

The Analysing and Techniques of the Anti-Phishing Technology

S.Suresh Babu

Lecturer, Dept of Computer Science, KBN College, Vijayawada, A.P, India

sunkarasuresh9@gmail.com

Abstract— Phishing is an attack that deals with social engineering methodology to illegally acquire and use someone else's data on behalf of legitimate website for own benefit (e.g. Steal of user's password and credit card details during online communication). It is affecting all the major sectors of industry day by day with a lot of misuse of user credentials. To protect users against phishing, various anti-phishing techniques have been proposed that follows different strategies like client side and server side protection. In this paper we have studied phishing in detail (including attack process and classification of phishing attack) and reviewed some of the existing anti-phishing techniques along with their advantages and disadvantages.

Keywords- Anti-phishing, Pharming, Phishing, Mutual authentication

I. INTRODUCTION

One of the primary goals of phishing is to illegally carry out fraudulent financial transactions on behalf of users using a forged email that contains a URL pointing to a fake web site masquerading as an online bank or a government entity. A phisher may lure a victim into giving his/her Social Security Number, full name, & address, which can then be used to apply for a credit card on the victim's behalf [1].

Attacker uses replica of original website as a bait that is send to the user. When user grabs the bait by filling and submitting his useful information attacker pulls the bait means saves the data for its own use illegally 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 emails to target users.

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

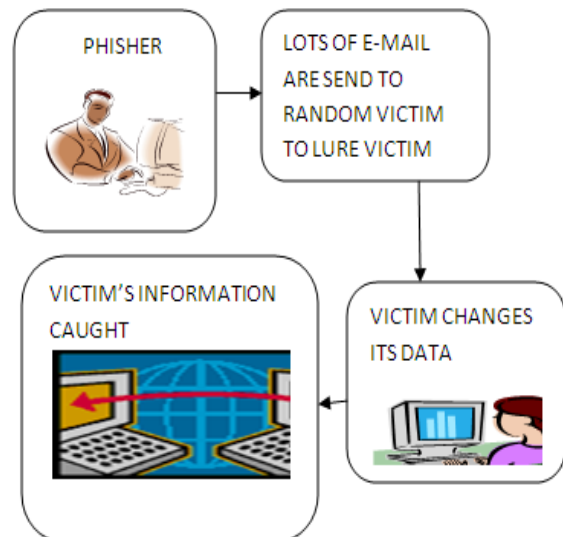


Fig 1: Process of phishing

According to a phishing activity trend report published by Ant phishing working group on 23 dec 2011, a lot of phishing attacks were done in first half of year 2011 as can be seen from fig 2. The number of unique phishing reports submitted to APWG in H1, 2011 reached a high of 26,402 in March, dropping to the half year low of 20,908 in April[2].



Fig 2: Phishing activity trend report

Financial Services continued to be the most targeted industry sector in the first half of 2011 as can be seen from figure 3.

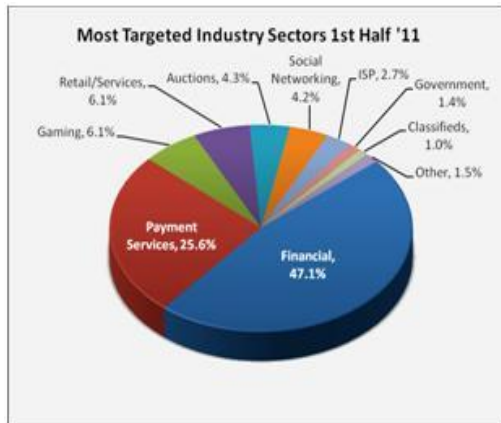


Fig 3: Industry sector area wise affect of Phishing

Retail sector appears to be third most vulnerable and classified as the least vulnerable to phishing attacks. So phishing attacks are emerging as one of the major area where immediate concern is needed as it is affecting all the major sectors of industry creating a lot of loss.

