ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 11, November 2018

Survey on Anti-Phishing Techniques

Pooja Shinde[1], Radhika Khamkar[2], Neha Sapkal[3], Poojarani Gurav[4]

B.E. Student, Department of Information Technology Engineering, VPKBIET Engineering College, Baramati, Maharashtra, India[1]

ABSTRACT: The number of phishing attacks against internet services has seen an everyday increase inflicting, for instance, a negative impact on the flexibility of banking and money establishments to deliver reliable services on the net. This paper presents a survey of latest anti-phishing techniques. we've studied totally different anti-phishing techniques like component similarity based, visual approach based, Server log info based, white list information based etc. we tend to propose associate degree approach that mixes a personalised white listing approach with machine learning techniques. The whitelist is employed as filter that blocks phish web content used to assume inoffensive user behaviour. The phishing pages that don't seem to be found by the whitelist pass are further filtered employing a Support Vector Machine classifier designed and optimized to classify these threats. Careful comparison of different techniques used to detect phishing websites is presented during this paper.

KEYWORDS: Phishing, sensitive information, whitelist, Machine learning, Support Vector Machine.

I. INTRODUCTION

The number of phishing websites is increasing day by day and because of their high similarity with original websites there's massive threat to user's personal data. Anti-Phishing working group (APWG) have declared that, phishing could be a highly criminal mechanism using each social engineering and technical misrepresentation to steal consumers personal identification data, together with monetary account credentials. Usually, phishers trick users with spoofed e-mails that seem to be from a trusty supply like a bank or a esteemed commerce agency. There are two main categories of phishing attacks malware primarily based phishing and deceptive phishing. Malware-based phishing strategies install malicious code by exploiting security holes within the user's system. This code then records confidential and sensitive data and relays it to the phisher.

Technique of sending attractive and misguiding emails which seems like coming from authorised sources is used by attackers in deceptive phishing. These phishing mails attracts users to access a website which actually is a fake website carefully designed to trick users into divulging targeted sensitive data. The criminals in this type of attack use different techniques to attract user and steal their important data [1].

Some of the types of these techniques are:

- Social engineering- This includes all methods and scenarios invented by phishers to create a convincing context. Phisher requires good knowledge of human behavior.
- Imitation- This consists of duplication of websites that is same in appearance and look like original website.
- E-mail spoofing- This allows a attacker to spoof the source address of an electronic mail.
- URL hiding – This allows attacker to hide the URL behind text and masks the URL to which a user is redirected.

Phishing Techniques:-

- Spear Phishing: spear phishing is a more targeted attack in which the hacker knows which specific individual or organization they are after.
- Email/Spam: Email is sent to millions of users with a request to fill in personal details. These details will be used for illegal activities.

- Web Based Delivery: This attack is a type of as “man-in-the-middle” attack. The attacker is positioned in between the original website
- Link Manipulation: In Link control phisher sends a connection to a pernicious site. At the point when the client taps on the false connection, it opens up the phisher's site rather than the site made reference to in the connection.
- Phishing through Search Engines- Some phishing coercion include web indexes where the client is coordinated to items locales which may offer ease items or administrations. At the point when the client endeavor’s to purchase the item by entering the Visa subtle elements, it's gathered by the phishing site. There are numerous phishing bank sites offering charge cards or advances to clients at a low rate yet they are really phishing sites. We have portray different enemy of Phishing Detect the phishing utilizing page part similitude, which examinations URL of phishing page are store in whitelist database for expanding the precision .phishing pages regularly keep its CSS include like their objective pages techniques[1].Detection of Phishing site depends on the investigation of real site server log data [3].Detection of a phishing site utilizing Novel Algorithm This identification calculation can discover the greatest number of phishing URLs since it executes various tests, for example, Blacklist look Test[4].Detect Phishing site is based on whitelist arrangements and dispense with the trouble of overseeing and refreshing a lot of information by utilizing a customized whitelist. They stores whole LUI data as opposed to just a URL of a site in the white-rundown to give a more secure condition If page is does not has a place with the whitelist it isn't named a phishing page [5].

