Classification Method For Mitigating Phishy E-Mails

Adu, Folashade*, Ikerionwu Charles, Nnonyelu Henry****

*Department of Information Technology, FEDERAL UNIVERSITY OF TECHNOLOGY, OWERRI IMO STATE. NIGERIA
**Department of Information Technology, FEDERAL UNIVERSITY OF TECHNOLOGY, OWERRI IMO STATE. NIGERIA
**Department of Information Technology, FEDERAL UNIVERSITY OF TECHNOLOGY, OWERRI IMO STATE. NIGERIA

DOI: 10.29322/IJSRP.10.02.2020.p98107
http://dx.doi.org/10.29322/IJSRP.10.02.2020.p98107

Abstract: Phishing is one of the most popular forms of social engineering attack that involves use of emails or text message from known associate or friend that has links or attachments, which automatically downloads a spyware unnoticed when clicked. This epidemic has been in existence for ages and pose a big threat to IT security, individual and organizations reputation. Thus, this study aims at developing an email application embedded with algorithm to categorize incoming emails into high, medium and low threats. The development process followed object-oriented analysis and design methodology (OOADM), while the system was developed with PHP (Hypertext Preprocessor) and MySQL (Structured Query Language). The validation of the developed application suggests that when implemented, it could mitigate phishy emails.

Keywords: Mitigate; Phishers; Phishing, Social Engineering

i. INTRODUCTION:

Intruders are always on the lookout for ways to gain access to confidential and valuable resources of people such as computer systems, or corporate or personal information on them that can be used maliciously for the attackers' personal gain. Sometimes, it happens when there are genuine gaps in the security that they can breach (Mcdermott, 2005). The use of phishing emails increases social engineering attacks; programs; virus and malware can be attached to these emails and a simple click on such emails can enable tricksters; hackers; virus to gain access to your personal files; the attack could literally spread to other inboxes; if damages do not occur, they can cause loss of productivity, valuable network resources will be used up. Hogben, (2007); warned that we need to be wary of any email asking you to click a link, especially if it comes out of the blue and particularly from someone that is unknown.

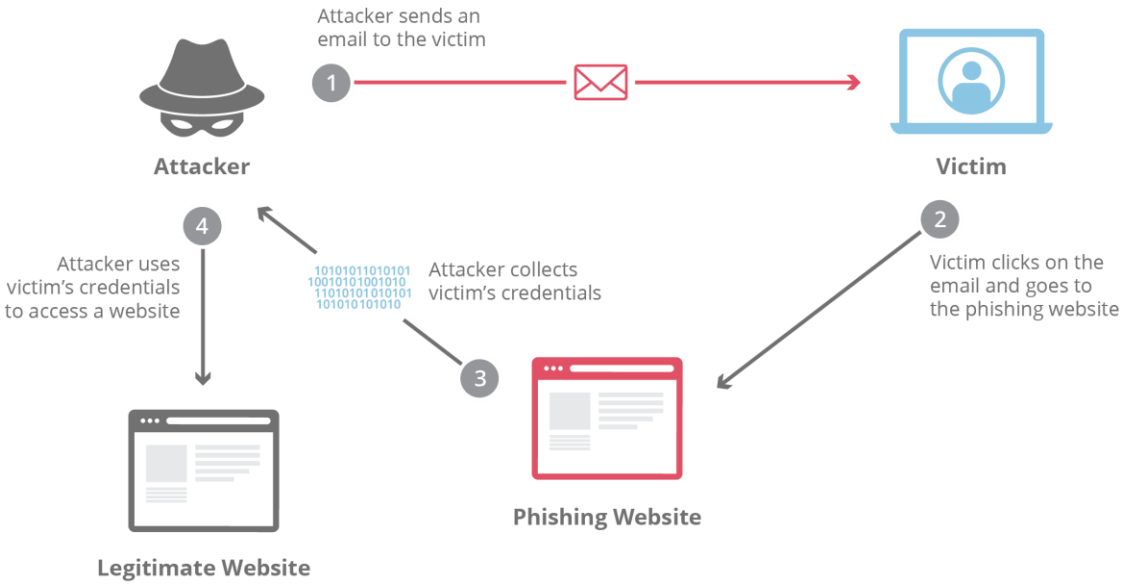

Fig 1: Process of phishing attack (Cloudflare Developer, 2018).

