Phising Detection Using Link Guard Algorithm

Madhav Prasad[1] Vijay Maheshawari[2]
[1,2]Shobhit University, Meerut (India)

Abstract—*An end-host based anti-phishing algorithm which we call Link Guard, based on the characteristics of the phishing hyperlink. Since Link Guard is character-based, it can detect and prevent not only known phishing attacks but also unknown ones. We have implemented Link Guard in Windows XP, and our experiments indicate that Link Guard is light-weighted in that it consumes very little memory and CPU circles, and most importantly, 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.*

Keywords :- Phising, Link Guard, Window XP, APWG

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 (or instant messages), 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 conform 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. Within one to two years, the number of phishing attacks increased dramatically. [1]

II. EXISTING SYSTEM

A. Detect and block the phishing Web sites in time: If we can detect the phishing Web sites in time, we then can block the sites and prevent phishing attacks. It's relatively easy to (manually) determine whether a site is a phishing site or not, but it's difficult to find those phishing sites out in time. Here we list two methods for phishing site detection.

1. The Web master of a legal Web site periodically scans the root DNS for suspicious sites (e.g.www.1cbc.com.cnvs.www.icbc.com.cn).
2. Since the phisher must duplicate the content of the target site, he must use tools to (automatically) download the Web pages from the target site.

It is therefore possible to detect this kind of download at the Web server and trace back to the phisher. Both approaches have shortcomings. For DNS scanning, it increases the overhead of the DNS systems and may cause problem for normal DNS queries, and furthermore, many phishing attacks simply do not require a DNS name. For phishing download detection, clever phishers may easily write tools which can mimic the behavior of human beings to defeat the detection.

B. Enhance the security of the web sites:

The business Websites such as the Web sites of banks can take new methods to guarantee the security of users' personal information. One method to enhance the security is to use hardware devices. For example, the Barclays bank provides a hand-held card reader to the users. Before shopping in the net, users need to insert their credit card into the card reader, and input their (personal identification number) PIN code, then the card reader will produce a onetime security password, users can perform transactions only after the right password is input. Another method is to use the biometrics characteristic (e.g. voice, fingerprint, iris, etc.) for user authentication. For example, PayPal had tried to replace the single password verification by voice recognition to enhance the security of the Web site. [2], [3].

C. Block the phishing e-mails by various spam filters:

Phishers generally use e-mails as 'bait' to allure potential victims. SMTP (Simple Mail Transfer Protocol) is the protocol to deliver e-mails in the Internet. It is a very simple protocol which lacks necessary authentication mechanisms. Information related to sender, such as the name and email address of the sender, route of the message, etc., can be counterfeited in SMTP. Thus, the attackers can send out large amounts of spoofed e-mails which are seemed from legitimate organizations. The phishers hide their identities when sending the spoofed e-mails, therefore, if anti-spam systems can determine whether an e-mail is sent by the announced sender (Am I Whom I Say I Am?), the phishing attacks will be decreased dramatically.

D. Install online anti-phishing software in user's computers:

Despite all the above efforts, it is still possible for the users to visit the spoofed Web sites. As a last defence, users can install anti-phishing tools in their computers.

The Anti-phishing tools in use today can be divided into two categories:

- Blacklist/white list based
- Rule-based
- Category I: When a user visits a Web site, the anti-phishing tool searches the address of that site in a blacklist stored in the database. If the visited site is on the list, the anti-phishing tool then warns the users. Tools in this category include Scam Blocker from the EarthLink Company, Phish Guard, and Net craft, etc. Though the developers of these tools all announced that they can update the blacklist in time, they cannot prevent the attacks from the newly emerged (unknown) phishing sites.

– Category II: This category of tools uses certain rules in their software, and checks the security of a Web site according to these rules. Examples of this type of tools include Spoof Guard developed by Stanford, Trust Watch of the Geo Trust, etc. Spoof Guard checks the domain name, URL (includes the port number) of Web site, it also checks whether the browser is directed to the current URL via the links in the contents of e-mails. If it finds that the domain name of the visited Web site is similar to a well-known domain name, or if they are not using the standard port, Spoof Guard will warn the users. In Trust Watch, the security of a Web site is determined by whether it has been reviewed by an independent trusted third party organization. Both Spoof Guard and Trust Watch provide a toolbar in the browsers to notify their users whether the Web site is verified and trusted. It is easy to observe that all the above defense methods are useful and complementary to each other, but none of them are perfect at the current stage.

