IJCSC

ISSN : 0973-7391

Vol. 3, No. 1, January-June 2012, pp. 215-219

BROWSER PREVENTION AGAINST PHISHING WEBSITE SECURITY RISK

Aanchal Malhotra[1], and Navdeep Kaur[2]

[1]Department of Computer Science Engineering (Information Security) Punjab Engineering College, university of technology Chandigarh *E-mail: aanchalmalhotra586@gmail.com*
[2]Department of Information Technology Punjab Engineering College, university of technology Chandigarh *E-mail: aulakh83@gmail.com*

ABSTRACT

Phishing is an online identity theft and is a malicious form of internet fraud with the aims to steal sensitive information such as credits, social security number, account information it is mainly by crafting a faux online presence to masquerade, This paper proposes a phishing detection approach, we use URLs and contents of a websites to identify through special symbol in using their Domain name. Anti Phish that aims to protect against spoofed web site based phishing attacks, phishing identification and explain various method to detect them.

Keywords: Types of phishing; anti phishing browser;

1. INTRODUCTION

Phishing attacks identify theft through phishing scams has become growing, it is an online identity theft that aims to steal sensitive information such as user names, credit card number and passwords. Today, online transaction worth billions of dollars are initiated every day. This is mainly by crafting a faux online presence to masquerade, hence the number of phishing attacks happens[1].

It has been seen often that till 2008 web browsers were foundto be implementing just the basic browsing. There were no such high security factors involved to save the browser from unsafe browsing[2]. Later on when companies like google, Mozilla focused their business on security risks, they found their browser unsafe enough to get hacked. Later on they started implementing the features to save the browsing from the phishy sites. They implemented a basic algorithm which matches the url from the database of their own server and they process the url if the url is safe. Still the basic implementation of the url hacking was missing from the browsers. As we all know that the hackers are very much active now a days and they are capable enough to distract the path of the browsing. The basic aim of this work is to save the browser from unsecured browsing if the browser gets corrupted. The hackers implemented a technique toz introduce special symbols in the url to hit the website they created. This happens in the sense that, suppose we are trying to open up a website say *http://www.helloworld.com.* Suppose the browser got hacked through a virus which can impart special symbols into the site So this site will take us to say. *http://www.hello$world.com* which for the time being is fishy. Although if we are browsing it through Google, we may run it safe because they update their database for the phishy websites every single hour. Still browsers like safari does not support this kind of facility[3][4]. Even though with Google also, sometimes we will find that the browser is taking you to uncertain website. Suppose we want to open up facebook.com, and the browser is infected by some virus then the url will be like face#book.com and the browser will take you to book.com not facebook.com.

2. TYPES OF PHISHING ATTACKS

A. Deceptive Phishing

Phishing referred to report robbery using messaging about the requirement to conform the account information unwanted account changes etc, recipients with the hope that un suspecting will react and click a link to or signing onto a bogus site where their secret information can be collected

B. Malware- Based Phishing

Refers to scams that involve running malicious software is usually attached to the email sent to the user by the phishers. Once you click on the link the malware will start working sometime it is attached to download able files.

C. Key Loggers and Screen Loggers

Key loggers refer to the malware used to identify input from the keyboard and throw applicable information to the hacker via the internet. To prevent key loggers from accessing personal information ,secure websites provide option to use mouse click to make entries through the virtual keyboard.

D. Session hacking

In session hacking, the phisher exploits the web session control mechanism to steal the information of the users .at that point the malicious software takes over and can undertake unauthorized actions such as transferring funds, without the user's knowledge[5].

E. Web trojan

When pop up invisibly when user are sttempting to login, the hacker collect the user credentials locally and transmit than to the phisher.

F. Host 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. The bulk of SMB users' PCs running a Microsoft Windows operating system look up the "host names" in their "hosts" file before responsibility a Domain Name System (DNS) lookup. "poisoning" the hosts file, hackers have a bogus address transmitted, taking the client unwittingly to a fake [6] "look alike" website where their information can be stolen.

G. System Reconfiguration Attacks

Alter settings on a user's PC for hateful purpose. For example: URLs in a favorites file might be customized to through client to look similar websites. For example: a bank website URL may be changed from "bankofabc.com" to "banlofabc.com".