Fig 1 shows the method phishing attack occurs. The attacker sends an email to targeted victim, the email has a link that directs it to a phishing website, the victims uploads its credential on the website and the attacker uses the credentials to access a legitimate website.

ii. CATEGORIES OF PHISHING

The Phishing methodology is known as the bait and hook method where the bait is flung with the anticipation that an unwary user will grab it and bite onto it just like a fish. (Emm, 2006).

1. **Spear Phishing:** Herley and Florencio (2008), explained that, a spear phishing attack involves the attacker impersonating as an organization employee, especially as the organization IT consultant; then tailor the message to make the attack less conspicuous (using same wording and signature as the consultant normally does), thereby deceiving the recipients into thinking it is an authentic message. The message makes the recipients change their password and provides them with a link that redirects them to a malicious page where the attacker now captures their credentials (Parson et al., 2013).
2. **Scareware:** It involves victims being bombarded with false threats and alarms. It preys on deception. It can be distributed through spam emails that send out bogus warning such as "Your computer may be infected with harmful spyware program" it offers the tools used to clean the infected system or directs a user to a malicious site where user computer becomes infected with virus. (Applegate, 2009).
3. **Vishing:** Vishing works like phishing but does not always occur over the internet, but it is carried out using voice technology. It can be conducted by voice email, VoIP (voice over IP), or landline or cellular telephone. (Ellie, 2018).
4. **Spoofing:** Phishers mimics a trusted person; company, domains and sends out sophisticated direct emails. spoofing types such as: IP spoofing, URL spoofing, Email spoofing, DNS spoofing, and MAC spoofing. (Bermlerand et al, 2005).

- **MEANS OF IDENTIFYING PHISHING ATTACKS**

According to Calluzzo & Cante, (2004), they note that some people tell their sensitive information even with their knowledge of pervasive attack and threats. Cormack (2008), stated some common likely phishing scenario to include

1. Emails from a familiar person that claims to be stranded in a foreign country and soliciting for your assistance to transfer money so they can travel back home.

2. Emails claiming to be sent from a reputable news organization about trending news; asking the recipient to click on a link so they can access the full story, which leads the user to a website with malicious content.

3. An Emails from a user bank appearing to warn their recipient that their account has been threatened, hence, urging the individual to a website where their username and/or password can be reset.

4. Emails that threaten to harm recipients unless money is been sent (Blackmail)

5. Emails claiming to be confirmation concerning a complaint made by the recipient; and the recipient will be inclined to check on the link sent with the email to find out the supposed complain.

