www.ijecs.in
International Journal Of Engineering And Computer Science ISSN: 2319-7242
Volume 3 Issue 5 may, 2014 Page No. 6255-6259

A survey on phishing detection and prevention technique

Archit Shukla [1], ***Lalit Gehlod*** [2]

[1] Institute of Engineering & Techology ,M.E. in CS,
DAVV,Indore, India
Shukla9190@gmail.com

[2] Institute of Engineering & Techology ,Dept. of CS,
DAVV,Indore, India
lalitgehlod@yahoo.co.in

Abstract : The first and greatest casualty of fraud is faith. According to [7] just over two-thirds (68%) of fraud victims say they are less willing to have faith on others after their fraud experience and 63% are less willing to make future investments. As far as, people with criminal intentions are concerned, identity theft is a conventional idea. A con man in police uniform, not cause many victims to become suspicious and they will comply with whatever they are told. Similarly, phishing is a form of online identity theft that aims to steal sensitive information from users such as online banking passwords and credit card information. This paper focuses on the phishing attacks which includes study of different contributions of recent research on phishing detection and prevention techniques. In addition of that based on the review, a new model for detection and prevention of phishing attacks is given in this paper.

Keywords: *phishing, information thief, machine learning , XSS, prevention model*

1. Introduction

Since Web-based fraud caused by unidentified theft, phishing attacks, malware, and other threats, costs organizations millions of dollars every year. Web fraud also negatively impacts users' perceptions of e businesses. In fact, 34% of victims reported avoiding certain merchants and 17% switched their primary bank as the result of a fraud event [8]. Reducing online fraud not only benefit consumers, but also helps businesses slash fraud recovery costs, avoid reputation damage, and prevents customer churns. Unfortunately, lots of Web fraud detection solutions require modifying the application. This additional development and testing can dramatically lengthen deployment processes. Threat Radar Fraud Prevention enables organizations to rapidly provision and manage fraud detection solutions without needing to update Web applications. With Threat Radar Fraud Prevention, the Secure Sphere Web Application Firewall (WAF) can transparently identify and stop fraudulent transaction. It also provides powerful monitoring and enforcement capabilities, allowing business to centrally manage WAF and fraud policies together.

Each month, more attacks are launched with the aim of making web users believe that they are communicating with a faithed entity for the purpose of stealing account information, logon credentials, and unique information in general. This attack method, commonly known as "phishing," is most commonly initiated by sending out emails with links to spoofed websites that steal information. Author present a method for detecting these attacks, which in its most general form is an application of machine learning on a feature set designed to highlight user-targeted deception in electronic communication. This method is applicable, with slight modifications, to detection of phishing websites, or the emails used to direct victims to these sites [10]. Author evaluate this method on a set of approximately 860 such phishing email, and 6950 non-phishing email, and correctly identify over 96% of the phishing emails will only mis-classifying on the order of 0.1% of the legitimate emails.

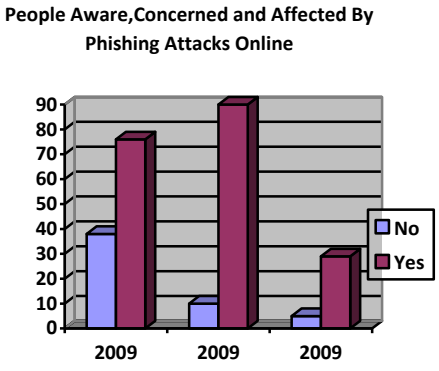

Fig 1.

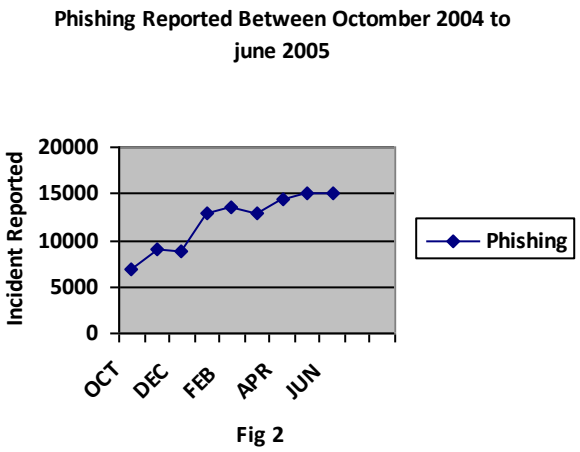

Fig 2

The name of the (electronic) street is Phishing; the process of tricking or socially engineering an organization's customer into imparting their confidential information for the wrong purpose. Riding on the back of mass-mailings such as Spams, or using bots to automatically target victim, any e business may find Phishers masquerading as them and targeting their customer information.While the security failures within SMTP are indeed a popular exploit vector for Phishers, there is an increasing array of communication channels available for malicious message delivery.As show in Fig1 that how many peoples are aware ,concerned and affected by phishing attacks.Similarly Fig 2 shows that phishing attacks are increasing .As with most criminal enterprises, if there is sufficient money to be made through phishing, the other message delivery avenue will be sought even if the holes in SMTP are eventually closed (although this is unlikely to happen within the next 3-5 years). With the high fear-factor associated with possible phishing scams, organizations that take a proactive stance in protecting their customers' personal information are likely to benefit from higher levels of faith and confidence in their services.