H. Data Theft

Unsecured PCs usually hold subsets of responsive information stored elsewhere on protected servers. PCs are used to access such servers and can be more easily compromised. By theft confidential communications, legal opinions, design documents, employee related records, etc., thieves profit from selling to those who may want to embarrass or cause economic damage or to competitors.

I. DNS-Based Phishing ("Pharming")

Pharming is the term given to hosts file alteration or Domain Name System (DNS)-based phishing. With a pharming scheme, hackers interfere with a company's host's files or domain name system so that requirements for URLs or name service return a bogus address and succeeding communications are directed to a fake site. The result: users are ignore that the website where they are incoming secret information is controlled by hackers[7] and is not even in the same country as the legitimate website.

J. Content-Injection Phishing

Phishing describes the situation where hackers replace part of the content of a real site with false material intended to deceive or misdirect the user into giving up their confidential information to the hacker.

K. Man-in-the-Middle Phishing

Phishing is difficult to distinguish between many other forms of phishing. In these attack hacker's location themselves between the user and the legitimate website[8]. They proof the information being entered but carry on passing it on so that users' transactions are not important. Later they can sell or use the information or credentials together when the user is not vigorous on the system.

L. Search Engine Phishing

Phishing occur when phishers make websites with good-looking, Users find the sites in the usual course of pointed for products and services are fool into charitable up their information. For example, scammers have set up false banking sites offering lower credit costs or improved attention rates than other banks.

3. PHISHING TECHNIQUES

In a phishing attack, the phisher sends a large number of fake e-mails[10] to random Internet users that seem to be coming from well-known organization (e.g. financial institutions, credit card companies, etc). A. Basic URL Obfuscation Ref [11], URL obfuscation misguide the users into thinking that a link and/or web site displayed in their web browser or HTML-capable email client is that of a trusted site. These methods tend to be technically simple highly effective, and are still used to some extent in phishing emails today.

A. Simple HTML Redirection

Simple techniques for obscuring the actual destination of a hyperlink is to use a legitimate URL within an anchor element but have its href attribute point to a malicious site. Thus clicking on a legitimate-looking URL then sends the user to a phishing site.

B. Use of Alternate Encoding Schemes

Hostnames and IP addresses can be represented in alternate formats that are less likely to be recognizable to most people. Alphanumeric characters can be changed to their hexadecimal representations.

C. Use of JPEG Images

Electronic mail rendered in HTML format is becoming more prevalent. Phishes are taking advantage of this by constructing phishing emails that contain a single image in JPEG format. When displayed, this image appears to

be real email from merchant site or an online bank. The image often contain official logos and text to add to the deception[12]. However, when users click on this image, they are directed to a phishing site.

D. Registration of Similar Domain Names

At initial glance, users may attempt to verify that the address displayed in the address or status bar of their web browser is the one for a real site. Phishes often register domain names that contain the name of their target institution to trick customers who are satisfied by just watching a legitimate name appear in a URL.A widely implemented version of this attack uses parts of a legitimate URL to form a new domain name as demonstrated below:

Legitimate URL *http://login.example.com*

Malicious URL *http://login-example.com*

E. Web Browser Spoofing Vulnerabilities

Over the past two years, several vulnerabilities in web browsers have provided phishers with the ability to obfuscate URLs and/or install malware on victim machines[13][14].

4. METHODS

The basic problem with the browser is that they do not avoid browsing to the phishy websites , or they do not block those sites which we don't want to open up while browsing[15]. The basic idea is to implement a SVM algorithm along with LRU algorithm to protect my browser from browsing to the fishy websites if the browser is affected through some virus attack or due to some undefined piece of code block which affects the browser while browsing .

Use of data mining for detecting phishing websites

The proposed technique contains two databases

- Main database
- Temporary database

1. ***Main database:*** contains all the possible criteria which help to detect phishing web sites. The features are given as follows
   - The main database contains all the phishing criteria. Phishing site is detected followed by which criteria that is stored in the temp database.
   - Counter value stores the iteration of criteria. How many times the phishing site is detected through this particular criteria like

Structure of the main database is

- It Contains index number
- Name of the criteria
- Counter value

2. ***Temporary database:*** contains the criteria index and counter value
   - Temporary database stores maximum 5 feature
   - Each feature has counter value
   - If temporary database has space say 4 features are already stored then next feature is added otherwise it will replace with the feature which has least counter value.