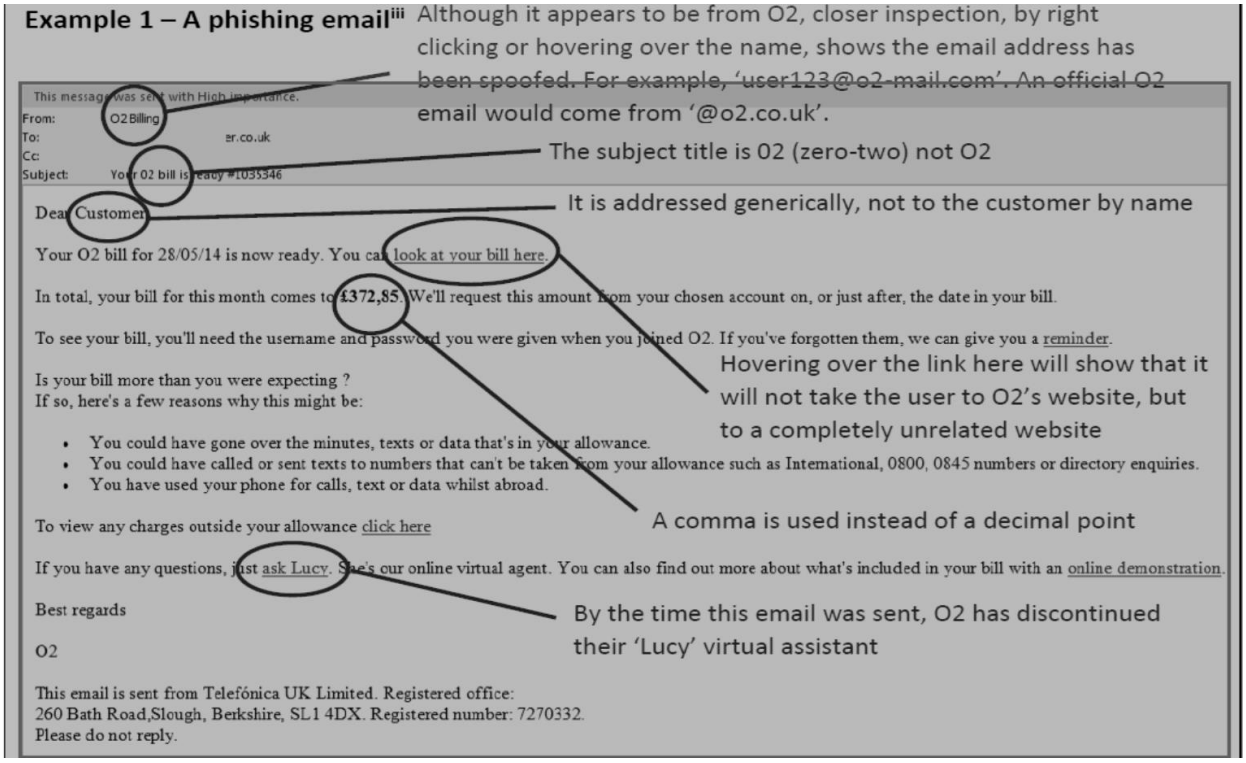

Fig 2: Sample of tricks deployed in the phishing scam (Chitrey, 2012).

Fig 2 shows the most sophisticated phishing campaign where scammers use a close replica of a legitimate website that tricked them into entering a username, password or other confidential information. By all appearances, this email that came from 'a billing company' you have an account with, and they are trying to provide information about the balance of your account. But note:

5. The misspelling of the words and discrepancies from the email address, to show the email is fraudulent and can be identified as a phishing scam.
6. The URL might be legitimate; the hyperlink would be directed to the phisher's webpage.
7. The mail has a generic salutation instead of being addressed by its customers' name like "Ms Joy" or "Mr Luke".

iii. DESIGN AND IMPLEMENTATION

This session simply refers to the filter features the model may look at, for it to either mark an email spam or non-spam:

a. Lower-casing: The entire email is converted into lower case, so that capitalization is ignored (e.g., IndIcaTE is treated the same as Indicate).

b. Stripping HTML: All HTML tags are removed from the emails. Many emails often come with HTML formatting; we remove all the HTML tags, so that only the content remains.

c. Normalizing URLs: All URLs are replaced with the text "httpaddr".

d. Normalizing Email Addresses: All email addresses are replaced with the text "emailaddr".

e. Normalizing Numbers: All numbers are replaced with the text "number".

f. Normalizing Dollars: All dollar signs ($) are replaced with the text "dollar".

g. Word Stemming: Words are reduced to their stemmed form. For example, "discount", "discounts", "discounted" and "discounting" are all replaced with "discount". Sometimes, the Stemmer strip off additional characters from the end, so "include", "includes", "included", and "including" are all replaced with "includ".

h. Removal of non-words: Non-words and punctuation have been removed.

i. All white spaces (tabs, newlines, spaces) have all been trimmed to a single space character.

