RESEARCH ARTICLE OPEN ACCESS

A Survey on Phishing and Anti-Phishing Techniques

L. Joy Singh, NIELIT IMPHAL

ABSTRACT

Phishing is one of the major online frauds used to collect personal information illegally through spamming or other deceptive means. It is an online theft of sensitive information that swindles innocent users into disclosing secret information such as user names, passwords, credit card numbersand other fraudulent activity. Phishing results negative impact on the economy through financial losses experienced by businesses and consumers, along with the adverse effect of decreasing consumer confidence in e-commerce. This paper gives brief information about existing phishing attacks including deceptive, malware-based, SMiShing, Vishing, PopUp window, web-Trojan, key loggers, cross site scripting, session hijacking, URL Obsifucation, Dos attack, Pharming, BotNet, Man-in-Middle attack, Search engine attack, etc. are highlighted and existing Anti-Phishing techniques based on OTP, CAPTCHA, HTTP secure, Digital certificate, Identity, Attribute, Genetic, Character and Content are also discussed. Each of existingAnti-Phishing techniques has both advantages and disadvantages. In this paper, a new Anti-Phishing techniquehas been proposed which can be used by directly installingto the workstations which is more cost effective and easy to maintain. It can avoid for installation of separate Anti-Phishing techniques & firewalls to all the client nodes within the specific network.

Keywords:- Phishing attacks, Anti phishing technologies, Countermeasures.

I. INTRODUCTION

Phishing is a process of procuring confidential information like usernames and passwords by fraud means like sending legitimate e-mails and fooling the unaware users. Mailing systems have evolved in a big way by filtering out spam to an extent. But, it may happen that a legitimate looking e-mail, asking for confidential details like usernames, passwords, social security numbers, address etc.

In this paper, we will identify several of the technical capabilities that are used to conduct phishing scams, review the trends in these capabilities over the past several years, and discuss currently deployed countermeasures. In general, phishing attacks are performed with the following four steps:

1) A fake web site which looks exactly like the legitimate Web site is set up by phisher

2) Phisher then send link to the fake web site in large amount of spoofed e-mails to target users in the name of legitimate companies and organizations, trying to convince the potential victims to visit their web sites.

3) Victims visit the fake web site by clicking on the link and input its useful information there.

4) Phishers then steal the personal information and perform their fraud such as transferring money from the victims' account.

II. CLASSIFICATION OF PHISHING ATTACKS

Phishing attacks can be classified into various types according to the way attack is done. According to many researchers, the various types of phishing attacks has been described below.

1. Deceptive Phishing- The word "phishing" initially referred to report robbery using immediate messaging but the most ordinary transmit method today is a misleading email message. Messages about the require to confirm account information, system failure requiring users to re-enter their information, fictitious account charges, unwanted account changes, new free services requiring quick action, and many other scams are broadcast to a broad collection of recipients with the hope that the

unsuspecting will react by clicking a link to or signing onto a bogus site where their secret information can be collected.

2. Malware-Based Phishing- Malware-based phishing involves running malicious software on the user's machine. The malware can be introduced as an email attachment or as a downloadable file exploiting security vulnerabilities. This is a particular threat for small and medium businesses (SMBs) who fails to update their software applications.

3.SMiShing: It is a form of Phishing that uses short messaging services (SMS) or text messages on mobile phones and Smartphone's. There are two main processes for the SMiShing scams; one involves receiving a text message which is purported to have originated from a known and trusted source, such as your bankers or your system administrator. The second one involves you receiving a vital text message about your identity been stolen or account number been frozen, it then goes ahead to direct you to a website or a phone number for the verification of the account information.

4. Vishing:"Vishing is the practice of leveraging IP-based voice messaging technologies (primarily Voice over Internet Protocol, or VoIP) to socially engineer the intended victim into providing personal, financial or other confidential information for the purpose of financial reward[4]. The term "Vishing" is derived from a combination of "voice" and "Phishing."

