Volume 6, Issue 8, August 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Email Security System for Phising Attack Using End-Host Based Algo

Akash Dutta*
Department of Computer Science,
Noida International University,
Greater Noida, U.P, India

Anuranjan Mishra
Head of Department of Computer Science,
Noida International University,
Greater Noida, U.P, India

Abstract— Email Phishing is an attack by attacker which creates a duplicate of an existing web page to make fool users in to submitting personal, financial, or password details data to what they think is their service provider's website. The concept based on end host anti-phishing algorithm, called the Link Guard, by usages of generic characteristics of the links in phishing attacks. The link Guard algorithm is the concept for finding the phishing emails sent by the phisher to grasp the information of the end user. The Link Guard is total based on the careful analysis of the characteristics of phishing hyperlinks. Each and every end user is implemented with Link Guard algorithm. After doing so that the end user recognizes the phishing emails and can avoid responding to such mails. Since Link Guard is characteristics based it can detect and prevent not only known phishing attacks but also unknown ones. The project uses the Java technologies and MYSQL.

Keywords— Email Attack, Email Phishing, phishing Algo, Email security, Internet security.

I. INTRODUCTION

Phishing is a new word produced from 'fishing', it refers to the act that the attacker allure users to visit a faked Web site by sending them faked e-mails and stealthily get victim's personal information such as user name, password, and national security ID, etc.

This information then can be used for future target advertisements or even identity theft attacks (e.g., transfer money from victims' bank account). The frequently used attack method is to send e-mails to potential victims, which seemed to be sent by banks, online organizations, or ISPs.

In these e-mails, they will make up some causes, e.g. the password of your credit card had been mis-entered for many times, or they are providing upgrading services, to allure you visit their Web site to confirm or modify your account number and password through the hyperlink provided in the e-mail.

If you input the account number and password, the attackers then successfully collect the information at the server side, and is able to perform their next step actions with that information (e.g., withdraw money out from your account).Phishing itself is not a new concept, but it's increasingly used by phishers to steal user information and perform business crime in recent years. Characteristics of phishing hyperlinks

1) The visual link and the actual link are not the same;
2) The attackers often use dotted decimal IP address instead of DNS name;
3) Special tricks are used to encode the hyperlinks maliciously;
4) The attackers often use fake DNS names that are similar (but not identical) with the Target Web site.

We then propose an end-host based anti-phishing algorithm which we call Link Guard, based on the characteristics of the phishing hyperlink.

Since Link Guard is ch0aracter-based, it can detect and prevent not only known phishing attacks but also unknown ones. It is very effective in detecting phishing attacks with minimal false negatives. Link Guard detects 195 attacks out of the 203 phishing archives provided by APWG without knowing any signatures of the attacks.

II. PROPOSED SYSTEM

I) Classification of the hyperlinks in the phishing e-mails:

In order to (illegally) collect useful information from potential victims, phishers generally tries to convince the users to click the hyperlink embedded in the phishing e-mail. A hyperlink has a structure as follows.

** Anchor text <\a>**

where 'URI' (universal resource identifiers) provides the necessary information needed for the user to access the networked resource and 'Anchor text' is the text that will be displayed in user's Web browser.

Examples of URIs are:

- http://www.google.com
- https://www.icbc.com.cn/login.html

- ftp://61.112.1.90:2345.

'Anchor text' in general is used to display information related to the URI to help the user to better understand the resources provided by the hyperlink. In the following hyperlink, the URI links to the phishing archives provided by the APWG group, and its anchor text "Phishing Archive" informs the user what's the hyperlink is about. <a href http://www.antiphishing.org/phishing-archive.html">Phishing Archive

Note that the content of the URI will not be displayed in user's Web browser. Phishers therefore can utilize this fact to play trick in their 'bait' e-mails. In the rest of the paper, we call the URI in the hyperlink the actual link and the anchor text the visual link. After analyzing the 203 (there are altogether 210 phishing e-mails, with 7 of them with incomplete information or with malware attachment and do not have hyperlinks) phishing email archives from Sep. 21st 33332003 to July 4th 2005 provided by APWG.

We classified the hyperlinks used in the phishing e-mail into the following categories:

1) The hyperlink provides DNS domain names in the anchor text, but the destination DNS name in the visible link doesn't match that in the actual link. For instance, the following hyperlink:

****
https://secure.regionset.com/EBanking/logon/

Appears to be linked to secure.regionset.com, which is the portal of a bank, but it actually is linked to a phishing site www.profusenet.net.

2) Dotted decimal IP address is used directly in the URI or the anchor text instead of DNS name. See below for an example.

<a href= "http://61.129.33.105/secured-site/www.skyfi.
Com/ index.html?MfclSAPICommand=SignInFPP&
UsingSSL= 1"> SIGN IN

3) The hyperlink is counterfeited maliciously by using certain encoding schemes.

There are two cases:

a) The link is formed by encoding alphabets into their corresponding ASCII codes. See below for such a hyperlink.