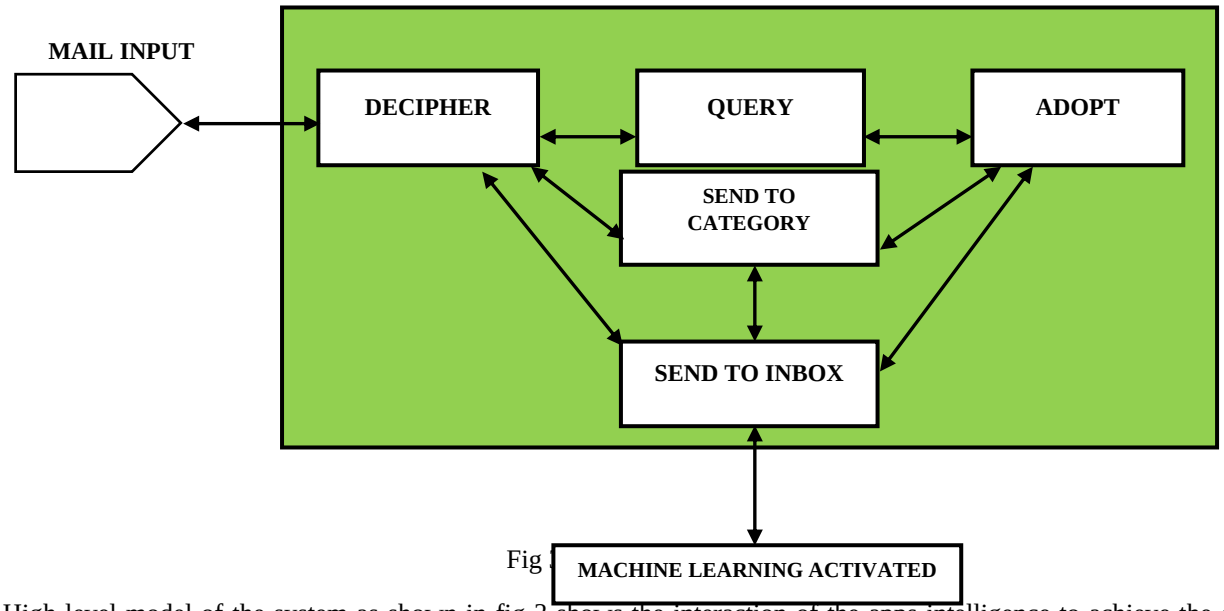

Fig .

High level model of the system as shown in fig 3 shows the interaction of the apps intelligence to achieve the desired result of the users. After the query the mail will be categorized accordingly (Low, Medium or high risk) but if after the query and the mail is not phishy, the mail will be sent to the inbox. The process is made possible using machine learning process.

- **PROCESS OF MITIGATING PHISHING EMAIL**

The researcher proposes categorization process of emails whereby email service providers (Gmail, yahoo mail, outlook etc.) should implement the categorizing suspicious emails into their risk factors. The mail system should filter through keywords and email addresses flagged as spam by people. The system should categorize them into: High risk mail, Medium risk mail and low risk mail.