A Types of phishing attacks

• Deceptive Phishing: - The term "phishing" originally referred to accounting theft using instant messaging but the most common broadcast method today is a deceptive email message. The message 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 actions, and many other scams are broadcast to a wide group of recipients with the hope that the unwary will respond by clicking a link to or signing onto a bogus site where their confidential information can be collected.

• Malware-Based Phishing: - It 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 Screen logger are particular varieties of malware that track keyboard input web site, or by exploiting known security vulnerabilities.

• System Reconfiguration: - Attacks modify settings on a user's PC for malicious purposes. For example: URLs in a favorite file might be modified to direct users to look same website. For example: a bank website URL may be changed from "bankof123.com" to "bancof123.com".

• Hosts File Poisoning.: - When a user types a URL to visit a website it must first be translated into an IP address before it's transmitted over the Internet.

• Data Theft: - Unsecured PCs often contain subsets of sensitive information stored elsewhere on secured servers. Manly PCs are used to access such servers and can be more easily compromised.

• DNS-Based Phishing ("Pharming"): - Pharming is the term given to hosts file modification or Domain Name System (DNS) based phishing. With a pharming scheme, intruders tamper with a company's host file or domain name system so that requests for URLs or name service return a bogus address and subsequent communications are directed to a fake site.

• Content-Injection Phishing: -Describe 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.

• Phishing through Search Engines: - Some phishing scams involve search engines where the user is directed to products sites which may offer low cost products or services.

• Phone Phishing: -In the phone phishing, phisher makes phone calls to the user and asks the user to dial a number.

• Malware Phishing: -Phishing scams involving malware require it to be run on the user's computer. Models Provided for detection and prevention..

• 2.Literature Survey

A. ***Protecting user against phishing using Antiphishing: -*** This paper presents a novel browser extension, AntiPhish, that aim to prevent users against spoofed web site-based phishing attack. To this end, AntiPhish tracks the sensitive information of a user and generates warnings whenever the user attempts to give away this information to a web site that is considered unfaithed. 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. However, expecting that all users will understand the phishing threat and surf accordingly is unrealistic. There will always be user that is tricked into visiting a phishing web site. Therefore, it is important for researchers and industry to provide solutions for the phishing threat. Most proposed phishing solutions are based on the crawling of websites to identify "clones" and the maintenance of black lists of phishing websites. Such solutions, however, require the antiphishing organizations to be much faster than the attackers [6].

B. **Learning to Detect Phishing Emails: -**Author present a method for detecting these attacks, which in the most general form is an application of machine learning on a feature set designed to highlight user-targeted deception in electronic communication. This method is applicable, with slight modification, to detect phishing websites, or the emails used to direct victims to these sites. Author evaluate this method on a set of approximately 860 such phishing emails, and 6950 non-phishing emails, and correctly identify over 96% of the phishing emails will only mis-classifying on the order of 0.1% of the legitimate emails. Author conclude with thoughts on the future for such techniques to specifically identify deception, especially with respect to the evolutionary nature of the attacks and information available [3].

C. **Phishing detection system for e-banking using fuzzy data mining: -** Detecting and identifying any phishing websites in real-time, particularly for e-banking services, is really a complex and dynamic problem involving many factors and criteria. Because of the subjective consideration and the ambiguities involved in the detection, of fuzzy data mining techniques can be an effective tool in assessing and identifying phishing websites for e-banking since it offers a more natural way of dealing with quality factors rather than exact values. In this paper, Author present a novel approach to overcome the 'fuzziness' in the e-banking phishing website assessment and propose an intelligent resilient and effective model for detecting e-banking phishing websites. The proposed model is based on fuzzy logics combined with data mining algorithms to characterize the e-banking phishing website factors and to investigate its techniques by classifying the phishing types and defining six e-banking phishing website attack criteria's with a layered structure [9].

