International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 9 (2019) pp. 2171-2175

A Review on Phishing Attacks

Akarshita Shankar
Computer Science Engineering,
RV College of Engineering Mysore Road,
RV Vidyaniketan Post,
Bengaluru, Karnataka - 560059, India.

Ramesh Shetty
Cyber Security and Risk Services,
Wipro Limited, No 72, Keonics, Hosur Rd, Electronic City, Bengaluru,
Karnataka - 560100, India.

Badari Nath K
Computer Science Engineering,
RV College of Engineering Mysore Road,
RV Vidyaniketan Post,
Bengaluru, Karnataka-560059, India.

Abstract:

Phishing is a cybercrime, which involves luring the user into providing sensitive and confidential informa- tion to the attacker. The information could include credit card details, username and passwords, bank details, etc. These phishing attacks occur through malicious emails, text messages and telephone calls. After obtaining the information, the attacker could commit crimes such as financial losses and identity thefts. The target could be an individual, an organization or a cluster in an organization. This paper provides an explanation on phishing attacks to create awareness and several countermeasures to overcome them.

Keywords: Phishing, deceptive and spear phishing, whaling, pharming, anti-phishing techniques.

I. INTRODUCTION

A field of Information Technology is Cyber Security that aims at the protection of data, systems, network, etc., from the various attacks. Cyber Security is one of the key concerns in today's information technology world. It also aims at the prevention of unauthorized access to sensitive data. Data is vulnerable to various attacks while in transit and while stored. These attacks, both existing and upcoming, pose a great threat to industries and individuals. Since, industries rely heavily on computers for their functionalities, confidential and sensitive information needs to be protected. Various Cyber Security techniques and tools provide this protection of data while it is stored and in transit.

The threats and vulnerabilities, along with the cost of detecting and fixing the bugs, have consistently been increasing over the past two decades. This has led to loss of intellectual property, loss of reputation and revenue, public exposure of security vulnerabilities, etc. The cost encountered for overcoming such security attacks have escalated from $27.4 billion to $66 billion in eight years.

The attackers use various methods like loophole in applications as gateway to exploit the vulnerabilities, which help them to gain unauthorized access to sensitive data.

Phishing involves sending fraudulent emails to a target that appear to come from a creditable source. An individual or several individuals known as phishers or attackers orchestrate the attack. The individuals who are affected by the attack are called victims or targets. The goal of phishing is to gather sensitive data, such as login credentials or bank account details or install malware into the target's system. Investigating such a complex attack is very challenging to the cybersecurity experts. Phishing attacks can be performed manually but to over- come the attack and to respond effectively to the attack requires a lot of time, intelligence and manpower. This may take days or even weeks to respond and analyze the attack in depth. Manual investigation has lot of dependency on the security analyst's talents and tools avail- able for investigation. Moreover, these manual investigations go wrong due to human errors.

Commonly known as the Amazon Prime Day phishing attack, the information of the customers of Amazon was compromised by a phishing attack. All the Amazon Prime members received an email that consisted of seemingly legitimate deals to them. On trying to purchase the 'deals', the transaction would fail, promoting the attackers to gain sensitive information on the user.

Another common example is Google Docs invitation. In May 2017, attackers sent fraudulent invitations to google users across the world to edit documents. When the recipients clicked the invitation, it led to a third party app that facilitated attackers to obtain confidential information.

II. PHISHING MECHANISM

The motive of phishing attack is to manipulate the attacker into providing confidential information about him/her. To perform such an attack, the attacker or phisher mimics a legitimate website. To mimic the web- site, he/she constructs a malicious site using a phishing website. This phishing

website would gather all the in- formation on the target and provide it to the attacker [11]. Usually, the targets are unable to distinguish between genuine and phishing websites causing them to fall into the traps set by the phisher.

Phishing attacks have several steps that attackers follow to obtain information. This can be explained in six steps. The steps are as follows:

- Plan
- Compose email
- Attack
- Gather data
- Fraud

The attacker starts the process by planning the attack. This step involves in deciding the legitimate website that has to be imitated and the victim whose information has to be gathered. Followed by planning, is composing an email that has to appear genuine for the victim to be lured into providing his/her data. The third part is sending the composed email to the target followed by gathering the information on the victim. The gathering of in- formation phase occurs only if the victim has been tricked by the phisher. Using the victim's information, the attacker commits cybercrimes such as credit card fraud, theft, etc.