The stores only those criteria index number through which phishing sites are detected

A. Work Flow

User enters the url in browser's address bar and press enter. Code starts scan whether it is legitimate site or phishing site. First time temporary database is empty and Main database is used for detecting phishing websites. The features through which phishing site is

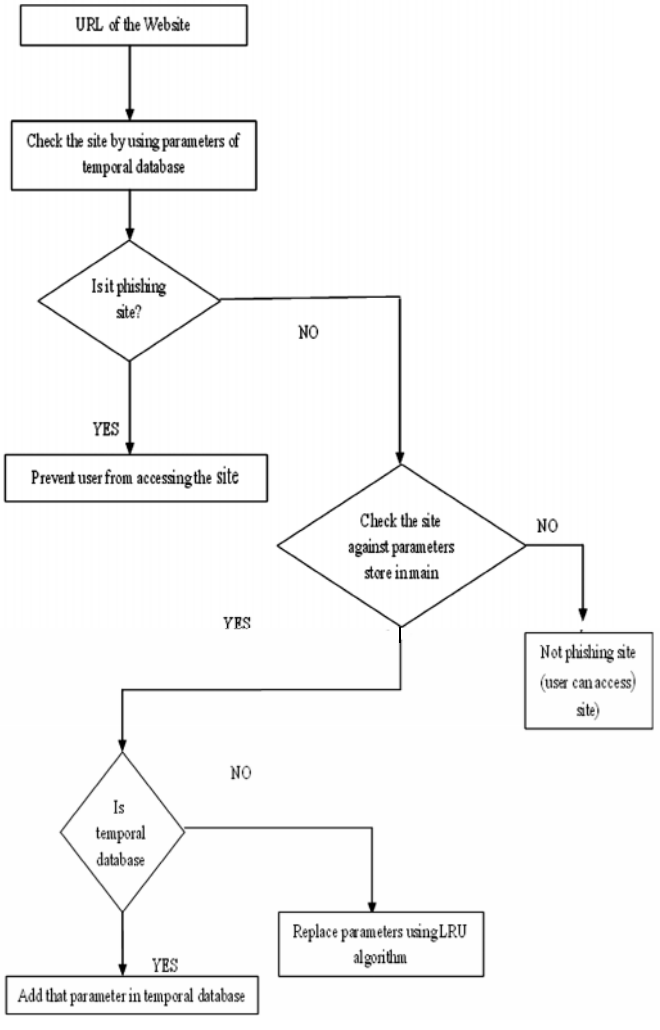

Figure 1: Flowchart of Proposed Technique

detected its index number is stored in temporary database and also its counter value increases Say the sites

is detected phishing through criteria using IP address now the index number of IP address in main database is 1. The index number 1 store in temporary database and counter value of IP address is increases in both databases (main and temporary databases). When anyone wants to access any website then again it detects whether the site is legitimate site or phishing site this time code uses temporary database criteria to detect the phishing website. If the features fulfilled then the site is phishing website otherwise it scans the main database

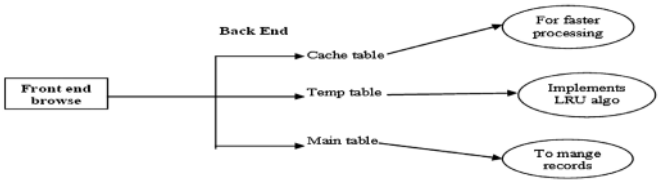

Figure 2 : Back End Flow Digram

B. Result

We have proposed our own browser named "Browseee" to block the phishy sites. It works like as other browser having secured and normal features Find out google search, yahoo search and MSNsearch in a same browser. On selecting smart security, the feature of detecting and blocking phishy web sites is enabled

Table 1
Comparison with Other Browsers

Browse Method	*Google chrome*	*Mozilla*	*Safari*	*Operation*	*Browsee My Browsee*
LRU	√	X	X	X	√
Phishing website detection & stopping	Δ	Δ	Δ	Δ	√
Manual testing	X	X	X	X	√
tabs working	√	√	√	√	X

On entering the phishy site in URL it scan and detect it as phishy website and block it and save the reason for future use.

