ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 11, November 2017

A Study User Behavior Using Phishing Education and Training

[1] Ms. Vidya Mhaske-Dhamdhere, [2] Dr. Prasanna Joeg, [3] Dr. Sandeep Vanjale
[1] PhD Scholar, [2,3] Professor,
Bharati Vidyapeeth Deemed University College of Engg. Pune, Maharashtra .India

Abstract:- **In this paper, we present results of surveys of users on their ability to recognize phishing emails as opposed to legitimate emails. User awareness is one of the most important approaches to phishing detection. We first conducted a survey of user confidence in their ability to detect phishing. We then designed a survey questionnaire that contained some legitimate emails and some phishing emails. We conducted a second survey to test the ability of a sample set of users to detect phishing. The results were mixed. After the second survey, we conducted a training session on the same group of users to train them to detect phishing. After the training we conducted a third survey with a somewhat different questionnaire to test their ability to detect phishing. Our results suggest that training considerably improves the ability to detect phishing. However, a significant minority are still not able to detect phishing.**

I. INTRODUCTION

Phishing is type of computer attack that exchange information through messages to humans via email or web page to perform certain actions for the attacker's benefits. Phishing is an example of social engineering techniques used to financial gain identity thefts. Phishers are targeting the customers of banks and online payment services, Emails. The risks phishing grows even larger in social media such as Facebook, twitter and Google+. The Phishing is typically carried out by email spoofing or instant messaging and it often directs user to enter details at a fake websites whose look and feel are almost identical to the legitimate. In existing system suffer from man –in-middle attack and with help of blacklisting or heuristics technics is not sufficient for zero day attacks. To solve this problem User awareness and to study behavior on internet is most important role. Users tend to ignore warning messages as well as not aware of latest trend of phishing attacks. So continuous training is required.

In section II related work done in this area are discuss. In section III proposed system is explain.in section IV method Tools that aim to combat phishing attacks must take into account how and why people fall fo r them in order to be effective. This study reports a pilot survey of 232 computer users to reveal predictors of falling for phishing emails, as well as trusting legitimate emails. Previous work suggests that people may be vulnerable to phishing schemes be cause their awareness of the risks is not linked to perceived vulnerability or to useful strategies in identifying phishing emails. In this survey, we explore what factors are associated with falling for phishing attacks in a role play exercise. Our data s uggest that deeper understanding of the web environment, such as being able to correctly interpret URLs and understanding what a lock signifies, is associated with less vulnerability to phishing attacks. Perceived severity of the consequences does not pred ict behavior. These results suggest that educational efforts should aim to increase users' intuitive understanding, rather than merely warning them about risks ology and results discuss.

1.1. Definition

Phishing is attempting to attack user to gain personal information. According to phish Tank "phishing is a fraudulent attempt, usually made through email to steal your personal information" [1-2].

