have a record rather than I telephonic communication whereby is you have your word against someone's.

V. Conclusion

In this research paper, a system that detects or identifies malicious character or words within the body of an email was developed. After that, we cautiously examined methods or frameworks that were previously employed by other researchers to detect malicious behavior on the internet globally.

An experiment was conducted using a database which was manually created to store malicious character on an email and this will definitely not be the end of the road for this research, we can go further in the future and be linked to online database such as www.cloud.google.com\sql

This is not the end of the road, it is just a bus stop in future the research will go even further to check the synonyms, case sensitivity, and even go further to block the sender to send any malicious email, and I hope these will fully encourage other researcher to continue conducting researches to solve this global problem because technology is developing rapidly, everyone is now shifting from physical way of communication e.g. letter to electronic system such as emails [1].

Acknowledgment

R. Moremi thanks Mr. Abiodun Modupe for his valuable inputs and contributions to this study.

References

[1] Ali, S. H. A., Ozawa, S., Nakazato, J., Ban, T., & Shimamura, J. (2015). An autonomous online malicious spam email detection system using extended RBF network. *Proceedings of the International Joint Conference on Neural Networks*, *2015-September*. http://doi.org/10.1109/IJCNN.2015.7280826.

[2] Amin, R. M., Ryan, J. J. C. H., & Van Dorp, J. R. (2012). Detecting targeted malicious email. *IEEE Security and Privacy*, *10*(3), 64–71. http://doi.org/10.1109/MSP.2011.154

[3] Dai, J. (2010). Detecting malicious software by dynamic execution.

[4] Dehinbo J. (2014). Teaching Students on How Software Development Project can be turned into a Research Project. In M. Searson & M. Ochoa

(Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2014* (pp. 2103-2109). Chesapeake, VA: AACE. held at Jacksonville, Florida. USA. 17-21 March 2014 Retrieved July 17, 2014 from
http://www.editlib.org/p/131099.

[5] Deshmukh, P. (2014). Detecting of Targeted Malicious Email, 199–202.

[6] Faculty, C. S., & Sciences, A. (2012). MALICIOUS URL DETECTION BY DYNAMICALLY.

[7] Feroz, M. N., & Mengel, S. (2015). Examination of data, rule generation and detection of phishing URLs using online logistic regression. *Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data 2014*, 241–250. http://doi.org/10.1109/BigData.2014.7004239

[8] Hsu, F., Chen, H., Ristenpart, T., Li, J., & Su, Z. (2006). Back to the future: A framework for automatic malware removal and system repair. *Proceedings - Annual Computer Security Applications Conference, ACSAC*, 257–266. http://doi.org/10.1109/ACSAC.2006.16

[9] Huang, T. K., Valler, N. C., & Faloutsos, M. (2010). Characterizing the scam hosting infrastructure. *GLOBECOM - IEEE Global Telecommunications Conference*. http://doi.org/10.1109/GLOCOM.2010.5683162

[10] Jain, K. (2014). A Hybrid Approach for Spam Filtering using Local concentration based K-means Clustering, 194–199.

[11] Khune, R. S., & Thangakumar, J. (2012). A cloud-based intrusion detection system for Android smartphones. *2012 International Conference on Radar, Communication and Computing, ICRCC 2012*, 180–184. http://doi.org/10.1109/ICRCC.2012.6450572

[12] Li, C., & Liu, J. (n.d.). COMBINING BEHAVIOR AND BAYESIAN CHINESE SPAM FILTER, 2–5.

[13] Modupe, A., Olugbara, O. O., & Ojo, S. O. (2011). Exploring Support Vector Machines and Random Forests to Detect Advanced Fee Fraud Activities on Internet. *2011 IEEE 11th International Conference on Data Mining Workshops*, 331–335.
http://doi.org/10.1109/ICDMW.2011.81

[14] Rathod, S. B., & Pattewar, T. M. (2015). Content Based Spam Detection in Email using Bayesian Classifier, 1257–1261.

[15] Ravi, J. (2004). "Note to users".in *2007 Proc. INTERMAG Conf.*, pp. 2.2-1–2.2-6.

[16] Robertson, M., Yin, P., & Bo, Y. (2010). A social approach to security: Using social networks to help detect malicious web content BT - Intelligent Systems and Knowledge Engineering (ISKE), 2010 International Conference on, 436–441.
http://doi.org/10.1109/ISKE.2010.5680839

[17] Song, J., Inoue, D., Eto, M., Kim, H. C., & Nakao, K. (2010). An empirical study of spam: Analyzing spam sending systems and malicious web servers. *Proceedings - 2010 10th Annual International Symposium on Applications and the Internet, SAINT 2010*, 257–260. http://doi.org/10.1109/SAINT.2010.20

[18] Wei, S., Adviser-Sethi, A., & Adviser-Mirkovic, J. (2009). Detecting anomalous internet clients via behavior profiles and reputations. Retrieved from http://dl.acm.org/citation.cfm?id=1751497