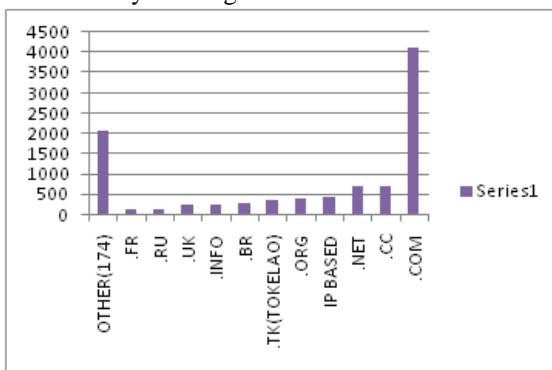


Fig 4: Domain name wise affect of Phishing

From the above figure it can be seen that .com sites are most vulnerable to phishing attacks. The figure also depicts that .net domain sites are also largely used by phisher for attack so it can be concluded that commercial site users becomes large victim of phishing attacks Further paper describes the literature survey and types of phishing and some of the anti-phishing techniques with their advantages and disadvantages

II. LITERATURE SURVEY

It has been done on implementing better and efficient tools on phishing detection and prevention, still it is very hard to completely eradicate the problem and to estimate no. of users that actually caught in bait of phishing logs that 311,449 people fall for phishing scams annually, costing around 350

million dollars . There are various techniques which defend against phishing. Some techniques give e-mail level protection and some provide security toolbars embedded with anti-phishing tools of indicators that identifies and distinguish legitimate sites from phishing sites. These indicators has been clustered into six criteria with their parameters of indication respectively as shown in table such as web link based identity, encryption and security based, source code and client side verification based, page layout and content based, web address bar based and social human factor based.

III. 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.

Deceptive Phishing- Messages about the need to verify account information, system failure requiring users to re- enter their information, fictitious account charges, undesirable account changes, new free services requiring quick action, and many other scams are broadcast to a wide group of recipients with the hope that the victim will respond by clicking a link to or signing onto a bogus site where their confidential information falls in this category.

Malware-Based Phishing- Refers to scams that involve running malicious software on users' PCs. Malware can be introduced as an email attachment, as a downloadable file from a web site, or by exploiting known security vulnerabilities.

Web Trojans- They pop-up invisibly when users are attempting to log in. They collect the user's credentials locally and transmit them to the phisher. **Hosts File Poisoning-** When a user types a URL to visit a website it must first be translated into an IP address before it is transmitted over the Internet. The majority of SMB(small and medium business organizations) users' PCs running a operating system look up these "host names" in their "hosts" file before undertaking a Domain Name System (DNS) lookup. By "poisoning" the hosts file.

System Reconfiguration Attacks- Modify settings on a user's PC for malicious purposes. For example: URLs in a favourites file might be modified to direct users to look alike websites. For example: a bank website URL may be changed from www.gmail.comto NS-Based Phishing "Pharming") With a pharming scheme, hackers tamper with a company's hosts files or(DNS)domain name system.

Content Injection Phishing-It describes the situation where hackers replace part of the content of a legitimate site with false content designed to mislead or misdirect the user into giving up their confidential information to the hacker. For example, phisher may insert malicious code to log user's credentials or an overlay which can secretly collect information and deliver it to the phisher.

Man-in-the-Middle Phishing- In these attacks phisher positions themselves between the user and the legitimate website or system. They record the information being entered but continue to pass it on so that users' transactions are not affected. Later they can sell or use the information or credentials collected when the user is not active on the system.

Search Engine Phishing- Occurs when phishers create websites with attractive (often too attractive) sounding offers and have 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. For example, scammers have set up false banking sites offering lower credit costs or better interest rates than other banks. Victims who use these sites to save or make more from interest charges are encouraged to transfer existing accounts and deceived into giving up their details.

IV. ANTI PHISHING

Anti-phishing refers to the method employed in order to detect and prevent phishing attacks. Antiphishing protects users from phishing. A lot of work has been done on anti-phishing devising various antiphishing techniques. Some techniques works on emails, some works on attributes of web sites and some on URL of the websites. Many of these techniques focus on enabling clients to recognize & filter various types of phishing attacks.

In general anti-phishing techniques can be classified into following four categories Content Filtering- In this methodology Content/email are filtered as it enters in the victim's mail box using machine learning methods

Such as Bayesian Additive Regression Trees (BART) or Support Vector Machines (SVM).

Black Listing- Blacklist is collection of known phishing Web sites/addresses published by trusted entities like google's and Microsoft's black list. It requires both a client & a server component. The client component is implemented as either an email or browser plug-in that interacts with a server component, which in this case is a public Web site that provides a list of known phishing sites [1].