5. Popup window: Popup windows are also used to steal the user's information. There are two ways to
activate popup windows, one of them is used to confirm something and are written in the HTML code like this < … onClick="window.open ('mypage.html')"> and it is by legal window and html. The other way to activate popup window is illegal because it is a JavaScript file used like: "Open Popup" onClick="javascript: Popup ('mypage.html')">.

6. Web Trojans-Web Trojans pop up when the users attempt to log in to an important website or performing any transaction. These web trojans are invisible to the users. They collect user's credentials locally and transmit them to the phisher.

7. Key loggers and Screen loggers: Key loggers and screen loggers are varieties of malware that track input from the keyboard and send relevant information to the hacker via the Internet. They can embed themselves into the user's browsers as small utility programs.

8. Cross Site Scripting: Cross-site scripting (XSS) is a type of computer security vulnerability typically found in web applications which allow code injection by malicious web users into the web pages viewed by other users. Examples are client-side scripts and HTML codes.

9. Session Hijacking:Session Hijacking is a kind of phishing attack where user's activities are monitored clearly until they log into a target account like the bank account and establish their credentials. At that point, the malicious software takes control and can undertake unauthorized actions, such as transferring funds, without the knowledge of the user.

10. URL OBSIFUCATION: - When a client types a URL to visit a website it must first be translated into an IP address before it's transmitted over the Internet. The bulk of SMB users' PCs running a Microsoft Windows operating system first look up these "host names" in their "hosts" file before responsibility aDomain Name System (DNS) lookup. By "poisoning" the hosts file, hackers have a bogus address transmitted, taking the user unwittingly to a fake "look alike" website where their information can be stolen.

11. Denial of service (DoS) Attack: This is an act by the criminal, who floods the bandwidth of the victims network

or fills his e-mail box with spam mail depriving him of the services he is entitled to access or provide Short for denial-of-service attack, a type of attack on a network that is designed to bring the network to its knees by flooding it with useless traffic.

12. Clone Phishing-In this type phisher creates a cloned email. He does this by getting information such as content and recipient addresses from a legitimate email which was delivered previously, then he sends the same email with links replaced by malicious ones. He also employs address spoofing so that the email appears to be from the original sender. The email can claim to be a re-send of the original or an updated version as a trapping strategy.

13. Spear Phishing-Spear phishing targets at a specific group. So instead of casting out thousands of emails randomly, spear phishers target selected groups of people with something in common, for example people from the same organization. Spear phishing is also being used against high-level targets, in a type of attack called "whaling".

14. BOTNET: Bot is a new type of malware installed into acompromised computer which can be controlled remotelyby BotMaster for executing some orders through thereceived commands . A botnet is a group ofcompromised computers (Bots), which are remotelycontrolled by attackers (BotMasters) to launch variousnetwork attacks such as distributed denial of service(DDOS), malware dissemination, phishing, and clickfraud .Botnets can also be used to spread other botnets in the network. It does this by convincing the user to download after which the program is executed through FTP, HTTP or email.

15.Web Spoofing-A phisher could forge a website that looks similar to a legitimate website, so that victims may think this is the genuine website and enter their passwords and personal information, which is collected by the phisher. Modern web browsers have certain built-in security indicators that can protect users from phishing scams, including domain name highlighting and https indicators.

16.E-mail spoofing-Email spoofing is email activity in which the sender address and other parts of the email header are altered to appear as though the email originated from a various source. Distributors of spam often use spoofing in an attempt to get recipients to open and possibly even respond to their solicitations.

17.DNS-Based Phishing ("Pharming") or Host file poisioning - With a pharming scheme, hackers tamper with a company's hosts files or (DNS) domain name system so that requests for URLs or name service return a bogus address and subsequent communications are directed to a fake site.

18. System Reconfiguration Attacks:

This is a kind of **phishing** attack where the settings on a user's PC are modified with bad intentions. For example: URLs in a favourites file might be modified to direct users to bogus websites that look alike. For example: a financial institution's website URL may be changed from "bankofxyz.com" to "bancofxyz.com".