<a href="http://034%02E%0333%34%2E%311%39%355%2E
%o340o31:%34%39%30%33/%6C/%69%6E%64%65%78
**%2E%68%74%6D"> www.citibank.com **

While this link is seemed pointed www.citibank.com, it actually points to http://4.34.195.41:34/l/index.htm.

b) Special characters (e.g. (in the visible link) are used to fool the user to believe that the e-mail is from a trusted sender. For instance, the following link seems is linked to amazons, but it actually is linked to IP address 69.10.142.34.

http://www.amazon.com:fvthsgbljhfcs83infoupdate@69.10.142.34.

4) The hyperlink does not provide destination information in its anchor text and uses DNS names in its URI. The DNS name in the URI usually is similar with a famous company or organization. For instance, the following link seems to be sent from PayPal, but it actually is not.

Since PayPal -cgi is actually registered by the phisher to let the users believe that it has something to do with PayPal.

****
Click here to confirm your account

5) The attackers utilize the vulnerabilities of the target Web site to redirect users to their phishing sites or to launch CSS (cross site scripting) attacks. For example, the following link

<a href="http://usa.visa.com/track/dyredirjsp?rDirl=
http://200.251.251.10/.verified/"> Click here <a>

Once clicked, will redirect the user to the phishing site 200.251.251.10 due to a vulnerability of usa.visa.com.

III. SOFTWARE & HARDWARE REQUIREMENTS

3.1 Hardware Requirements

- Hard disk : 500 GB and above
- RAM : 2Ghz MB and above
- Processor speed : 2.6 GHz and above

3.2 Software Requirements

- Operating System : Windows 2000/XP
- Documentation Tool : Ms word 2000

3.3 Technologies Used

- JSP
- Servlets
- Apache Tomcat 5.5

3.4 Database

- Mysq l

Table **1**: The Categories of Hyperlinks In Phishing E-Mails

Category	Number of links	Percentage
1	90	44.33%
2	85	41.87%
3.a	19	9.36%
3.b	16	7.88%
4	67	33%
5	4	2%

Table 1 summarizes the number of hyperlinks and their percentages for all the categories. It can be observed that most of the phishing e-mails use faked DNS names (category 1,44.33°O) or dotted decimal IP addresses (category 2, 41.87%).Encoding tricks are also frequently used (category 3a and 3b, 17.24%). And phishing attackers often try to fool users by setting up DNS names that are very

Similar with the real e-commerce sites or by not providing destination information in the anchor text (category 4) . Phishing attacks that utilize the vulnerability of Web sites (category 5) are of small number (2%) and we leave this type of attacks for future study. Note that a phishing hyperlink can belong to several categories at the same time. For instance, an attacker may use tricks from both categories 1 and 3 at the same time to increase his success chance. Hence the sum of percentages is larger than 1.

Once the characteristics of the phishing hyperlinks are understood, we are able to design anti-phishing algorithms that can detect known or unknown phishing attacks in real-time.

IV. CONCLUSION

Phishing has becoming a serious network security problem, causing finical lose of billions of dollars to both consumers and e-commerce companies. And perhaps more fundamentally, phishing has made e-commerce distrusted and less attractive to normal consumers. In this paper, we have studied the characteristics of the hyperlinks that were embedded in phishing e-mails.

We then designed an anti-phishing algorithm, Link-Guard, based on the derived characteristics. Since Phishing-Guard is characteristic based, it can not only detect known attacks, but also is effective to the unknown ones. We have implemented Link Guard for Windows XP. Our experiment showed that Link Guard is light-weighted and can detect up to 96% unknown phishing attacks in real-time.

We believe that Link Guard is not only useful for detecting phishing attacks, but also can shield users from malicious or unsolicited links in Web pages and Instant messages.

REFERENCES

[1] Androutsopoulos, J. Koutsias, K.V. Chandrinos, and C.D. Spyropoulos. An Experimental Comparison of Naive Bayesian and Keyword-Based Anti-Spam Filtering with Encrypted Personal E-mail Message.In *Proc. SIGIR 2000*, 2000.

[2] The Anti-phishing working group. http://www.antiphishing.org/.

[3] Neil Chou, Robert Ledesma, Yuka Teraguchi, Dan Boneh, and John C.Mitchell. Client-side defense against web-based identity theft. In *Proc.NDSS 2004*, 2004.

[4] Cynthia Dwork, Andrew Goldberg, and Moni Naor. On Memory-Bound Functions for Fighting Spam. In *Proc. Crypto 2003*, 2003.

[5] EarthLink. ScamBlocker. http://www.earthlink.net/software/free/toolbar/.

[6] David Geer. Security Technologies Go Phishing. *IEEE Computer*, 38(6):18–21, 2005.

[7] John Leyden.Trusted search software labels fraud site as 'safe' http://www.theregister.co.uk/2005/09/27/untrusted search/.

[8] Microsoft.Sender ID Framework.http://www.microsoft.com/mscorp/safety/technologies/senderid/default.mspx.

[9] Netcraft. Net craft toolbar. http://toolbar.netcraft.com/.

[10] PhishGuard.com.Protect Against Internet Phishing Scams .http://www.phishguard.com/.

[11] Jonathan B. Postel. Simple Mail Transfer Protocol.RFC821:http://www.ietf.org/rfc/rfc0821.txt.

[12] Georgina Stanley. Internet Security-Gone phishing. http://www.cyota.com/news.asp?id=114.

[13] MengWengWong. Sender ID SPF. http://www.openspf.org/whitepaper.pdf. 7

