ISSN 2348 – 8034
Impact Factor- 3.155

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

Phishing Attack Protection Using Similarity Matrix

Phulwade Sayaji P.*[1] , Gunjal Yogesh S.*[2], Ghandat Arun B.*[3]

*[1], Lecturer, Department of Information Technology, Jaihind Polytechnic, India.
*[2], Principal, Department of Computer Engineering, Jaihind Polytechnic,kuran,india.
*[3], Assistance professor, Department of computer engineering, Amrutvahini College of engineering,

ABSTRACT

This paper provides an anti-phishing strategy that uses site similarity to detect potential phishing sites. Phishing is an attempt to steal confidential information such as bank username passwords, credit card information etc. for financial gain or fraudulent activities. In recent years, there is rapid growth of phishing attacks. There are various previous researches that provide anti-phishing strategies which include methods like blacklists of phishing sites or method that recognizes phishing patterns using statistical learning. The dynamic nature of phishing attacks, development of anti-phishing strategies is difficult task. This system assumes that each ISP will invoke search on each request that is requested by client, and searches for similar site using reverse searching. If there is any site having most of matching content then it will flag that site as phishing site and warns user about the event.

Keywords: *phishing, cyber security, ISP.*

I. INTRODUCTION

According to research conducted in [1], the graph of phishing attack victims is rising in each year and it is mostly targeted at financial sites. In phishing attack, attacker usually falsely convinces victim about for entering their confidential information such as, credit card number, password [2]. Another kind of phishing attack mainly uses malicious codes that specifically target user account information, also there are number of tools available in market for phishing attack. The causes for phishing vulnerability can be any of the following such as, weak authentication schemes browser vulnerabilities, security Flaws, unsecure desktop tools, Lack of user awareness, etc. The anti-phishing strategies falls into two types either, silently eliminating threat and warning user about the existence of threat. Various anti-phishing solutions are given in [2-3], it focuses on searching for webpage under attack to detect phishing attack, but this is very difficult task. The problem with existing system is that it suffers from high false alarms rates and most of the solutions keeps burden on client side, which is not effective as it may slow down their device. Further, it is important to respond dynamically to various phishing attacks. In this paper, we focus on anti-phishing detection model that can be used to safe guard user from viewing the attack page. This system uses site similarity measures for detecting phishing attack.

II. LITERATURE SURVEY

Most of the real world anti-phishing strategies use blacklists of phishing domains that are given to browser which refuses to visit. For example, IE7 browser is integrated anti-phishing solution that uses blacklist of web site domain. Browser periodically updates its list from Microsoft web server, thus making sure that phishing sites are not accessible to user. But, Microsoft has not released any details about its detection strategy. Google has also provides phishing tool such as Google Safe Browsing [25], Net Craft tool bar [18], eBay tool bar [7] and McAfee Site Advisor [12]. Google Safe browsing tool also uses blacklists of web domains for identify phishing sites while browsing. The problem with this strategy is that, non-blacklisted phishing sites are not detected. Spoof Guard searches for phishing signatures (e.g., obfuscated URLS) in web pages and if found raises alert. PwdHash is a tool that creates domain-specific passwords that useless if they are submitted to another domain.

While, Net Craft [3] extracts probability of phishing of a visited site by determining how old registered domain is? The database of registration of site is maintained by web servers, thus it can be accessed easily, but the problem with this approach is that new sites may be falsely identified as phishing sites. Site Advisor [4] is not truly an anti-phishing tool, but it is designed for protection against malware-based attacks (e.g., spyware, Trojan horses, etc.). It crawls the web, performs analysis and creates threat rating of each visited site. VeriSign [5] also uses blacklist of web sites which is updated from their web server [28]. Another approach for phishing attack detection uses feature extraction, also DNS based detection is also proposed in paper [7], genetic algorithm based detection [6], Thiyagarajan [9] used the anti-phishing approaches on the e-banking services, are some researches that had been proposed for the phishing webpage detection.

There are several approaches which tries to tackle root cause of phishing attack, such as email, thus preventing emails from reaching to user for that filters (such as Bayesian filters)are used in email systems. Some approaches attempt to solve the phishing problem at the e-mail level. But, filters are not sufficient to alone tackling the problem, some emails manages to be undetected. AntiPhish [10] monitors user actions and if it founds that user is entering any confidential information on web forms that is untrusted then it warns user and operation is canceled.