II. RELATED WORK DONE IN THIS AREA

Humans are often the weak link in any cybersecurity defense. People behave unpredictably because we are sometimes driven by emotion and by an innate desire to trust and please other people. Also, we tend to take the path of least resistance, even if that path inadvertently creates a cybersecurity risk. Attackers understand these human traits, which is why they are frequently successful in exploiting people to get around more predictable machine-based defenses. As an example, consider phishing. It's that globally, 8 million phishing email messages are opened every day, and of those, 800,000 recipients of the malicious messages click on the embedded links. Ten percent of the people who click on a link actually give their information, such as login credentials for personal applications or their employer’s applications. According to Verizon, 89% of business-related security breaches in 2015 came about as a result of phishing attempts by organized criminals. What’s more, Microsoft estimated in 2014 that the annual worldwide impact of phishing could be as high as $5 billion. Earlier this year, the security awareness company published its which indicates phishing is still a large and growing problem for organizations of all sizes. In the survey behind the report, 60% of respondents reported the rate of phishing attacks has increased overall. Incidents of spear phishing increased 22% from 2014 to 2015, with two-thirds of the survey respondents saying they experienced spear phishing in the past year. While phishing is a serious concern for many organizations – especially those without a strong security awareness program – it’s certainly not the only area where end users tend to be the weakest link. For its report “ Wombat analyzed the answers to nearly 20 million questions asked and answered around various topics in its Security Education Platform over the past two years. These questions are part of different organizations’ end user training and security awareness programs. According to the report, many people still struggle in the following areas: negative reduce 34% to 17%...315 users were still unable to identify between phishing and legitimate websites. Xun Dong, John A. Clark [5] have designed phishing emails of the following two types: recognition of scam as in business opportunities and making easy money. Language, country, student, IP address these parameters are considered. A decision support system is used for user awareness education. The emails are shown offline to the users. 92% of users gave authentication credentials on the phishing website and 7.14% of users gave both personal and authentication credentials. User behavior analysis is done in training and detection mode [6]. Mainly user name, password, domain name, IP address, email id this data is collected from user profile. Most false positive cases are related to some specific websites .with help of predication model users avoid 76.94% phishing threats. Main drawback is that we cannot manipulate user behavior. Steve Sheng[7] has exploring user behavior on browser using prediction techniques . But only context aware user behavior on phishing detection is possible. Phishing educational materials reduced users personal information shared on phishing site or webpages by 40% and 52.3% participants clicked on the simulated phishing emails.

Jussi-Pekka Erkkila has explained users are categories into two categories, which are with in US and outside US. For user phishing education popular training materials are used like anti-phishing Phil, phish guru, cartoon [8]. This training is given continuously till user ignore security warnings and age and gender these two demographic predict the phishing. Users do not care about security warnings.do not use anti phishing tools to detect malicious websites. Pop up warnings produced significantly better results than passive toolbars warnings. Users have lack of knowledge and attenuation. User has false assumption that security is not necessary [9]. Also users are not aware of the structure of computer system and internet. Total 28 days continues training is given some group of users then only 50% users better identified phishing websites and attacks.

Nalin Asanka,Steve Love has discuss on understand relationship between victims background and phishing attacks.to study victims backgrounds user centric design and to identify phishing attacks technology centric design [13] done. Phishers are aware of recent trends and phishing. K. Parsons,A.McCormac has design mobile game to enhance the user’s avoidance behavior through motivation to protect themselves against phishing threats[13]total 10 URL used out of that 4 are real and 6 are phishing .big fish is player who is going to eat small fish. If big fish eat good worm then user gain 10 points which is associate with legitimate URL. If player eat bad worm then users loses 100 point which is indicate phishing URL. Mainly these test 20 participants participated and consider gender age in range 18-25 years, experience using mobile device and average hours per week on the internet. A karakasiliotis ,S.M. Furnell has done experiment on 117 participants to mange 50 emails consisting of 25 genuine emails and 25 phishing emails. The email represented a range of topics such as banking, shopping and social networking and an effort was made to select comparable genuine and phishing email of each topic. For each email participants were asked to respond to the question with one of the four replies [14]

Leave the email in the box and flag for follow up.
Leave the email in inbox.
3. Delete the email.
4 Delete the email and block the sender.

For phishing email 48% and genuine email 60% accuracy. In this paper for experiment 20 mix legitimate and phishing emails taken and asked participant to classify these emails. total 179 participant give responses out of that 36% successfully identified legitimate email and 45% emails classified as phishing emails[15]

III. PROPOSED WORK

We conducted a first survey of a large population to understand their internet usage, their basic knowledge of security, their awareness of phishing.. We then created a set of email examples some of which were legitimate and the rest illegitimate. The examples were designed keeping in in the frequency with which different types of phishing is done. We then conducted a second survey to test the ability of a focused group of users to detect illegitimate emails. After that we conducted training on detecting phishing and then a third survey to test the user's ability to detect phishing again. The content of the survey was derived from findings from an open - ended interview study. The survey consisted of several sections: an email role play where respondents responded to images of emails and web sites, a URL evaluation section where respondents identified features of URLs, a section asking how respondents would react to different warning messages, a knowledge section where respondents interpreted the meaning of lock icons and jargon words, past experience with web sites, and ratings of potential negative consequences of phishing