III. SYSTEM REQUIREMENTS FOR PROPOSED SYSTEM

A. Software Requirements

The minimum requirements for detection and prevention of phishing attacks are:

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

B. Hardware Requirements

The minimum hardware requirements are:

– Hard disk : 20 GB and above
– RAM : 256 MB and above
– Processor speed : 1.6 GHz and above

IV. LINK GUARD ALGORITHM

Phishing can generally occur with Banking websites or e-shopping websites. This project explains the implementation of the Link Guard algorithm using a mail-box system.There are three modules involved in this project:

– Creation of a mail system and database operations
– Composes, send and receive a mail
– Implementation of the Link Guard algorithm

1) MODULE 1

This module deals with the user interface for the home page, sign-in, sign-up and forgot your password pages. This module enables a new user to Sing-Up. It also enables an existing user to Sign-In. The user may use the Forget password link if he did forget his password. The password is retrieved on the basis of security question and answer given by the user. Database operation manages the users. Every time a new user signs in his details are written in to the database.

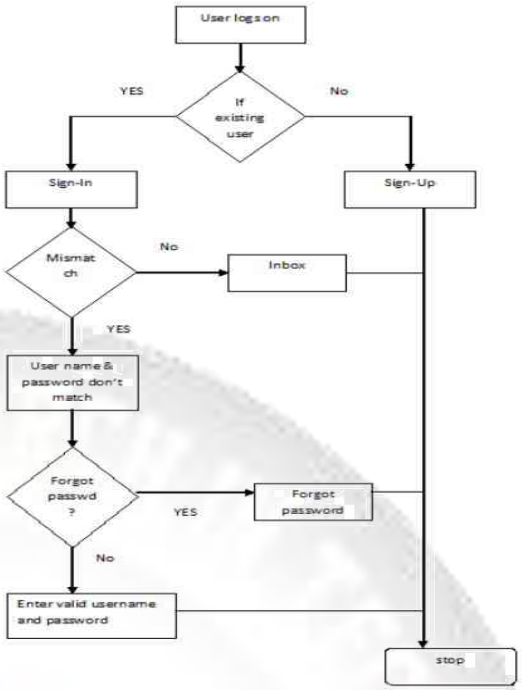


Fig.1: CREATION OF A MAIL SYSTEM AND DATABASE OPERATIONS

2) MODULE 2

The module 2 enables the user to compose and send a mail. It also allows the user to read a received mail. Once a mail is sent the date and the subject of the mail gets displayed. The received mail can be checked if it is phishing or not, the implementation of which is given in the next module. The compose mail option contains an option for spoof id. The spoof id allows the mail of the composer to be delivered with a different from address. This is being incorporated to demonstrate the Link Guard algorithm.[4],[5],[6].

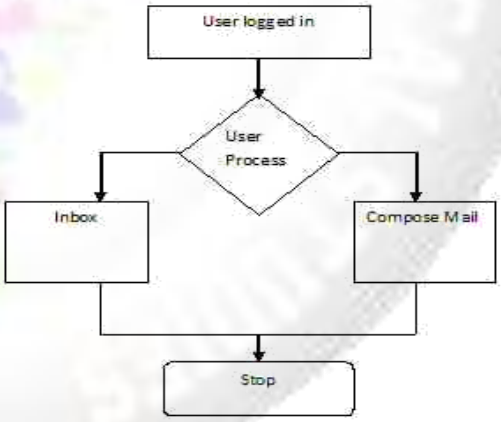


Fig 2: COMPOSES, SEND AND RECEIVE A MAIL

3) MODULE 3

The module contains the implementation of the Link Guard algorithm. It is possible for the user to add domain names and categorize them as either white list or black list under settings. Whenever a mail is detected as phishing the domain name in that mail automatically gets added as black list. The Link Guard algorithm checks if the domain names fall under any of the 5 categories of hyperlinks for phishing emails. It also refers to the database of black and white list entries and sets the status of the mail as either **Phishing** or

Non-Phishing. Once the mail is categorized as Phishing the user can take care that he does not open the link or submit any personal, critical information on to the website. [7],[8].

V. TESTING

Testing is a process, which reveals errors in the program. It is the major quality measure employed during software development. During testing, the program is executed with a set of conditions known as test cases and the output is evaluated to determine whether the program is performing as expected.

Software testing is the process of testing the functionality and correctness of software by running it. Process of executing a program with the intent of finding an error.

A good test case is one that has a high probability of finding an as yet undiscovered error. A successful test is one that uncovers an as yet undiscovered error. Software testing is usually performed for two reasons.

- Defect detection
- Reliability estimation

A. Testing Objectives

- Testing is a process of executing a program with the intent of finding an error.
- A good test case is one that has a high probability of finding an as yet undiscovered.
- A successful test is one that uncovers an as yet undiscovered error.

B. Testing Principles:

- All tests should be traceable to customer requirements.
- Tests should be planned large before testing begins.
- Testing should begin “In the Small” and progress towards “In the Large”.