Figure 2.1 shows the phishing mechanism and how the attackers manage to collect sensitive information about the target. It shows several steps involved in the attack.

Step 1 shows that the phisher (attacker) composes an email with the help of a phishing website. This email is composed such that it appears to be genuine and legitimate. In step 2, the attacker sends this composed email to the victim (target). Step 3 indicates that the victim, unable to differentiate between genuine emails and phishing emails, tends to open the email. The email then directs her to the phishing website. The victim enters her login credentials in the webpage oblivious of the fact that it is a malicious site. The phishing website then pro- vides the login credentials to the attacker. This is illustrated in step 4. In the last step, the phisher, using the data he has obtained from the phishing website, logs into the target website. Now, he would be able to access all the information of the victim. Thus, the process of phishing is completed.

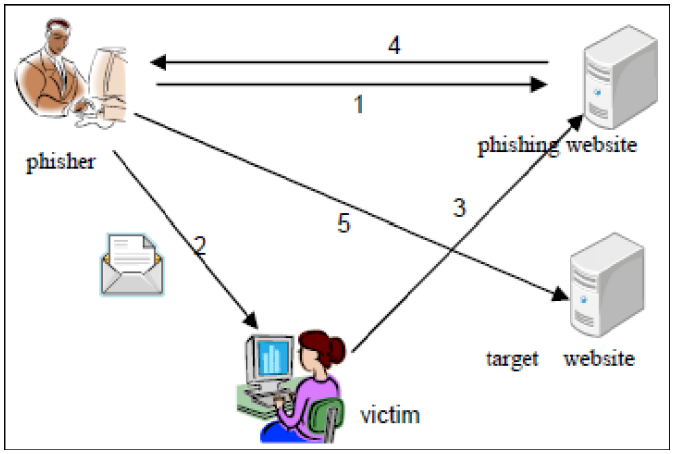

Figure 2.1 Phishing Attack Mechanism

(Source: www.researchgate.net/figure/steps-in-a-deceptive-phishing-attack)

III. TYPES OF PHISHING ATTACKS

The attack can be performed in several ways. The mo- tive for all the kinds are the same. The only variation amongst the types are the number of targets and the mechanism used to obtain the data.

The various types of phishing attacks are [12]:

- Deceptive Phishing
- Spear Phishing
- Whaling
- Pharming
- *Deceptive Phishing*

Deceptive phishing is the most prevalent type of phishing attack. It involves imitating a legitimate website and sending an email to the target appearing it to be genuine. The email sent would contain a malicious URL or link. It would instruct the target to click on the URL. Upon following the instruction, the phishing website gathers all the login credentials and other sensitive information about the target and forwards it to the attacker [11].

For example, testuser@amazon.com uses a lowercase 'a' that could be removed. Hence, testuser@mazon.com could trick the target and thereby obtain data.

- *Spear Phishing*

This variety of phishing is nearly identical to deceptive phishing. The only difference is the target. Unlike deceptive

phishing, spear phishing targets one individual only. The attacker aims at one person and lures him/her into providing confidential data. The fraudsters customize the email according to the individual. The email would consist some of the target's information such as the person's name, company he/she is working in, designation, etc. The most common platforms where spear phishing takes place is social media sites such as LinkedIn where it is easy for them to obtain information on the individual's profession [2].

- *Whaling*

Whaling attacks occur when the phisher targets an individual at an executive position like CEO. The attacker would be profiling the victim for a considerable period before performing the attack. The attacker, similar to other types, would send an email to the target and manipulate him/her into providing information to the attacker. Whaling is considered a very dangerous attack since the people in executive bands have access to the organization's most confidential information.

- *Pharming*

Pharming is another variation of phishing. Unlike, the other techniques, it is not necessary to target individuals. The attack can victimize a large number of people with- out having to be targeted individually.