3.1 General Survey of users (experiment 1)

First we surveyed a broad range of users on the their internet usage, awareness about email phishing, how often they change passwords, how often they tend to click on links embedded in emails and how often they shared personal information to close colleagues. We sent questionnaire to various users such as Students, teachers, IT professionals, and managers,

A total of 4412 user's data was collected. We collected data on their occupation, education, gender, age in range, daily internet usage, and frequencies of sharing passwords and understanding of security

Occupation – 1575 were students which accounted 36 9%, 1541 were management professionals which accounted for 35%, 1227 were IT professionals which accounted for 28 %. 69 were teachers which accounted for 1%. 87% were male and 13 % were females, and education, internet daily usages, age in range [16].

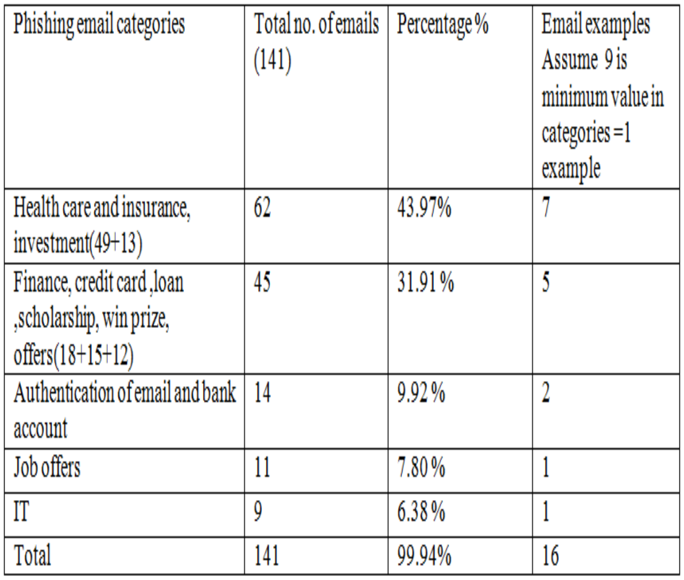
Our findings are as follows:

Only 169 users said they were aware about email and website phishing and 831 users are aware of email phishing and 2066 users are aware of website phishing. Only 17 %(769) users are aware that we have to update bank information personally. 22% (972) users never change password and 42 %(1856) user change password once in a month and 3 %(149) users change password when they forget. only 6 %(268) users never click on email embedded links and 128 %(1217) users always click on email embedded links. [16]

4 Methodology used for designing training phishing emails:

According to [10] phishing email categories, total emails and percentage are shown in below table. On scale of 5, according to percentage each categories of phishing email examples are decided for workshop 1 (before training) and workahop2 (after training). No. of email for each categories decided with help of table 1.

Table1. Phishing email categories (no. of email example for survey 2&3)



Phishing email categories	Total no. of emails (141)	Percentage %	Email examples Assume 9 is minimum value in categories =1 example
Health care and insurance, investment(49+13)	62	43.97%	7
Finance, credit card ,loan ,scholarship, win prize, offers(18+15+12)	45	31.91 %	5
Authentication of email and bank account	14	9.92 %	2
Job offers	11	7.80 %	1
IT	9	6.38 %	1
Total	141	99.94%	16

V. EXPERIMENTAL SETUPS

In this experimental setup, we divide into three parts.experiment 1 is called as before training approach, actual training and lastly after training approach. In this experiment mainly we want to provide user education though online training about phishing emails.