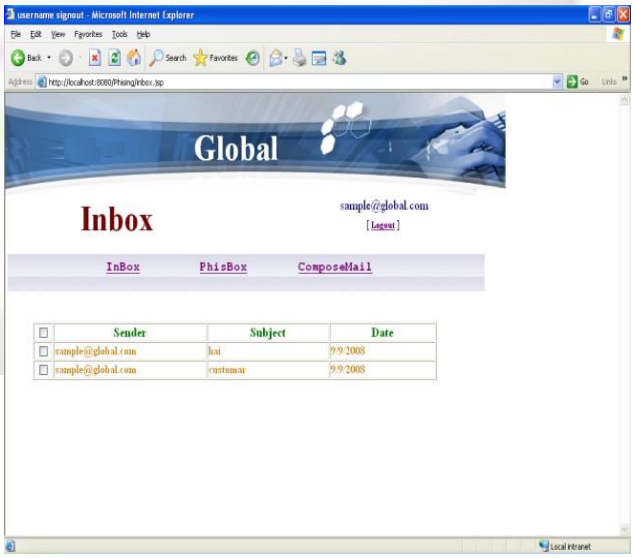
VI. RESULT & SCREENS

A. LOGIN PAGE:

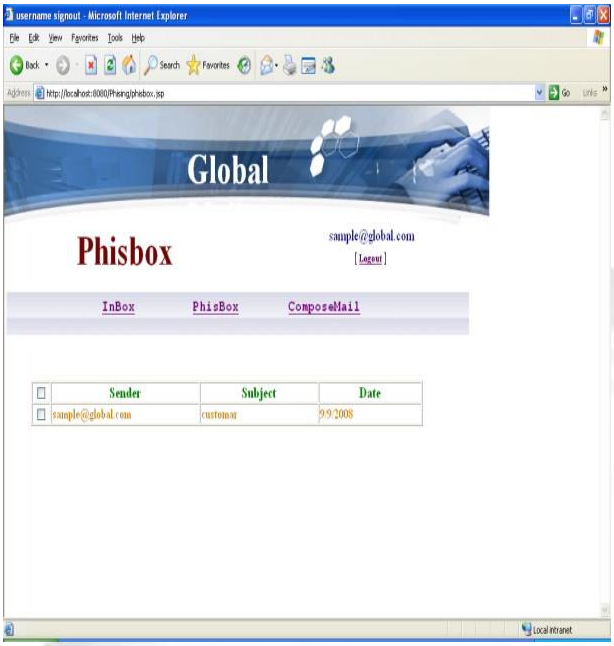
B. REGISTRATION FORM:

C. USER ID:

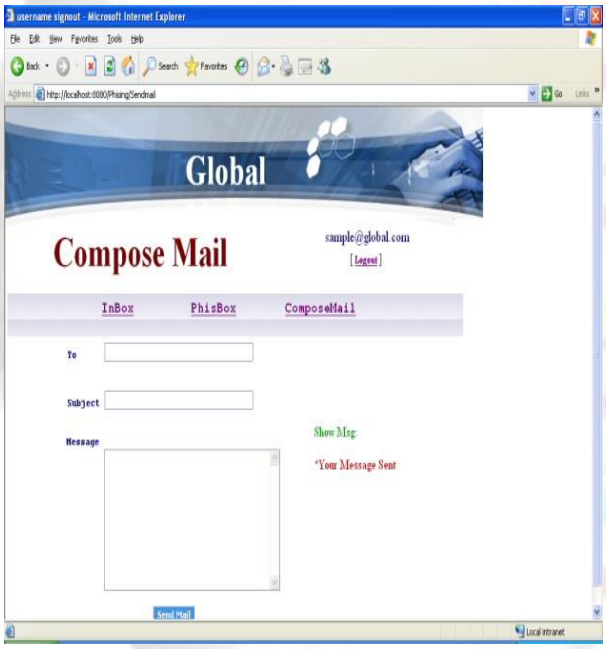
D. INBOX:



E. PHISBOX:



F. COMPOSE MAIL:



VII. 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. Our future work includes further extending the Link Guard algorithm, so that it can handle CSS (cross site scripting) attacks.

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 MoniNaor. On Memory-Bound Functions for Fighting Spam. In Proc. Crypto 2003, 2003.

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

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

[7] J. Leyden. Crooks harvest bank details from Net kiosk. The Register. http://www.theregister.co.uk/2003/01/27/crooks harvest bank details/, 2003.

[8] W. Liu, X. Deng, G. Huang, and A.Y. Fu. An Anti-phishing Strategy Based on Visual Similarity Assessment. In IEEE Internet Computing, Vol. 10, No. 2, "March/April" 2006.