Symptom-Based Prevention- Symptom-based prevention analyses the content of each Web page the user visits and generates phishing alerts according to the type and number of symptoms detected.

Domain Binding- It is an client's browser based techniques where sensitive information (eg. name, password) is bind to a particular domains. It warns the user when he visits a domain to which user credential is not bind.

V. ANTI-PHISHING TECHNIQUES

Attribute based anti-phishing techniques Attribute-based antiphishing strategy implements both reactive and proactive anti-phishing defenses. This technique has been implemented

in Phish Bouncer tool. The various checks that phish bouncer does has been shown in figure 5.

The Image Attribution check [4] does an 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[4]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.

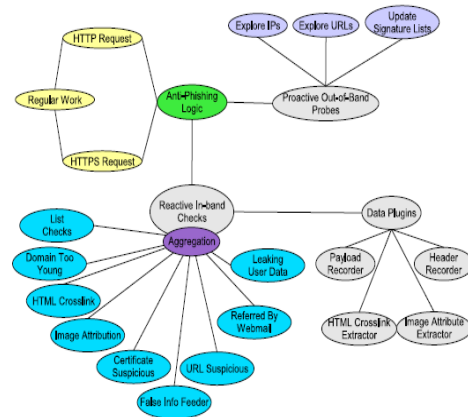


Fig 5: Use Case Diagram Showing Check, Probes and Data plugins

Advantage: As attribute based antiphishing 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. 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:

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
 }

This rule can be explained as:

if there exists an IP address of the URL in email 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.

Advantage: It provide mutual authentication for server as well as client side. Using this technique user does not to reveal his credential password in whole session except first time when the session is initialized.

Disadvantage: In identity based ant phishing if a hacker gain access to the client computer and disable the browser plug-in then method will be compromise against phishing detection Character Based Anti-Phishing Approach

Many time phishers tries to steal information of users by convincing them to click on the hyperlink that they embed into phishing email. A hyperlink has a structure as follows. Anchor text where 'URI' (universal resource identifiers) provides the actual link where the user will be directed and 'Anchor text' is the text that will be displayed in user's Web browser and represents the visual link.

Character based antiphishing technique uses characteristics of hyperlink in order to detect phishing links. Link guard [6] 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 Link Guard, 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

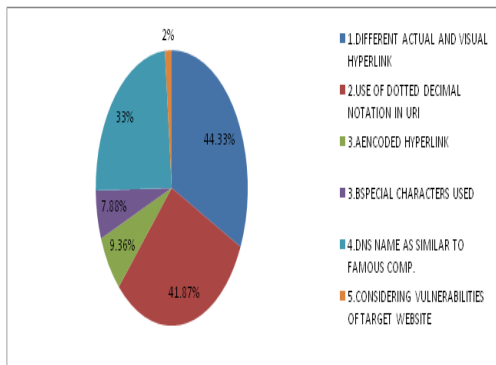


Fig 6: Linkgaurd Analysis In Various Classified Hyperlinks

If the actual dnsis 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 third step is to input the converted text into a search engine to retrieve results and analyse the page rank.

Advantages:

Generally Gold Phish does not result in false positive and provides zero day phishing.

Disadvantages:

Gold Phish delays the rendering of a webpage. It is also vulnerable to attacks on Google's PageRank algorithm and Google's search service.

VI. CONCLUSION AND FUTURE WORK

In the above study we can conclude that most of 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

VII. REFERENCES

- [1] Hicham Tout, William Hafner "Phishpin: An identity-based antiphishing approach" in proceedings of international conference on computational science and engineering, Vancouver, BC, pages 347-352, 2009 .
- [2] Phishing activity trend report 1st half /2011, <http://www.antiphishing.org>.
- [3] Michael Atighetchi, ParthaPal "Attribute-based prevention of phishing attacks" Eighth IEEE international symposium on network computing and application, 2009.
- [4] V.Shreeram, M.Suban, P.Shanthi, K.Manjula "Anti-phishing detection of phishing attacks using genetic algorithm" in proceedings of Communication control and computing technology (ICCCCT), IEEE international conference, Ramanathapuram , pages 447-450, 2010.