II. SURVEY OF ANTI-PHISHING TECHNIQUES

JAIN MAO, WENQIAN TIAN and ZHENKAI LIANG et.al [1] has designed systems which detect the phishing using page component similarity, which analyses URL of phishing page are store in whitelist database for increasing the accuracy .phishing pages typically conduct its CSS feature similar to their target pages. Based on this observation, a straightforward approach to detect phishing pages is to compare all CSS rules of two web pages and calculate similarity between two pages, It prototyped Phishing-Alarm as an extension to the Google Chrome browser and demonstrated its effectiveness in evaluation using real-world phishing samples.The Anti-Phishing Working Group (APWG ; www.antiphishing.org) [2] recently reported that the number of attacks is increasing by 50 present per month, with roughly 5 present of recipients falling victim to them. Phishing Web pages generally use similar page layouts and style sheet.Wenyin Liu, Xiaotie Deng, Guanglin Huang et.al Fu has proposed system which use anti-phishing tech that uses a visual approach to find bogus Web pages. The visual feature extractor module begins by segmenting the true and suspicious Web pages into sets of salient blocks.it checks Block-Level Similarity, Layout Similarity, Overall Style Similarity.[3]JUN HU, YUCHUN JI and HANBING YAN This method to detect Phishing website is based on the analysis of legitimate website server log information. every time a victim opens the phishing website, the phishing website will refer to the legitimate website by asking for resources. Then, there will be a log, which is recorded by the website server and from this logs Phishing site can be DetectedVARSHARANI RAMDAS, V.Y. KULKARNI and R.A.RANE [4] proposed a system to detect a phishing website using Novel Algorithm. Maximum number of phishing URLs are detected by this algorithm because it executes multiple tests such as Blacklist search Test, Alexia ranking test, and different URL features test. this solution is effective only for HTTP URLs.Y. Cao, W. Han and Y. Le et.a1[5] They maintain the accuracy of whitelist solutions and eliminate the difficulty of managing and updating large amounts of data by using a personalized whitelist. They stores entire LUIinformation rather than only a URL of a web site in the white-list to provide a more secure environment If page is does not belongs to the whitelist it is not classified as a phishing page. Checks the similarity between two pages if a page has a low similarity with the pages in the whitelist, this page is transformed into a feature vector to be classified .SVM used for classification [6], Naive Bayesian classifier is used to automatically maintain the white-list for the user" Blacklist/whitelist " solutions The blacklist solution is typically deployed as a toolbar or extension in web browsers. Examples of tools that implement this type of solution as Mozilla’s Firefox [7], Google’s safe browsing [8], and Phish Tank [9].Kirdaet.al [10] presents a novel browser extension, Anti-Phish that aims to protect users against spoofed web site-based phishing attacks. Anti-Phish tracks the sensitive information of a user and generates ominous whenever the user [attempts] to give away this information to a web site that is considered untrusted.Samuel Marchal, Jerome François et.al[11]This paper introduces Phish blast, an efficient phishing URL detection system bank on URL lexical analysis. The approach is based on the intra-URL

DOI: 10.15680/IJIRCCE.2018. 0611030

relatedness. This relatedness reflects the relationship among the words blended into a URL and particularly into the part of the URL that can be freely defined and the registered domain.

III. DISCUSSION

TABLE COMPARISON BETWEEN SERVER SIDE FILTERS AND CLASSIFIERS TECHNIQUES FOR PHISHING WEB DETECTION

No	Techniques used	Advantages	Disadvantages
1	Bag-of-words model	Build good scanner between users mail transfer Agent and mail user agent	Mostly working with supervised learning algorithm, rules are fixed ,week detection of attack
2	Multi Classifiers algorithms	Provide all idea about the effective level of each classifier on phishing emails	Mostly working with supervised learning algorithm ,rules are fixed ,week detection of attack
3	Module Based Classifiers	High level of accuracy and create new type of features.	Many algorithm and features are used for classification, large number of mail servers are used, time consuming.
4	Clustering of Phishing Email	Fast classification process.	Less accuracy because it depends on unsupervised learning.
5	Multi-layered System	High level of accuracy, through previous classifiers.	Time consuming.
6	Evolving Connectionist System	Fast and less consumption of memory, and high accuracy.	Need feed continuously.
7	Evolving White-list database	High level of accuracy	Test websites must be in whitelist database.

IV. CONCLUSION

In this activity we have studied various anti-phishing techniques those are used to detect suspicious phishing pages using analysis of URL. The comparative chart of various phishing techniques provides at a glance analysis of anti-phishing methods. By study of various methods we conclude that associate degree approach that mixes a personalised white listing approach with machine learning techniques can provide the better results. Machine learning approach will improve the accuracy of phishing detection with increase in phishing examples.

REFERENCES

1. JIAN MAO1,WENQIAN TIAN1, PEI LI1, TAO WEI2, AND ZHENKAI LIANG3 “Phishing-Alarm: Robust and Efficient Phishing Detection via Page Component Similarity” July 25, 2017.
2. Wenyin Liu, Xiaotie Deng, Guanglin Huang, and Anthony Y.Fu” An Antiphishing Strategy Based on Visual Similarity Assessment”.
3. Jun Hu, XiangzhuZhang,YuchunJi, Hanbing Yan, Li Ding, Jia Li and HuimingMeng Detecting Phishing Websites Based on the Study of the Financial Industry Webserver Logs
4. VarsharaniRamdasHawanna, V. Y. Kulkarni and R. A. Rane A Novel Algorithm to Detect Phishing URLs.
5. Y. Cao, W. Han and Y. Le, "Anti-phishing Based on Automated Individual Whitelist", DIM’08, October 31, 2008, Fairfax, Virginia, USA.
6. C. C. Chang and C.-J. Lin, "LIBSVM:a library for support vector machines". ACM Transactions on Intelligent Systems and Technology, 1–27 (2011).
7. Mozilla. Phishing protection, http://www.mozilla.com/enUS/firefox/phishingprotection/, query date: March 2011.
8. Googlesafe browsing: http://code.google.com/intl/fr/apis/safebrowsing/, query date: March 2011.

DOI: 10.15680/IJIRCCE.2018. 0611030

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 11, November 2018

9. http://www.phishtank.com/ query date: March 2011

10. E. Kirda and C. Kruegel, “Protecting users against phishing attacks with antiphish,” in Proceedings of the 29th Annual International Computer Software and Applications Conference - Volume 01, ser. COMPSAC ’05.Washington, DC, USA: IEEE Computer Society, 2005.

11. Samuel Marchal, Jerome François, Radu State, and Thomas Engel” PhishStorm: Detecting Phishing With Streaming Analytics”, IEEE, VOL. 11, NO. 4, DECEMBER 2014