After blocking the website it displays the message "Fishy website from cache".

5. CONCLUSION

Phishing differs from traditional scams primarily in the scale of the fraud that can be committed. Con artists have been around for centuries, but E-mail and the World Wide Web provide them with the tools to reach thousands or millions of potential victims in minutes at almost no expense. Phishing has become a major threat to information security and personal privacy. we represented new Web Browser based on URL domain identity. It first identifies the related authorized URL. We used approximate LRU algorithm for check the counter value of symbols. The proposed approach is inspired by the AntiPhish browser plug in and thesolution addresses the shortcomings of these approaches and aims to make these systems more effective. When checking for domain names, we consider features that are visually perceived by users because, as reported in literature, victims are typically convinced that they are visiting a legitimate page by judging the look-and-feel of a web site.

REFERENCES

[1] Vieira, M., Antunes, N., Madeira, H., "Using Web Security Scanners to Detect Vulnerabilities in Web Services", *Intl. Conf. on Dependable Systems and Networks*, Lisbon, 2009.

[2] Simson Garfinkel with Gene Spafford, "Web Security & Commerce" 1-56592-269-7, Order Number: 2697, 1st Edition June 1997.

[3] Deng Liwu, Xu Ruzhi, Jiang Lizheng and Lv Guangjuan., "A Database Protection System Aiming at SQL Attack" *School of Computer Science & Technology*, North China Electric Power University, Beijing, China IEEE 2009.

[4] Sadia Afroz and Rachel Greenstadt. "PhishZoo: Detecting Phishing Websites By Looking at Them". *Department of Computer Science*, Drexel University, Philadelphia, PA 19104 IEEE 2011.

[5] Brad Wardman , Gaurang Shukla, and Gary Warner. "Identifying Vulnerable Websites by Analysis of Common Strings in Phishing URLs". Computer Forensics Lab, University of Alabama at Birmingham. *IEEE* 2009.

[6] Insoon Jo, Eunjin (EJ) Jung and Heon Y.Yeom., "Yor're not who you Claim to be : Website Identity Check for Phishing Detection", *School of Computer Science and Engineering*, University of San Francisco, Seoul, Korea IEEE 2010.

[7] JungMin Kang and DoHoon Lee . "Advanced White List Approach for Preventing Access to Phishing Sites". *Electronics and Telecommunications Research* Institute (ETRI), Yuseong , Daejeon, South Korea. IEEE 2007.

[8] Engin Kirda and Christopher Kruegel. "Protecting Users Against Phishing Attacks with AntiPhish". *Technical University of Vienna, IEEE* 2005.

[9] T. Moore and R. Clayton, "Examining the Impact of Website Take-Down on Phishing", in Proceedings of the Anti-Phishing Working Groups 2nd Annual eCrime Researchers Summit. ACM, 2007, pp. 1-13.

[10] A. Herzberg and A. Gbara, "Security and Identification Indicators for Browsers Against Spoofing and Phishing Attacks", Cryptology ePrint Archive, Report 2004/155, 2004, http://eprint.iacr.org/

[11] "Phishing Attack Trends Report, June 2004", Anti-Phishing Working Group, http://www.antiphishing.org APWG_Phishing_Attack_Report-Jun2004.pdf.

[12] Y. Zhang, J. Hong, and L. Cranor, "Cantina: A Content Based Approach to Detecting Phishing web Sites", *in Proceedings of the 16th International conference on World Wide Web*, 2007.

[13] C. Whittaker, B. Ryner, and M. Nazif, "Large-Scale Automatic Classification of Phishing Pages", in NDSS'10, 2010.

[14] K. T. Chen, J. Y. Chen, C. R. Huang, and C. S. Chen, "Fighting Phishing with Discriminative Keypoint Features", *IEEE Internet Computing*, **13,** No. 3, pp. 56-63, 2009.

[15] Thomas Raffetseder, Engin Kirda, and Christopher Kruegel. "Building Anti-Phishing Browser Plug-Ins: An Experience Report" Secure Systems Lab, Technical University of Vienna, *IEEE* 2007.