19. Content-Injection Phishing-Content-injection phishing means inserting malicious content into a legitimate website. The malicious content can redirect to other websites or may install malware on a user's computer and also insert a frame of content that will redirect data to the phishing server.

20. Man-in-the-Middle Phishing-Man-in-the-Middle Phishing is hard to detect than many other forms of phishing. In these attacks hackers sit between the user and the website or the system. They record the information being entered by the user but continue to pass the user on to the next steps so that user transactions are not affected and the user remains unaware.

21. Search Engine Phishing- Occurs when phishers create websites with attractive (often too attractive) sounding offers and help them indexed legitimately with search engines. Users find the sites in the normal course of searching for products or services and are fooled into giving up their information.

22. Phone phishing-

This type of phishing refers to messages that claim to be from a bank asking users to dial a phone number regarding problems with their bank accounts. Traditional phone equipment has dedicated lines, so Voice over IP, being easy to manipulate, becomes a good choice for the phisher. Once the phone number, owned by the phisher and provided by a VoIP service, is dialed, voice prompts tell the caller to enter her account numbers and PIN. Caller ID spoofing, which is not prohibited by law, can be used along with this so that the call appears to be from a trusted source.

III. ANTI-PHISHING TECHNIQUES

Anti-phishing refers to the method employed in order to detect and prevent phishing attacks. Anti-phishing protects users from phishing.Anti-phishing software consists of computer programs that attempt to identify phishing content contained in websites and e-mail. It is often integrated with web browsers and email clients as a toolbar that displays the real domain name for the website the viewer is visiting, in an attempt to prevent fraudulent websites from masquerading as other legitimate web sites. Anti-phishing functionality may also be included as a built-in capability of some web browsers.

A. One time password -One-time passwords can be generated in several ways and each one has different benefits in terms of security, convenience, cost and accuracy. A more convenient way for users is to use an OTP token which is a hardware device capable of generating OTP. Some of these devices are PIN protected, offers an additional level of security. The user enters the OTP with other identity credentials (typically user name and password) and an authentication server validates the login request.

B. CAPTCHA: Completely Automated Public Turing test to tell Computers and Humans Apart, is a method recently adopted in some banking systems whose objective is to render automated attacks against authenticated sessions ineffective. This method requires the legitimate user to input information conveyed as scrambled images which are difficult for automated robots to process and recognize.

C. Hypertext transfer protocol secure -Hypertext Transfer Protocol Secure (HTTPS) is a communications protocol for secure communication over a computer network, with especially on the Internet. Technically, not a protocol in and of itself; rather it is the result of simply layering the Hypertext Transfer Protocol (HTTP) on top of the TLS protocol, thus adding the security capabilities of SSL/TLS to standard HTTP communications. The security of HTTP secure is therefore that of the underlying TLS, It uses long term public and secret keys to exchange a short term session key to encrypt the data flow between client and server. Important property in this context is perfect forward secrecy (PFS), so the short term session key cannot be derived from the long term asymmetric secret key; however, PFS is not widely adopted. To guarantee one is talking to the partner one wants to talk X.509 certificates are used.

D. Digital Certificates: Digital certificates are used to authenticate both the users and the banking system itself[3]. This kind of authentication depends on the existence of a Public Key Infrastructure (PKI) and a Certificate Authority (CA), which represents a trusted third-party who signs the certificates attesting their validity.

E. Attribute based anti-phishing techniques - Attribute-based anti-phishing strategy implements both reactive and

proactive anti-phishing defences. This technique has been implemented in Phish Bouncer tool.

