Phishing Email Filtering Techniques A Survey

P.Rohini [#1], K.Ramya *[2]

[#]*M.E Student,* [*]*Assistant Professor,*
[1,2]*Department of Computer Science and Engineering, Hindusthan College of Engineering and Technology, Coimbatore, Tamilnadu, India.*

Abstract

The most interesting species of Internet fraud is Phishing. Email Phishing is a vulnerable activity which is referred as E-mail fraud, includes web link or form and Asks for confidential information such as password, account details. The email will be classified as phishing email and legitimate email by various phishing email filter techniques based on their functional activities. Various Anti phishing Mechanisms and tools are used for user's protection against this fraudulent act by using heuristics method and machine learning algorithm by (SVM) support vector machine classifier. The phishing problem is highly effective and no single solution exists to mitigate all the vulnerabilities effectively. This survey relies on recently developed anti phishing mechanisms and tools.

Keywords **– Phishing Email, legitimate Email, vulnerability, Anti-phishing, SVM, Heuristics, Machine Learning Algorithm.**

I. INTRODUCTION

Phishing is a new word retrieved from 'fishing' which refers that the attackers act as a fisher and the victim who is affected by this email phishing act as a hunt of fish. In other words "Ph" is replaced as "f" by a common hacker and is a dip into the original form of hacking, known as "phreaking". The Two different types of phishing attacks are Malware-based phishing and deceptive phishing. In The malware-based phishing type capturing the user confidential details, and send to the phisher. The Deceptive phishing, in which a phisher sends out deceptive emails pretending to receive from a reputable institution such as bank, industrial etc.. The frequently used attack method is to send e-mails to each victim, which considered and sent by banks, or ISPs. Here the user is also behalf of this kind of attack. (e.g.) the credit card password had been wrongly entered as long as , or they are providing upgrading services, to insist us visit their Web site to conform or modify our account number and password through the hyperlink provided in the e-mail. If we put the account number and password, the attackers can successfully collect our information at the server side, and are able to perform their next step actions such as withdraw money from our account. Phishing concept is familiar one but it's increasingly used attacker here to steal user information and perform business crime in recently.

Within one to two years, the number of phishing attacks increased dramatically. A Phishing emails are usually sent in large batches for time consumption. So we don't get type all recipients names out and send emails one-by-one. If we can't see our name, it'll be suspicious. Even if a link having name that you recognize somewhere in it, it doesn't mean that the link of real organization.

Roll your mouse over the link and see if it matches what appears in the email. If there is an inconsistency, we haven't click on the link. Also, legitimate websites where it is safe to enter personal information begin with "https" here the "s" stands for secure. Most probably we could get aware before provisioning our personal information asked through any email request which is probably a phishing attempt. For that purpose the browsers having toolbar as add on. The Net craft [1] toolbar for the Mozilla browser is to just send URL to their respective servers where all necessary processing is done. After finding the result it is sent back to the toolbar which indeed displays the result in that respective browser. Merely this process takes an amount of time to reduce this real time processing at end user level. And Black List is maintained by browsers like Google Chrome [2].

These Black list acts as a database of suspicious email or URL. Recently Multiple surveyor published their opinion about phishing email is that no banks and visa send emails asking for customer personal information by Jamieson survey. Also He noted there are 156 million phishing emails send out globally every --ECrime trends report [14] shows that phishing attacks are increasing at a rapid rate such as phishing in

Quarter 1 (Q1) of 2011 grew by 12% over that in Quarter 1 (Q1) of 2010.The phishing scam also fooling a very good friend of ours. The subject of "clever chase bank customer survey phishing scam" is getting $50 Reward from the Customer survey.

II. PHISHING EMAIL PREVENTION

Numerous anti-phishing Technologies are running as a real world process. All those Mechanisms are based on only a Specific Algorithm and Heuristics based. Yahoo keep their Anti Phishing Technique through their Sign-In Seal [3] which is a secret message or an image selected by the user that Yahoo displays on the user's computer . Every time the user visits Yahoo from the same machine and enables the user to make sure that they are on a genuine Yahoo site. Whenever we receiving an Email means we should check the following Features sometimes the impersonal greetings like 'Dear User,' or 'your email address' make us as a victim. E.g. PayPal email will always greet you by your first and last name only.

We have to consider the domain name which is crosscheck e-mail messages to verify their origin. The same anti- phishing handled by Gmail through the 'report spam' that any messages reported as spam get sent to a separate folder. Google's Anti-spam software is notified one more thing is this always requires the user's smart password only. The security key factor is a Password complexity of any of your online accounts The password should be more general, longer, more complex, and more random password such as 12 to 15 characters, with lower-case, upper-case, numbers, and special characters (with the exception of "!","@","#", and "$").