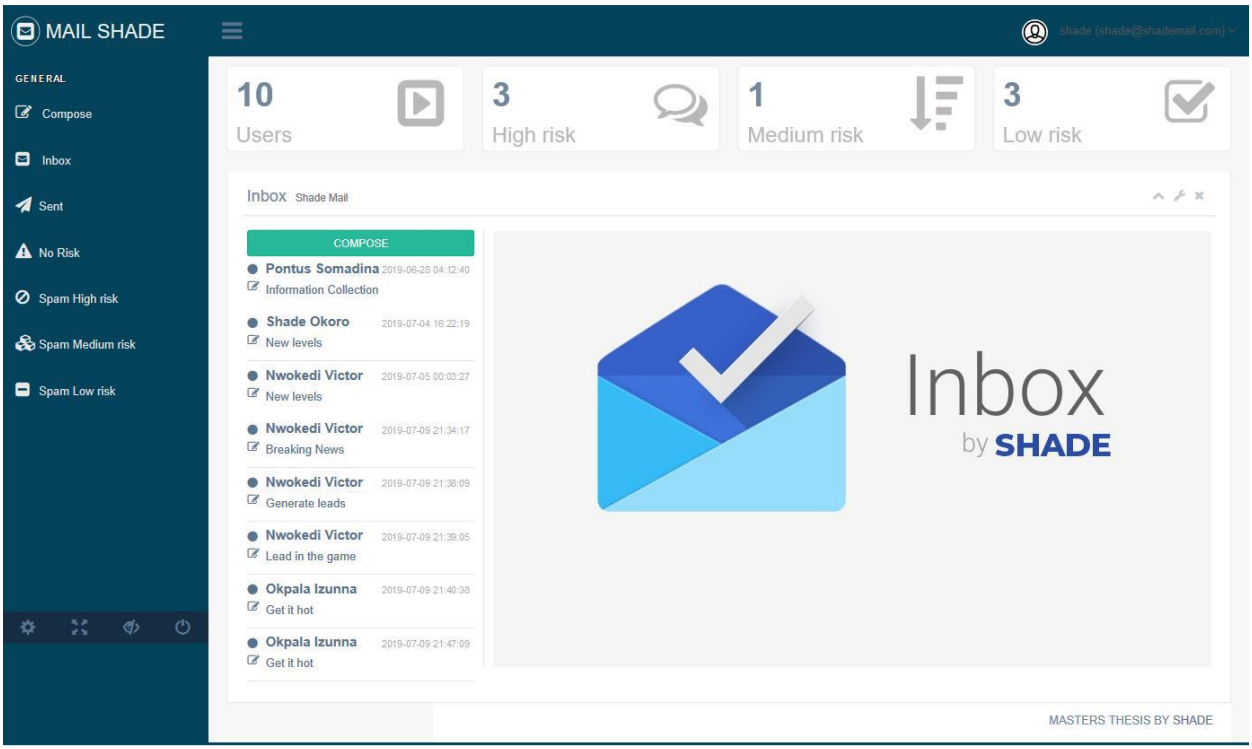

Fig 4: Overall email dashboard

Fig 4 shows the sample of the working app. The user mails are shown, and the app will work on the mails to decipher the phishing emails. The phishing emails are filtered and categorized according to the level of threat while the user will decide the action to take on them.

1. High Threat Email

High threat email that has 90% chance of causing virus infection, harm and possibly death to the victim. When a high threat mail is identified it is advisable for the user to delete the mail completely or report to law officials.