The Image Attribution check does a comparison of images of visiting site and the sites already registered with phish bouncer. The HTML Crosslink check looks at responses from nonregistered sites and counts the number of links the page has to any of the registered sites A high number of cross-links is indicative of a phishing site. In false info feeder check ,false information is input and if that information is accepted by site then it is probable that link is phished one. The Certificate Suspicious check validates site certificates presented during SSL handshake and extends the typical usage by looking for Certification Authority (CA) consistency over time.URL suspicious check uses characteristics of the url to identify phishing sites.

Advantage: As attribute based anti-phishing considers a lot of checks so it is able to detect more phished sites than other approaches. It can detect known as well as unknown attacks.

Disadvantage: As multiple checks perform to authenticate site this could result in slow response time.

F. Genetic Algorithm Based Anti-Phishing Techniques - It is an approach of detection of phishing web pages using genetic algorithm. Genetic algorithms can be used to evolve simple rules for preventing phishing attacks. These rules are used to differentiate normal website from anomalous website. These anomalous websites refer to events with probability of phishing attacks. The rules stored in the rule base are usually in the following form [5]:

if { condition } then { act }

For example, a rule can be defined as:

If { The IP address of the URL in the received e-mail finds any match in the Ruleset }

Then

{ Phishing e-mail

}[5]

This rule can be explained as: if there exists an IP address of the URL in e-mail and it does not match the defined Rule Set for White List then the received mail is a phishing mail.

Advantage: It provides the feature of malicious status notification before the user reads the mail. It also provides malicious web link detection in addition of phishing detection.

Disadvantage: Single rule for phishing detection like in case of url is far from enough, so we need multiple rule set for only one type of url based phishing detection. Likewise for other parameter we need to write other rule this leads to more complex algorithm.

G. An Identity Based Anti-Phishing Techniques - This technique follows mutual authentication methodology where both user and online entity validates each other's identity during handshake. It is an anti-phishing technique that integrates partial credentials sharing and client filtering technique to prevent phishers from easily masquerading as legitimate online entities. As mutual authentication is followed, there would be no need for users to re-enter their credentials. Therefore passwords are never exchanged between users and online entities except during the initial account setup process.

H. Character Based Anti-Phishing Approach - Character based antiphishing technique uses characteristics of hyperlink in order to detect phishing links. Linkguard is a tool that implements this technique. After analysing many phishing websites, the hyperlinks can be classified into various categories as shown in fig 6. For detection of phishing sites LinkGuard, first extracts the DNS names from the actual and the visual links and then compares the actual and visual DNS names, if these names are not the same, then it is phishing of category 1 If dotted decimal IP address is directly used in actual DNS, it is then a possible phishing attack of category 2.

If the actual link or the visual link is encoded (categories 3 and 4), then first the link is decoded and

then analysed. When there is no destination information (DNS name or dotted IP address) in the visual link then the hyperlink is analysed. During analysis DNS name is searched in blacklist and white list .if it is present in whitelist then it is sure that the link is genuine and if link is present in blacklist then it is sure that link is phished one.

If the actual DNS is not contained in either whitelist or blacklist, Pattern Matching is done. During pattern matching first the sender email address is extracted and then it is searched in seed set where a list of address is maintained that are manually visited by the user. Similarity checks the maximum likelihood of actual DNS and the DNS names in seed-set. The similarity index between two strings are determined by calculating the minimal number of changes needed to transform a string to the other string.

I. **Content Based Anti-Phishing Approach** - GoldPhish[7] tool implements this technique and uses google as its search engine This mechanism gives higher rank to well-established web sites. It has been observed that phishing web pages are active only for short period of time and therefore will acquire low rank during internet search and this becomes basis for content based anti-phishing approach. The design approach can be broken down into three major steps. The first step is to capture an image of the current website in the user's web browser. The second step is to use optical character recognition techniques to convert the captured image into computer readable text. The third step is to input the converted text into a search engine to retrieve results and analyse the page rank.