Google rolled out its 2-factor authentication In February of 2011 option in an attempt to beef up account security. We can access this setting by clicking the gear icon in the top right corner of our Gmail account screen. Spammers could be alerted that your address is valid and being monitored. They could send out more spam by redirection and URL redirection activity that the attackers have limited resources and usually keep reuse property only.

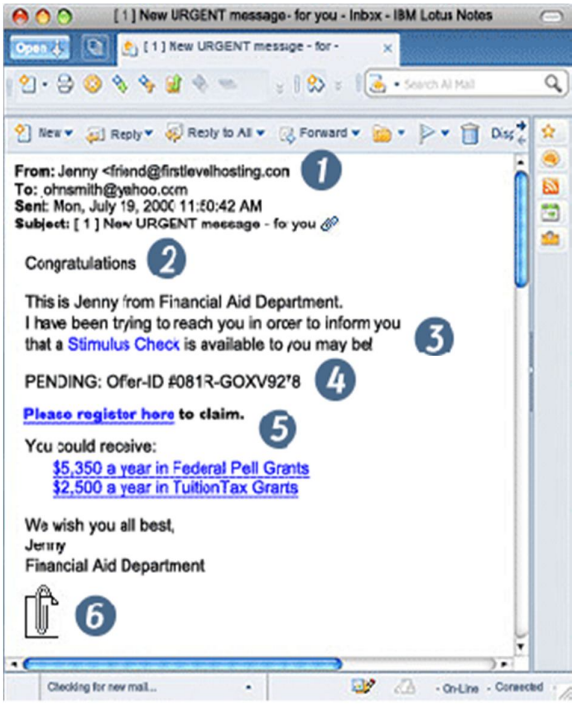

Fig 1. fake email

So their URL redirect chains frequently share the same URLs. The anti-phishing internet services are built into e-mail servers and web browsers and available as web browser toolbars such as Spoof Guard Toolbar1 [4], Trust Watch Toolbar2, and Net craft Anti-Phishing Toolbar3) [5]. The Spoof Guard developed by Stanford, it checks the domain name, URL (includes the port number) of a given Web site which shown up by the hyperlink in user email, it also checks whether the browser is directed to the current URL via the links in the contents of e-mails.

If it is not like that means the Spoof Guard will warn the users. In Trust Watch developed by of the Geo Trust, the security of a Web site is determined whether it has been reviewed by an independent trusted third party organization. Still there are multiple technologies are prevent the phishing email and phishing website.

III. LITERATURE SURVEY

A. Impact of Hyperlink in Email phishing

Obviously the anti-phishing security feature disables "hypertext links" inside e-mails which supposed to be pointing the phishing pages. To combat this problem, we compute the similarity of phishing pages and authentic pages at their

presentation level. In this image-based page matching and page classification we first take a snapshot of a suspect Web page and we treat it as an image throughout the detection process.

By using contrast context histogram (CCH) descriptors invariant information around discriminative key points [6] on the suspicious page will be captured. Then match the descriptors with those of authentic pages that are often targeted by attackers, where the pages are stored in a database compiled by users and authoritative organizations, such as the APWG. Matching CCH descriptors yields a similarity degree for a suspect page and an authentic page. Finally, the similarity degree between the two pages to determine whether the suspect page is a counterfeit. If the similarity degree between a suspect page and an authentic one is greater than a certain threshold, we consider the suspect page to be a phishing page for the authentic one or genuine if it's not a phishing page for any authentic pages in the database.

B. Modern email malware

The survey of the new modern email malwares[7], [8], [9], [10], [11] are focused on modeling the propagation dynamics which is a fundamental technique for developing counter measures to reduce email malware's spreading speed and prevalence. The most email malware extracted by two features only. "Reinfection", is one of the features which are occurred by infected user sends out malware copies whenever this user visits the malicious hyperlinks or attachments. In The Second "self-start" feature the malware copies are sends out by infected user when certain events like PC restart are triggered that can be infected multiple times. These two critical processes Reinfection and self-start solved by introducing a group of difference equations and virtual nodes, And SII (susceptible-infected-immunized) model [12] .Through this Process we can neutralize the modern email malware such as Melissa and Love letter from more vulnerability.

C. Anti phishing algorithm

An end-host character based Link guard anti-phishing algorithm in Windows XP depends on the characteristics of the phishing hyperlink such as,

1) The visual link and the actual link should not be same.

2) The attackers often use IP address instead of DNS name.

3) Special tricky encoding process used for the hyperlinks maliciously.

4) Fake DNS will be used instead of original domain name.

It can detect and prevent not only for known phishing attacks but also unknown ones. On regarding this algorithm, Link Guard is a light-weighted that it also consumes very little memory and little CPU cycles per second. And most importantly it is very effective in detecting phishing attacks with true positive rate. This method also relies on the black list and white list techniques for their suggestion whether the given information is phishing or not. These are all user based default settings.