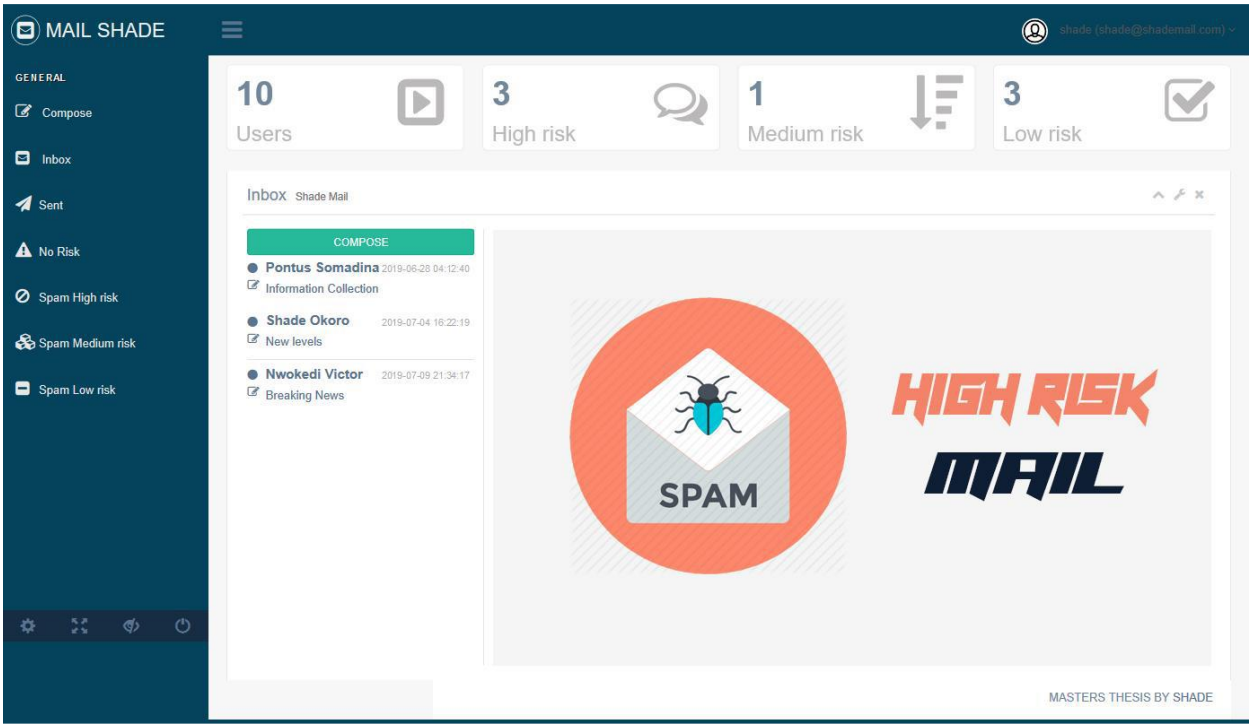

Fig 5: High Risk Interface

Fig 5 shows the layout of the mail categorized as a high-risk mail. The user will determine the next course of action to be done on the mails found in this section.

2. Medium Threat Email

The medium threat email poses a minor risk unlike the high threats. The email is focused on monetary values and ways to gather free advert, plays on people greed and ignorance.