Pharming can be performed in two ways:

i) The first method involves a code that is sent to the target via email that modifies all the local host files in the system. The URLs would be converted by the host files to number strings, used by the system to access web sites. This causes the target to be redirected to the malicious site in spite of entering the correct URL.

ii) The second method of performing pharming is through a technique called DNS Poisoning. In this method, the system's local host files are not corrupted but the domain name system table is modified. This results in the target being redirected to malicious websites without their knowledge. The target would be assuming they are accessing the legitimate websites, but because of DNS Poisoning, they would be accessing the malicious website.

Thus, the motive for all the variations of phishing are the same. Only the method and the technique used to obtain the information varies from one type to another.

IV. ANTI-PHISHING TECHNIQUES

In [1], a solution was proposed where Automated Individual White-List (AIWL), an automated list, attempts to maintain a white list that consists of every familiar Lo- gin User Interfaces (LUI) of the user's. When a user submits his/her login credentials or sensitive information to a LUI that is missing from the white-list, AIWL will warn the user of the possible trap and will warn him/her of the consequent attack.

In [2], the authors proposed a solution to defend phishing attacks using a combination of visual similarity based techniques and white list. The Computer Vision (CV) tool called Speed up Robust Features (SURF) detector. This detector uses square shaped filters for extracting discriminative key point features. These features are extracted from both – suspicious and genuine web- sites. The features extracted from the websites are then compared for calculating a similarity degree. The similarity degree then helps in determining if the website is legitimate or not. If the similarity degree was high, it was considered malicious since the legitimate website was trying to be imitated.

In [3], a different solution was proposed by the use of Support Vector Machines (SVM) to detect if the mail is malicious or not. The SVM extracted common characteristics of the mail such as language used, layout of the mail, structure of the mail, etc. It then compares the extracted details with the details present in the system to check the similarity accuracy. If the accuracy exceeds a certain threshold, it marks the email as malicious.

The study conducted in [4] used a unique technique of Natural Language Processing (NLP) to determine if the mail was malicious or not. In this paper, they extracted and compared common characteristics using NLP tools. PhishNet-NLP utilized natural language techniques along with all information present in an email, namely the header, links, and text in the body. PhishSnag used information extracted from the email to detect phishing. Phish-Sem used NLP and statistical analysis on the body for labelling the mail as phishing or non-phishing.

A more advanced technique of filtering and classification was used in [5]. In this paper, the authors tested the URLs and verified whether it was malicious or not. They used an automated approach for detecting phishing. It had two phases- Pre-filtering and Classification phase. In the pre-filtering phase, the URL was compared against a black list using the domain part of the URL. If the URL was present in that list then it was classified as malicious and would not be proceeding to the Classification Phase. In the next phase, two main features were checked for consistency- Randomness of the URL (RU) and the Position of the domain token. Based on

the results of Classification Phase, the URL was classified as malicious or non-malicious.

In [6], the authors used text mining to extract distinct features from emails. The emails could be phishing or genuine emails for better detection of the attack. The strategy followed was an initial conversion of the email to a vector representation followed by feature selection techniques for classification. The evaluation was per- formed using data sets accumulated from the HamCorpus of SpamAssassin project (legitimate e-mail) and the publicly available PhishingCorpus (phishing e-mail).

The extraction and classification method was further developed in [7]. In this paper, the vulnerabilities were differentiated into three categories based on the structure of the email. The three categories were Page-content vulnerability, Domain vulnerability and Code-scripting vulnerability. The evaluator model used was Anti-Phishing Effectiveness Evaluator Model (APEE Model) which is used to analyze the effectiveness of the Anti-Phishing Mechanisms that have been implemented. The reputation of the vulnerabilities from the three categories are tested which help in determining whether the mail is a phishing email or not.

The method used in [8] is marginally different from the other techniques. In [8], they classify the mails as junk or not junk based on the spam filter. When an email arrives to the mailbox, the spam filter performs its filtering function and verifies if the mail is spam or not. The spam filtering is performed based on the reputation of the URL present in the mail. If the URL seems to be unsafe or suspicious, the filter marks the mail as 'junk'. The URL(s) in the mail are deactivated and the mail is then moved to junk. If the mail is genuine, the mail is moved to inbox for the user to open it safely.