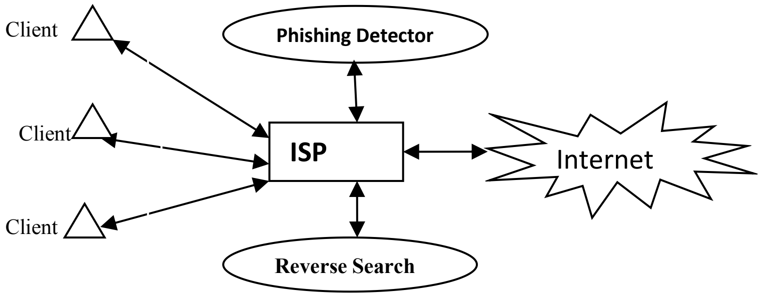

Figure 3.1 Typical System Architecture

III. SYSTEM ARCHITECTURE

The architecture of this system is given in figure 3.1. The architecture focuses on the structural connectivity of client for staying connected with outer world. Each client depends on its ISP for connection to the Internet, and thus ISP can be used for providing anti-phishing services. It minimizes risk of phishing attack, because large number of clients does not even aware of the phishing threat and its implications.

For successful detection of phishing attacks, this system is designed in such a way that, before fulfilling every request of client, ISP checks every web page for detecting phish pages, if the page is phished and the page is designed to take input from user such as credit card number, amount, etc. like confidential information, then ISP warns user about his activity and never delivers the faulty page, further ISP can help user for right selection of web link. The system finds overall similarity of web page with high ranked pages on internet, if the similar page is found having high similarity values then such page is flagged as phished page, weighting scheme is used for similarity detection in which more weight is given to those pages where a form parameters are found to be similar (as most of the phishing attacks are targeted at financial web sites and login pages). The anti-phishing strategies are never-ending battle against cyber threat.

IV. RESULT & DISCUSSION

When the similarity between a suspicious page and the query exceeds a threshold T for any of the three metrics, the system reports the page as a probable phishing page. To test our approach's ability to avoid false alarms, we treated the six true pages as queries and searched for visually similar pages in the test data set. Table 5 lists similarity values for eight pairs of true and phishing pages.

Table 4.1. Result

Sr No	Threshold	FPR	DR
1	0.3	94%	100%
2	0.4	90%	100%
3	0.5	84%	86%
4	0.6	75%	86%
5	0.7	60%	70%
6	0.8	40%	70%

V. CONCLUSION

This system designed by observing that phishing attack uses site similarity for falsely convincing user, thus client end solution is required. The ISP is good for this purpose, as it handles computational overhead by itself and result obtained shows (using a proxy server) that it can identify most of the phishing site using just site similarity matrix.

The future work will focus on using visible and invisible links on received email, thus preventing user to falsely open phishing site, as spoofed emails addresses are generally main source of phishing attacks.

REFERENCES

1. *M. Naor and A. Shamir, "Visual cryptography", in EUROCRYPT '94 Proceedings, Lecture Notes in Computer Science, Springer-Verlag, vol. 950, pp. 1-12, 1995.*
2. *Y. Liu, W. Liu, and C. Jiang, "User Interest Detection on Webpages for Building Personalized Information Agent," Proc. 5th Int'l Conf. Web-Age Information Management (WAIM 04), LNCS 3129, Springer-Verlag, 2004, pp. 280–287.*
3. *W. Liu et al., "Phishing Webpage Detection," Proc. 8th Int'l Conf. Document Analysis and Recognition, IEEE Press, 2005, pp. 560–564.*
4. *A.Y. Fu, W. Liu, and X. Deng, "EMD-based Visual Similarity for Detection of Phishing Web pages," presented at the Int'l Workshop on Web Document Analysis, 2005; available at www.cs.cityu.edu.hk/~liuwy/publications/WDA -EMD-Anti-Phishing-final.pdf.*
5. *Hicham Tout, William Hafner "Phishing: An identity-based ant phishing approach" in proceedings of international conference on computational science and engineering, Vancouver, BC, pages 347- 352, 2009.*
6. *V.Shreeram, M.Suban, P.Shanthi, K.Manjula, "Anti-phishing detection of phishing attacks using genetic algorithm" in proceedings of Communication control and computing technology(ICCCCT),IEEE international conference, Ramanathapuram , pages 447-450, 2010.*
7. *Juan Chen, Chuanxiong Guo," Online Detection and Prevention of Phishing Attacks (Invited Paper)"in proceedings of Communicational and networking in china, first international conference, Beizing, pages 1-7, 2007.*
8. *E. Kirda and C. Kruegel, "Protecting users against phishing attacks," The Computer Journal, 2005.*
9. *Thiyagarajan, P.; Venkatesan, V.P.; Aghila, G.; "Anti-Phishing Technique using Automated Challenge Response Method"', in Proceedings of IEEE- International Conference on Communications and Computational Intelligience, 2010.*
10. *Sun Bin.; Wen Qiaoyan, Liang Xiaoying, "A DNS based Anti-Phishing Approach," in Proceedings of IEEE- Second International Conference on Networks Security, Wireless Communications and Trusted Computing, 2010*