D. ***Collaborative Detection of Fast Flux Phishing Domains:*** - Author propose two approaches to correlate evidence from multiple DNS servers and multiple suspect FF domain. Real-world experiment show that our correlation approaches speed-up FF domain detection, based on an analytical model that present to quantify the number of DNS queries needed to confirm a FF domain. Author also show how our correlation schemes can be implemented on a large scale by using a decentralized publish-subscribe correlation model called LARSID, that is more scalable than a fully centralized architecture. In conclusion, FF domains are extremely difficult to detect in a timely and accurate manner, due to the use of a screen of proxies to shield the FF Mothership. Author present a theoretical model to analyze the FF detection problem by quantifying the number of DNS queries needed to retrieve a certain number of unique IP addresses [11].

E. **A Prior-based Transfer Learning Method for the Phishing Detection: -** In this paper, Author present a priority-based transfers learning method for our statistical machine learning classifier which based on the logistic regression to detect the phishing sites that relies on our selected features of the URLs. Because of the mismatched distribution of the features in different phishing domain, Author employ multiple models for different regions. But it is impossible for us to collect enough data from a new region to rebuild the detection model and adjust the existing model by the transfer learning algorithm to solve these problems. Our URL-based method in phishing detection is an appropriate solution. To fulfill the detection requirements of features' mismatched, so proposed a transfer learning method to generate an adaptive model for the new detection scenario since it is impossible to train a new model with a limited training sample [12].

3.Methodology

Cross-Site Scripting (also known as XSS) Is one of the most common application-layer web attacks. XSS vulnerabilities target scripts embedded in a page which are executed on the client-side (in the user's web browser) rather than on the server-side. XSS in itself is a threat which is brought about by the internet security weaknesses of client-side scripting languages such as HTML and JavaScript. The concept of XSS is to manipulate client-side scripts of a web application to execute in the manner desired by the malicious user. Such a manipulation can embed a script in a page which can be executed every time the page is loaded, or whenever an associated event is performed.

XSS is the most common security vulnerability in software today. This should not be the case as XSS is easy to find and easy to fix. XSS vulnerabilities can have consequences such as tampering and sensitive data theft.

Key Concepts of XSS

XSS is a Web-based attack performed on vulnerable Web applications, In XSS attacks, the victim is the user and not the application .In XSS attacks, malicious content is delivered to users using JavaScript.

Explaining Cross-Site Scripting

An XSS vulnerability arises when Web applications take data from users and dynamically include it in Web pages without first properly validating the data. XSS vulnerabilities allow an attacker to execute arbitrary commands and display arbitrary content in a victim user's browser. A successful XSS attack leads to an attacker controlling the victim's browser or account on the vulnerable Web application. Although XSS is enabled by vulnerable pages in a Web application, the victims of an XSS attack are the application's users, not the application itself. The potency of an XSS vulnerability lies in the fact that the

malicious code executes in the context of the victim's session, allowing the attacker to bypass normal security restrictions.

Impact of Cross-Site Scripting

When attackers succeed in exploiting XSS vulnerabilities, they can gain access to account credentials. They can also spread Web worms or access the user's computer and view the user's browser history or control the browser remotely. After gaining control of the victim's system, attackers can also analyze and use other intranet applications.

By exploiting XSS vulnerabilities, an attacker can perform malicious actions, such as:

Hijack an account, Spread Web worms, Access browser history and clipboard contents, Control the browser remotely, Scan and exploit intranet appliances and applications, Identifying Cross-Site Scripting Vulnerabilities, XSS vulnerabilities may occur if: Input coming into Web applications is not validated. Output to the browser is not HTML encoded.

For example, the HTML snippet:

<title>Example document: %(title)</title>

is intended to illustrate a template snippet that, if the variable title has value Cross-Site Scripting, results in the following HTML to be emitted to the browser:

<title>Example document: XSS Doc</title>

Table (1) XSS Examples

4.Limitation

In recent year's uses of internet are rapidly increases, the number of internet users are also increasing in the same manner. On the other hand internet user now becomes more aware about the internet based frauds and scams. But the numbers of phishing attacks are increases as the internet users are increases and awareness about phishing is increases. To detect and prevent the phishing attacks, the browsers currently usage SSL/TLS, but these techniques is not much effective and still allows web spoofing, i.e. misleading users by impersonation or misrepresentation of identity or of credentials. Indeed, there is an alarming increase in the amount of real-life web-spoofing attacks, usually using simple techniques. Often, the swindlers lure the user to the spoofed web site, e.g. impersonating as financial institution, by sending her spoofed E-mail messages that link into the spoofed web-sites; this is often called a phishing attack. The goal of the attackers is often to obtain user-ID's (Identity), passwords/PIN (Personal Identification Number) and other personal and financial information, and abuse it e.g. for identity theft. Thus, the significant improvement in detection of spoofed sites is required.