In [9], Anti-phishing technique was developed with the help of advanced heuristic approach. In this technique, when a suspicious website was encountered, it was immediately updated in the black list. If a legitimate web- site is found, it updates the same in the white list. There- fore, when the user open a website, it was first verified if the website was a phishing website or not and accordingly provided access to the same. This technique used PHP Programming Language along with a Database to maintain the two lists. According to this technique used, 2519 URLs were tested and 2510 were correctly classified.

The authors talk about reusable components for anti-phishing components layer in [10]. These reusable components are used for converting webpages to feature vectors using heuristic methods and external repositories. The finite feature vectors that provide as input to these vector machines train the support vector machine. With the training provided by these inputs, the support vector machine classified and determined various web pages as legitimate or a phishing web page. This was experimented with the mixture of heuristics in identifying a phishing webpage.

V. CONCLUSION

Phishing is a technique to gather sensitive information about the target using malicious links and emails. It is one of the most dangerous cyber-attacks that occurs in organizations, personal devices, etc. It is often difficult to distinguish between genuine emails and phishing emails. There are several methods that can be used to avoid this attack. Periodical updating of anti-phishing tools and platforms can prove to be very powerful. This study provides an in-sight to phishing, the mechanism of the attack, various forms it can occur in and the possible solutions to overcome them.

REFERENCES

[1] Ye Cao, Weili Han and Yueran Le - Anti-phishing based on automated individual white-list, *Proceedings of the 4th ACM workshop on Digital Identity Management*, pp. 51-60, October 2008.

[2] Routhu Srinivasa Rao and Syed Taqi Ali - A Computer Vision Technique to Detect Phishing Attacks, *5th International Conference on Communication Systems and Network Technologies, IEEE*, October 2015.

[3] Madhusudhanan Chandrasekaran, Krishnan Narayanan and Shambhu Upadhyaya - Phishing E-mail Detection based on Structural Properties, *IEEE*, November 2015.

[4] Rakesh Verma, Narasimha Karpoor, Nabil Hossain and Nirmala Rai - Automatic Phishing Email Detection based on Natural Language Processing Techniques, *Research Gate*, 2016.

[5] Yi-Shin Chen, Huei-Sin Liu, Yi-Hsuan Yu and Pang-Chieh Wang, Detect Phishing by Checking Content Consistency, *IEEE*, 2017.

[6] Masoumeh Zareapoor, K.R. Seeja, Text Mining for Phishing E-mail Detection, *Intelligent Computing, Communication and Devices: Advances in Intelligent Systems and Computing*, vol. 308, pp. 65-71, August 2016.

[7] Sankhwar S., Pandey D., Khan R.A - A Novel Anti-phishing Effectiveness Evaluator Model, *Smart Innovation, Systems and Technologies*, *Springer*, vol 84, Cham, 2018.

[8] Xavier Joseph, Mitchell Aime M, Tsang Brian J, Herbert, George A, Savastano, Hernan I, Khandelwal Lubdha, Pengelly Robert C. J, Novitskey Robert, Grant Stanley - Anti-phishing protection, United States Patent 10,069,865 B2, Sept 4th, 2018.

[9] Okunoye, O.B, Azeez, N.A, Ilurimi F.A - A Web Enabled Anti-Phishing Solution Using Enhanced Heuristic Based Technique, *FUTA Journal of Research in Sciences*, vol. 13 (2), pp. 304-321, October – 2017.

[10] Anna L. Buczak, Erhan Guven - A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection, *IEEE Communications Surveys & Tutorials*, vol. 18, pp. 1153-1176, June – 2016.

[11] Dr. Radha Damodaram - Study on Phishing Attacks and Anti-Phishing tools, *International Research Journal of Engineering and Technology (IRJET)*, vol. 3, pp. 700-705, January – 2016.

[12] Gaurav, Madhuresh Mishra, Anurag Jain - Anti-Phishing Techniques: A Review, *International Journal of Engineering Research and Applications (IJERA)*, vol. 2, pp. 350-355, April – 2012.