5.1 Experiment 1 (Before Training)

According to [7,9], the following groups are more susceptible to phishing: (1)age group 18 to 25 years, (2) those who do not have a computer education and (3) females with less technical knowledge. For our survey therefore we chose final year and graduate students of computer engineering/IT. The number of participants was 149. In experiment 1 (before training) only 3 legitimate and 13 phishing emails were used. Because our motto is to identify phishing emails by users. This experiment we selected According register media phishing categories as 5 types which are shown in table 1 with percentage of year 2016.also for training total examples for each category also given.

In experiment 1 we shown 16 emails, which is really received on authors emails... we selected participants according to groups 1, and 2 mention and in 5.1 that tis computer engineering final year students. This participants are from age group in range from 18 years to 25 years and they are having computer knowledge .then we gather all participant s those are welling ly interested to participant in this experiment and explain all 3 experiments in depth with motivation .then we show to participants all 16 email s to identification. Also we are provided answer sheet hard copy with question numbers .while showing email we told participant to write answer as three options which is fixed like option A is phishing .option B legitimate, and option C.

We show 16 emails out of that 3 are legitimate and 13 are phishing emails.after showing emails we told participant to write down answer as mentions above.

5.2 Experiment 2 (actual training for identification of phishing emails)

In this experiment we continue all participant as it is for training .first we discuss all emails ,which is shown in before training ask answer orally and explain why particular email is phishing or legitimate, also how you are going to identify email is legitimate and phishing. Also does and don't. we give phishing email identification tips like to check sender is unknown, observed email header, email subject, email footer, email language,multi-color in email,email is embedded with link or not if yes then weather asking personal information like bank details and user name and password. Also some online email [11-12] examples are discussed.

5.3 Experiment 3(after phishing email identification training)

In experiment participant are very excited to identification of emails. To check confidence level of participant on scale of 5. We summarized and consolidated and generated graph for before and after training confidence level of participant, which is shown in fig.1.

For experiment 3 we selected 16 emails according to table1 also we repeated 40% email of before training as it is and 60% emails are newly introduced.

Experiment 1 and experiment 3 participants responses are notes and compared which is shown in result as next section.

6 Experimental 1 and 3 results comparison

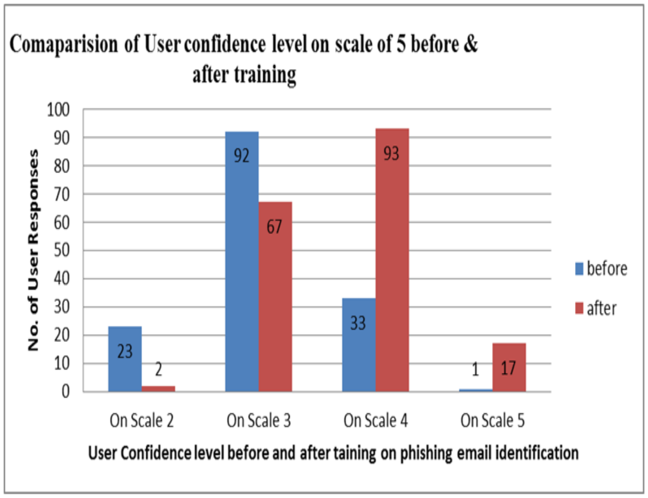


Fig.1 Confidence level of participant before and after training

We ask participant confidence level 1 and 3 on scale of 5. We summarized and consolidated and generated graph for before and after training confidence level of participant, which is shown in fig.1.in experiment 1 92 participant confidence levels is 3.but after training 67 participant confidence levels is 3 and 93 participant confidence level is 4. So it indicates that after training confidence level of identification of phishing meal is increases Only 49%before training and 38% after training legitimate emails are classify by participant which is less. Before training 42% and after training 59% phishing emails are identified.it indicated that after training phishing email identification is increased by 17%.