5.Proposed Model

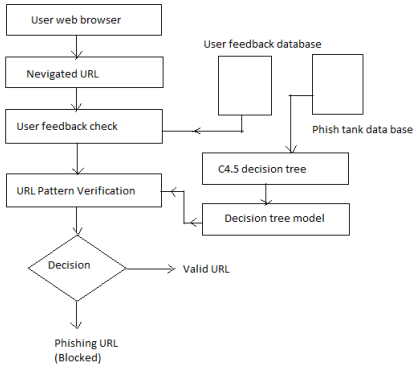

Fig 3

The description of the above Fig 3 is given below.

Phish tank Database: That is an updated database where the entire phish reported web URLs is stored, proposed system contains a relational data table which store these web URL patterns and used to build a data model form algorithm selected.

Universal Database: This database in common for all guests who use proposed tool, this database contains user feedback about URLs.

Navigated URL: That is, a user interface where a user navigated URL information is stored and provides the various user experiences about the navigated page.

USER Feedback: A user interface provided in the proposed model to submit feedback for a URL if required to report and this is taken in databases.

Algorithm selection: This module contains algorithms and user select an algorithm for consuming phish tank database and develop a data model for navigation.

Data model: The developed data model is a decision tree which is grown using a phish tank database and used to analysis the URL pattern which is found in the database. After analysis of web URL decision is reached.

6.Conclusion & Future Work

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. To guard against these threats, user need to be educated on the dangers of advanced malware and the forms it can take today. In addition, security teams need advanced technologies that can detect and stop the advanced threats that are currently bypassing their conventional defenses. This paper further endorses the recommendations in support of the fight against phishing and identity theft as a whole. It did not discuss the growing trend towards outsourced email. Communication and log analysis across organizational boundaries can be challenging. In the longer terms we expect that other digital payment activity will also be the victim of attacks. We advise internet banking services to seriously research these problems before attacks are carried out in the wild. A control that protects all crucial internet banking activity and the information involved in this activity is required.

7.References

1. "Phishing: Challenges and Issues in Malaysia" Madihah Mohd Saudi, Islamic Science University of Malaysia (USIM), Negeri Sembilan, Malaysia Shaharudin Ismail, Islamic Science University of Malaysia (USIM), Negeri Sembilan, Malaysia Emran Mohd Tamil, University Malaya, Malaysia Mohd Yamani Idna Idris, University Malaya, Malaysia.
2. "Online Detection and Prevention of Phishing Attacks (Invited Paper)" Juan Chen Institute of Communications Engineering Nanjing 210007, P.R. China icechj@msn.com Chuanxiong Guo Institute of Communications Engineering Nanjing 210007, P.R. China xguo@ieee.org
3. "Learning to Detect Phishing Emails" Ian Fette School of Computer Science Carnegie Mellon University Pittsburgh, PA, 15213, USA icf@cs.cmu.edu Norman Sadeh School of Computer Science Carnegie Mellon University Pittsburgh, PA, 15213, USA Anthony Tomasic School of Computer Science Carnegie Mellon University Pittsburgh, PA, 15213, USA
4. "A Comparison of Machine Learning Techniques for Phishing Detection" Saeed Abu-Nimeh1, Dario Nappa2, Xinlei Wang2, and Suku Nair1 SMU HACNet Lab Southern Methodist University Dallas, TX 75275
5. NISR The Phishing Guide Understanding & Preventing Phishing Attacks
6. "Protecting Users Against Phishing Attacks with AntiPhish" Engin Kirda and Christopher Kruegel Technical University of Vienna
7. Canadian Securities Administrators | www.csa-acvm.ca ,Innovative Research Group, Inc. | www.innovativeresearch.ca
8. Proactively Stop Web Based Fraud and Fraudlent Transaction With ThreatRadar Fraud Prevention, www.imperva.com
9. Modeling and Preventing Phishing Attacks by Markus Jakobsson, Phishing detection system for e-banking using fuzzy data mining by Aburrous, M. ; Dept. of Comput., Univ. of Bradford, Bradford, UK ; Hossain, M.A. ; Dahal, K. ; Thabatah, F.
10. Learning to Detect Phishing Emails : Authors AnIan Fette (Carnegie Mellon University),Norman Sadeh (Carnegie MellonUniversity),Thony Tomasic (Carnegie Mellon University
11. Collaborative Detection of Fast Flux Phishing Domains Chenfeng Vincent Zhou, Christopher Leckie and Shanika Karunasekera Department of Computer Science and Software Engineering, The University of Melbourne, Australia.
12. A Prior-based Transfer Learning Method for the Phishing Detection Jianyi Zhang1,2,3, Yangxi Ou2,3, Dan Li2,3, Yang Xin2,3 1Beijing Electronic Science and Technology Institute, Beijing, China 2Information Security Center, Beijing University of Posts and Telecommunications, Beijing, China 3Beijing Safe-Code Technology