TABLE I
SUMMARY COMPARISON OF THIS SURVEY WITH EXISTING WORK

No	Technique used	Advantages	disadvantages
1	contrast context histogram (CCH)	High level of accuracy – collect new type of features and detect suspicious pages	fraction of legitimate pages have already been Immunized .This method applicable for even legitimate before finalized.
2.	Modern email malware	SII act as a good scanner between user's mail transfer Agent (MTA) and mail user gent (MUA)	specially for reinfection and self start only
3.	Link guard Anti phishing algorithm	Fast in classification process	Need feed continuously
4.	Detection phishing emails using features decisive values	Provide clear idea about the effective level of each classifier on phishing email	18 features not enough to assign whether the given email is phishing or not.

-D. Phishing email detection by feature decisive values

The survey of this paper using FEFDV (feature existence & feature decisive value criteria).The phishing email could be detected by the 18 feature decisive values from email header and body content. And these feature values are analyzed by using training and testing data classifier. In this classifier the features are all considered as binary values as 0 and 1's. The feature existence will be referred as and 0, for not found case. Based on the features weight the phishing email will be detected [13].

IV. CONCLUSION

In this paper we surveyed a number of novel features that are particularly well-suited to avoid from phishing emails. This survey improves the awareness of the phishing emails problem, prevention and their solution space efficiently. Approaches are given in the literature still has much limitation on performance, especially from the phishing email attack. The security industry has taken up the challenges and today we have multiple solutions to the phishing email problem. We need to move towards effective solutions without over-burdening the user.

ACKNOWLEDGMENT

The work was supported in part by Ms.K.RAMYA Assistant Professor, Department of Computer Science and Engineering, Hindusthan College of Engineering and Technology Coimbatore.

REFERENCES

[1] http://www.netcraft.com/anti-phishing/.

[2] http://www.google.com/tools/firefox/safebrowsing

[3] Personalized Sign-In Seal – Yahoo Inc. https://protect.login.yahoo.com.Accessed: May 2, 2010

[4] M. Wu, R. C. Miller, and S. L. Garfinkel, "Do security toolbars actually prevent phishing attacks?" in Proc. SIGCHI Conf. Human Factors Comput. Syst., Montreal, QC, Canada, Apr. 2006, pp. 601–610.

[5] N. Chou et al., "Client-Side Defense Against Web-Based Identity Theft," Proc. Network and IT Security Symposium, Internet Soc., 2004; http://cryptoanford.edu/SpoofGuard/webspoof.pdf

[6] "Fighting Phishing with Discriminative Keypoint Features" Kuan-Ta Chen, Chun-Rong Huang, and Chu-Song Chen Institute of Information Science, Chen Columbia University Published by the IEEE Computer Society 1089-7801/09/$25.00 © 2009 IEEE INTERNET COMPUTING

[7] C.C. Zou, D. Towsley, and W. Gong, "Modeling and Simulation Study of the Propagation and Defense of Internet E-Mail Worms," IEEE Trans. Dependable and Secure Computing, vol. 4, no. 2, pp. 105- 118, Apr.-June 2007.

[8] Z. Chen and C. Ji, "Spatial-Temporal Modeling of Malware Propagation in Networks," IEEE Trans. Neural Networks, vol. 16, no. 5, pp. 1291-1303, Sept. 2005.

[9] C. Gao, J. Liu, and N. Zhong, "Network Immunization and Virus Propagation in Email Networks: Experimental Evaluation and Analysis," Knowledge and Information Systems, vol. 27, pp. 253-279, 2011.

[10] S. Wen, W. Zhou, Y. Wang, W. Zhou, and Y. Xiang, "Locating Defense Positions for Thwarting the Propagation of Topological Worms," IEEE Comm. Letters, vol. 16, no. 4, pp. 560-563, Apr. 2012.

[11] J. Xiong, "Act: Attachment Chain Tracing Scheme for Email Virus Detection and Control," Proc. ACM Workshop Rapid Malcode (WORM '04), pp. 11-22, 2004.

[12] IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, VOL. 11, NO. 4, JULY/AUGUST 2014 Modeling and Analysis on the Propagation Dynamics of Modern Email Malware Sheng Wen, Student Member, IEEE, Wei Zhou, Jun Zhang, Member, IEEE.

[13] International Journal of Advanced Research in Computer Science and Software Engineering Research Paper Available online at: www.ijarcsse.com "Detection Phishing Emails Using Features Decisive Values" Noor Ghazi M. Jameel Loay E. George Computer Science Institute Assistant Professor Sulaimani Polytechnic University College of Science Kurdistan Region.

[14] IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 15, NO. 4, FOURTH QUARTER 2013 A Survey of Phishing Email Filtering Techniques Ammar Almomani, B. B. Gupta, Samer Atawneh, A. Meulenberg, and Eman Almomani .