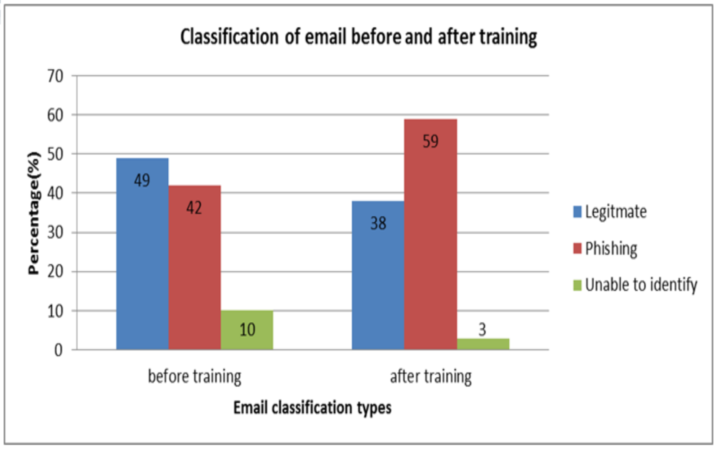


Fig.2 Before training legitimate email correctly and incorrectly classify in percentage.

As before training approach phishing and legitimate email correctly identification performance is less than 50% .so that we apply training approach for same participant we repeat ion of 40% email example of before training and see performance of users. In training all example which is included in before training with do's and don'ts are explain also phishing identification tips are given.

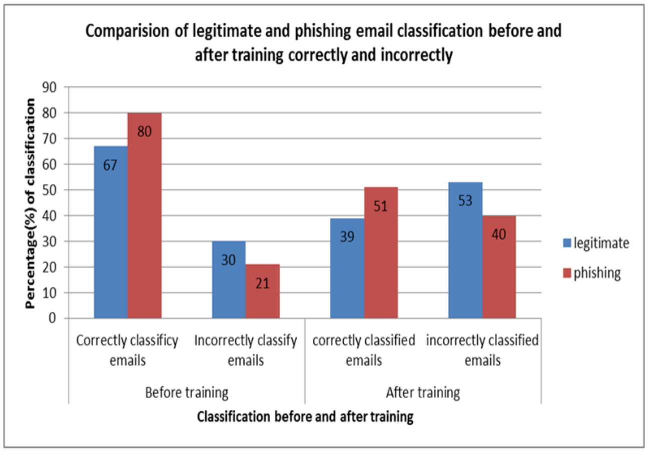


Figure.3 comparison of Phishing email classified as phishing correctly before and after training

After training 67 % users correctly identify legitimate email and 80 % phishing emails are identified. If we compare before and after training approach only 28% users legitimate email correctly identification is improvement and 39% phishing email identification improvement ,which is very less so that we required to solve this problem machine learning algorithms are required. After training we take review of users why they incorrectly classify legitimate email as phishing and phishing email as legitimate. They give reason like multicolor are used in email, email embedded URL is given, sender is unknown, email signature is not proper, domain and subdomain is not register.

According to reason given by participant we categories as follows.

Type1 –visual look and feel of email

In this type of email different colors are used as well as email signature is formal and organization logo is same.

Type 2- technical parameters of email

In this categories email embedded link is given, there is no https in URL, sender email address is known.

Type 3- email header and body

Personalized data ask in email, typing mistake or error, promoting offers, us of urgent or force full language.

Almost 90 % to 97% phishing emails are categories belong from type 2 and type 3.and only type 1 only 3% emails are phishing.

VII. CONCLUSIONS AND FUTURE WORK

User awareness about email and websites phishing is one of most important aspects. Existing literature survey user education was done online or offline. User education should provide continuously. In existing user education targeted only age group 18 to 25 years old, gender, and country, which was not sufficient parameter analysis the performance of user.to fill up this research gap we are going to include more parameter like age in different range, education, profession, daily work internet usages.

If we compare before and after training approach only 28% users legitimate email correctly identification is improvement and 39% phishing email identification improvement ,which is very less so that we required to solve this problem machine learning algorithms are required.