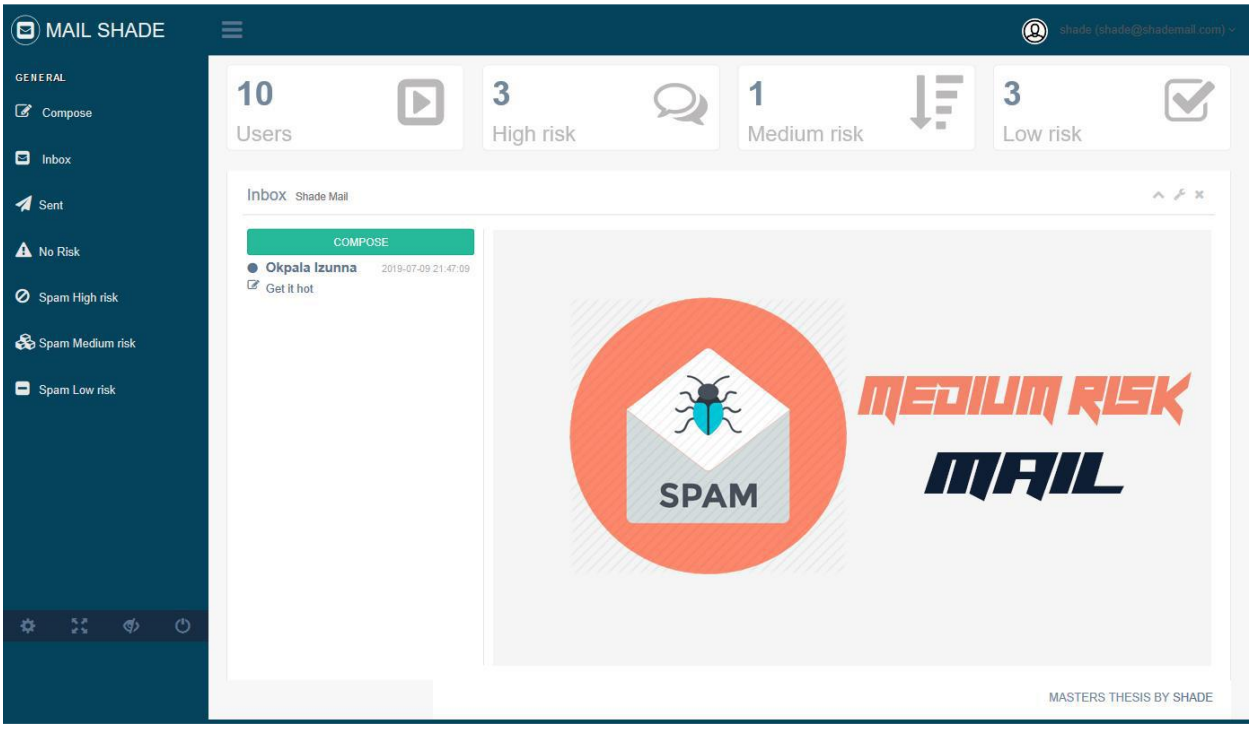

Fig 6: Medium Risk Interface

Fig 6 shows the layout of the mail categorized as a medium-risk mail. The user will determine the next course of action to be done on the mails found in this section.

3. Low Threat Email

The low threat email will have little or no consequences to the target; the target should analyze the originator of the mail.

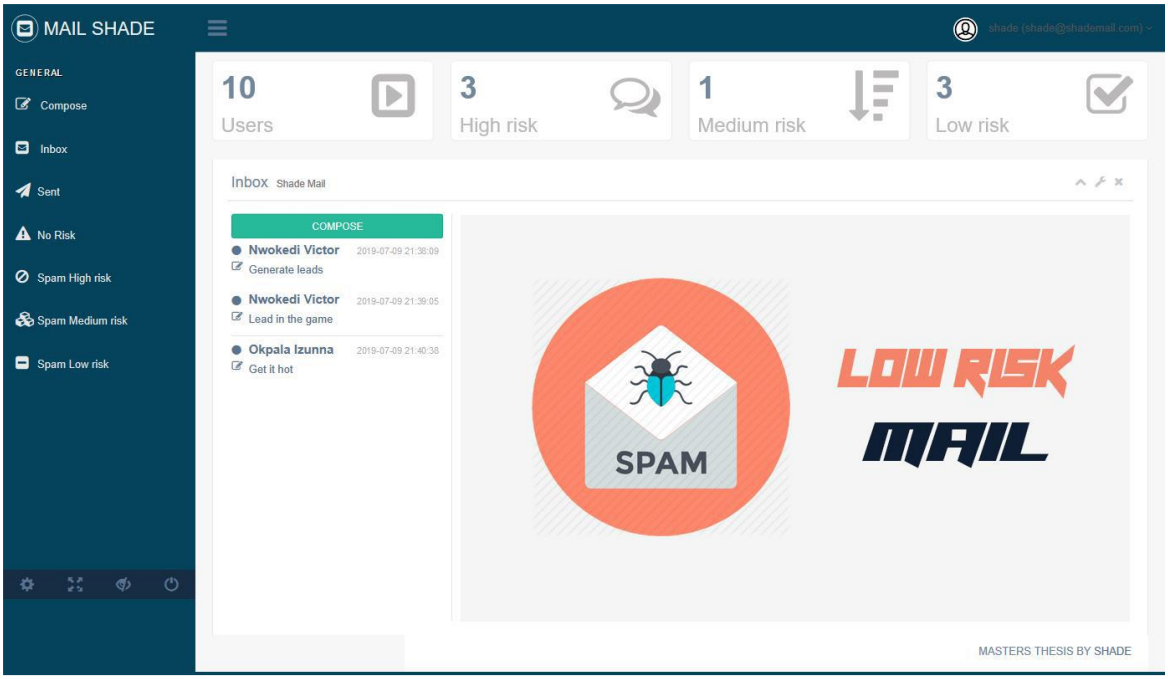

Fig 7: Low Risk Mail

Fig 7 shows the layout of the mail categorized as a low-risk mail. The user will determine the next course of action to be done on the mails found in this section.