IV. EFFECTIVE COUNTERMEASURES

After having a discussion over the different techniques that can be used for the task of session hijacking. One must be aware about that techniques too that works against it. Thus the countermeasures for it are:

- Use SSL to have secure communication channel.
- There must be logout function for session termination.
- Trust HTTPS connection for passing authentication cookies.
- Always pass encrypted data between user and webservers.
- Adopt a secure protocol.
- Regeneration of Session ID after log in.
- Reduce having remote access.
- Emphasis on Encryption.
- Reduce the life span of session or cookie.
- Always create session keys with lengthy strings or random numbers.
- Try preventing Eavesdropping.
- Expire the session as soon as user logs out.
- Do not access links received through mails.
- Use firewall and browser settings to restrict cookies.
- Make sure website which we are accessing is certified by certified authority.
- Clear history, offline contents and cookies from browser after every secret or sensitive transaction.

V. CONCLUSION

Nowadays, criminals are using phising as the highly profitable activity. Over the past several years, there has been an increase in the technology, diversity, and sophistication of these attacks in response to increased user awareness and countermeasures, in order to maintain profitability.Phishing differs from traditional scams primarily in the scale of the fraud that can be committed.The problem of Phishing does not have a single solution as of today. Phishing is not just a technical problem and Phishers would keep coming up with new ways of attacking the users. Online users should undertake periodic vulnerability analysis to identify and plug weaknesses that can lead to a successful Phishing attack.

From the above study it can conclude that generally the anti-phishing techniques focus on contents of

web age, URL and email. Character based anti-phishing approach may result in false positive but content based approach never results in false positive. Attribute based approach consider almost all major areas vulnerable to phishing so it can be best anti-phishing approach that can detect known as well as unknown phishing attack. Identity based anti-phishing approach may fails if phisher gets physical access to client's computer.

As a future work on phishing, it must be better trend to work on server side security. In the server side security policy , dual level of authentication can use for user by which only authentic user can get the access of his account, and to educate the user about this policy will results in avoiding user to give his sensitive information to phished web site.

In order to combat phishing, business and consumers need to adopt best practices and practice awareness, educate themselves about phishing and anti-phishing techniques, use current security protection and protocols, and report suspicious activities. The most effective solution to phishing is training users not to blindly follow links to web sites where they have to enter sensitive information such as passwords.

VI. REFERENCES

[1] IBM Internet Security Systems.

[2] [APJU] The Anti-Phishing Working Group, "Phishing Attacks Trend Report, June 2004".

[3] Microsoft Security Bulletin MS05-001, January 11, 2005.

[4] "Phishing attacks and countermeasures", Ramzan, Zulfikar (2010).

[5] AmmarAlmomani, B. B. Gupta, SamerAtawneh, Meulenberg, and EmanAlmomani "A Survey of Phishing Email Filtering Techniques" IEEE Communications Surveys & Tutorials, Vol. 15, No. 4, 2013.

[6] HALLER, N. A One-Time Password System (RFC 2289). Internet Engineering Task Force.[S.l.].1998.

[7] Neil Chou, Robert Ledesma, Yuka Teraguchi, Dan Boneh, and John Mitchell."Client-side defense against web-based identity theft".In 11th Annual Network and Distributed System Security Symposium (NDSS'04), San Diego, 2005.

[8] [APWG] The Anti-phishing Working Group, "Proposed Solutions to Address the Threat of E-mail Spoofing Scams," December 2003.

[9] Weiwei Zhuang1,2, Qingshan Jiang2,3*, Tengke Xiong2 "An Intelligent Anti-phishing Strategy Model for Phishing Website Detection" 1545-0678/12 © 2012 IEEE.

[10] Anti-Phishing Working Group, http://www.antiphishing.org.

[11] "HTML Code Injection and Cross-site scripting", Gunter Ollmann, 2001.

[12] "Web Based Session Management", Gunter Ollmann, 2002.

[13] "URL Encoded Attacks", Gunter Ollmann, 2002

[14] NISR The Phishing Guide Understanding & Preventing Phishing Attacks

[15] "Anti-Phishing: Best Practices for Institutions and Consumers", McAfee, March 2004.