In future we are going to apply machine learning algorithm to detect phishing emails .also we are going to design simulated phishing email attacks for users training.

•

someone stole yo ur social security number or your identity your computer automatically sent bad software to everyone in your online address book someone you didn't know could see everything that you typed on your computer your computer started to crash several times a day 3. Most participants (85%) were PC users, and most of the rest (12%) were Mac users. Almost all (98%) respondents had made a purchase on the web at some point in the past. Most had also engaged to some degree with altering the prote ctions on their computer, such as installing anti virus software (93%) or adjusting their security preferences (79%), although the latter could be making the security settings tighter or more lax. A minority had experienced some negative consequences of t he type threatened by phishing, although not necessarily as a result of phishing. They reported having had a credit card stolen (21%), having had information stolen or compromised (14%), and a small number reported having had their social security number s tolen or had someone try to steal their identity (3%)

REFERENCES

[1] M. Khonji, Y. Iraqi, and A. Jones, "Phishing detection: a literature survey," Communications Surveys & Tutorials, IEEE, vol. 15, pp. 2091-2121, 2013.

[2] Ammar Almomani, B. B. Gupta, Samer Atawneh, A. Meulenberg, and Eman Almomani "A Survey of Phishing Email Filtering Techniques" IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 15, NO. 4, FOURTH QUARTER 2013 page no 2070 -2090.

[3] Tzipora Halevi, James Lewis," A Pilot Study of Cyber Security and Privacy Related Behavior and Personality Traits", International World Wide Web Conference Committee (IW3C2), May 13–17, 2013, ACM.

[4] Vishkha pawar, pritish tijare ," User Security Awareness against Phishing", International Journal of Advance Research in Computer Science and Management Studies, Volume 2, Issue 3, March 2014.

[5] Xun Dong, John A. Clark, User Behavior Based Phishing Websites Detection", Proceedings of the International Multiconference on Computer Science and Information Technology pp. 783–790 IEEE-2008.

[6] Lung-Hao Lee, Kuei-Ching Lee,Yen-Cheng Juan," Users' Behavioral Prediction for Phishing Detection", ACM 978-1-4503-2745-9/14/04. April 7–11, 2014.

[7] Steve Sheng,1 Mandy ,Holbrook,Ponnurangam Kumaraguru, Lorrie Cranor,Julie Downs," Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions", Copyright 2010 ACM.

[8] Jussi-Pekka Erkkila," Why we fall for Phishing", Copyright 2011 ACM.

[9] Ali Darwish, Ahmed El Zarka and Fadi Aloul," Towards Understanding Phishing Victims' Profile2013 IEEE.

[10] https://regmedia.co.uk/2016/05/12/dbir_2016.pdf

[11]https://blog.returnpath.com/10-tips-on-how-to-identify-a-phishing-or-spoofing- email-v2

[12] http://www.techrepublic.com/blog/10-things/10-tips-for-spotting-a-phishing-email/

[13] Nalin Asanka Gamagedara Arachchilage , Steve Love, Konstantin Beznosov," Phishing threat avoidance behaviour: An empirical investigation", Computers in Human Behavior 60 (2016) 185e197, 0747-5632/© 2016 Elsevier Ltd

[14] K. Parsons1, A. McCormac, M. Pattinson, M. Butavicius1 and C. Jerram," Using Actions and Intentions to Evaluate Categorical Responses to Phishing and Genuine Emails", Proceedings of the Eighth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2014)

[15] A Karakasiliotis, S M. Furnell,"Assessing end-user awareness of social engineering and phishing", Australian Information Warfare and Security conference, 2006.

[16] Vidya Mhaske Dhamdhere,Prasanna Joeg," To Study of phishing attacks and user behavior", IEEE,2ND INTERNATIONAL CONFERENCE ON INVENTIVE COMPUTATION TECHNOLOGIES,2017