- **CATEGORIZATION MODULE**

This module is a programming code developed with PhP which demonstrates how the risk emails are filtered; identified and fetched from email.

```
<?php
include "includes/config.php";
$query_catch= "SELECT * FROM `catch_words`";
$result_catch = $conn->query($query_catch);
$count = 0;
while($rs = $result_catch->fetch_array()){

        $str = $message;
        $sub = $subject;
        $word_count = str_word_count($str);

        if (strpos($str, $rs['words']) !== false) {
                $count++;
                if ($rs['type'] == 3){
```

```
                                        $high = 3;
                                }
                }else if (strpos($sub, $rs['words']) !== false){
                                $count++;
                                if ($rs['type'] == 3){
                                        $high = 3;
                                }

                }
                else{

                }
}

//To determine the rank in which the mail resides, this calculation is done
$percent = ($count/$word_count) * 100;
if (isset($high)) {
                //echo "Highest";
                $risk = 3;
}
else if ($percent == 0) {
                //No risk;
                //echo $percent;
                $risk = 0;
}
else if ($percent <= 30) {
                //lowest risk;
                $risk = 1;
}
else if ($percent >= 50) {
                //High risk
                $risk = 3;
}
else{
                //Medium risk
                $risk = 2;
}

?>
```

iv. CONCLUSION

The process will not eliminate social engineering and phishing attack but will reduce it. The email service provider should make the categorization process to be user friendly; understandable economical; accurate, efficient and possess optimal function capabilities. The research on phishing emails keeps evolving, but this research shows the practical relevance to reducing phishing emails in corporations and the mail filter system according to risk category.

ACKNOWLEDGMENT

We give acknowledge God, our family, our friends, our FUTO colleagues and associates. Many thanks.

REFERENCES

[1] McDermott, J. (2005). *"Social engineering – The Weakest Link in Information Security".*

Retrieved from: http://www.secinf.net/Network_Security/Social-Engineering-TheWeakestLink.html

[2] Hogben, G. (2007) *"Security issues and recommendations for online social networks*," ENISA position paper, vol. 1, 2007.

[3] Cloudflare Developer (2018). *What is phishing attack?*

[4] Emm, D. (2006), *"Phishing update, and how to avoid getting hooked",* Network Security, Vol. 2006 No. 8, pp. 13-15.

[5] Herley, C., & Florencio, D. (2008). *Phishing as a Tragedy of the Commons.* NSPW, Lake Tahoe, CA.

[6] Parsons, K., McCormac., A., Pattinson, M., Butavicius, M., & Jerram, C. (2013). *Phishing for the truth: A scenario-based experiment of usersˆa AZ behavioural response to emails*. Volume 405 of IFIP advances in Information and communication Technology. Pages 366 - 378. Springer Berlin Heidelbegg.

[7] Applegate, S. D. (2009). *"Social engineering: hacking the wetware!"* Information Security Journal: a Global Perspective, vol. 18, 2009, pp. 40-46.

[8] Ellie, W. (2018). *5 ways to reduce the risk of cyber fraud and prevent social engineering*. Retrieved From: https://smile-pass.com/news/5-ways-reduce-risk-cyber-fraud-prevent social engineering

[9] Bermlerand, A., & Levy, H. (2005). "*Spoofing prevention Method,"* INFOCOM'05.

[10] Calluzzo, V. J., & Cante, C. J. (2004). *Ethics in information technology and software use*. Journal of Business Ethics, 51(3): 301–312.

[11] Cormack, G.V. (2008), *"Email spam filtering: a systematic review",* Foundations and Trends in Information Retrieval, Vol. 1 No. 4, pp. 335-455.

[12] Chitrey, A. (2012). *A Comprehensive Study of Social Engineering Based Attacks in India to develop a Conceptual Model*. International Journal of Information and Network Security, 46.

AUTHORS

First Author – Ms. Adu Folashade, B.Sc Computer Science; PGD Information Management Technology, xtiana1202@gmail.com

Second Author – Dr. Ikerionwu Charles, BSc, MSc, MCA, PhD, Certified Novell Administrator (CNA), MIEEE, MCPN, MAOPN, charles.ikerionwu@futo.edu.ng

Third Author – Mr. Nnonyelu Henry, B.Tech Information Management Technology, nnonyeluhenry@gmail.com