Results of Real Time Detection of Phishing Websites

[*]Raj Kumar Singh, [#]Dr. Rashmi Jha

[*,#]Asst. Prof. IITM Janakpuri, New Delhi, India.

Abstract: Web Spoofing lures the user to interact with the fake websites rather than the real ones. The main objective of this attack is to steal the sensitive information from the users. The attacker creates a 'shadow' website that looks similar to the legitimate website. This fraudulent act allows the attacker to observe and modify any information from the user. This paper proposes a detection technique of phishing websites based on checking Uniform Resources Locators (URLs) of web pages. The proposed solution is able to distinguish between the legitimate web page and fake web page by checking the Uniform Resources Locators (URLs) of suspected web pages. URLs are inspected based on particular characteristics to check the phishing web pages. The detected attacks are reported for prevention. The performance of the proposed solution is evaluatedusing Phistank and Yahoo directory datasets. The obtained results show that the detection mechanism is deployable andcapable to detect various types of phishing attacks maintaining a low rate of false alarms.

Keywords: Phishing Attack; URL; Real Time Model; Phishing Detection

I. INTRODUCTION

Social engineering attack is a common security threat usedto reveal private and confidential information by simplytricking the users without being detected [1]. The mainpurpose of this attack is to gain sensitive information suchas username, password and accounts numbers. According to[2], phishing or web spoofing technique is one example ofsocial engineering attack. Phishing attack may appear inmany types of communication forms such as messaging,

SMS, VOIP and fraudster emails. Users commonly havemany user accounts on various websites including socialnetwork, email and also accounts for banking. Therefore,the innocent web users are the most vulnerable targetstowards this attack since the fact that most people are

unaware of their valuable information, which helps to makethis attack successful.Based on the report prepared by the Anti-Phishingworking group organization [2], there were about 163,333phishing attacks reported in 2014. A recent study byMcAfee Lab [3] showed that there were about 30,000,000new suspected URLs in Quarter 3 for the year 2014. Thesereports also showed that web browser was classified as thetop most network threat which was about 26% compared tothe other network threats. For some crime groups, phishingattack is actually a business. Billions of dollars have beenreported stolen from banks in US, Russia and Eastern

Europe.Typically phishing attack exploits the socialengineering to lure the victim through sending a spoofedlink by redirecting the victim to a fake web page. Thespoofed link is placed on the popular webpages or sent viaemail to the victim. The fake webpage is created similar tothe legitimate webpage. Thus, rather than directing thevictim request to the real web server, it will be directed tothe attacker server. Figure 1 shows the steps involved in web spoofing attack.

There are many researchers conducted to detect web spoofing attacks. However, these researches are not effective enough to stop the sophisticated attack of web spoofing. The use of various media communication such s social network leads to the increase of the numbers of attacks. According to [4], 70% of successful phishing attacks are launched through social network. In fact, the lack of awareness and education on web spoofing attack causes the fall of the victims. Inability to distinguish between the fake and legitimate web pages is still a challenge in the existing prevention solutions of web spoofing.

 DOI : 10.35291/2454-9150.2019.0228

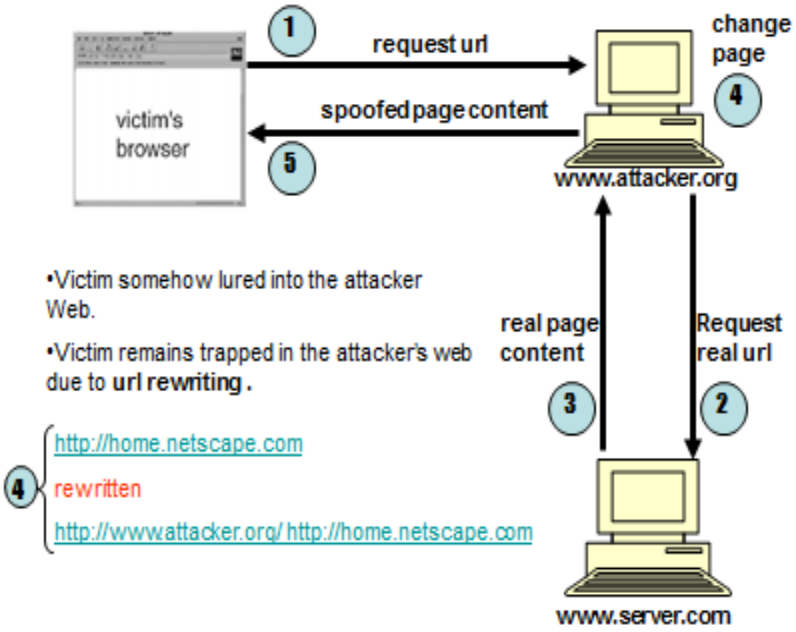

Figure.1: Steps involved in web spoofing attack

II. PHISHING LIFECYCLE

There are various phases to the phishing cycle. However, there are three main phases in phishing cycle repeated by various phishers [1, 2, 4-9]. In first phase, the phisher explores organizations and selects a target and then, creates a phishing website and send numerous spam emails among the various users in Internet community. Second phase starts with reading of these emails. Whenever the user "bites" on the phish i.e. click on the link, third phase starts and user is redirected to the phishing site. In this section, we briefly described about the phishing campaign in which phisher uses the advantage of ignorance about the communication channel in common users, as described in figure 11. In mailing system, every email first passes through the DNS based blacklist filters. If the domain of sender is found in blacklist, the email is blocked before reaching the SMTP mail server. Based on structural properties of emails, various solutions filters email before it reaches to the user's inbox. There are also various solutions available to check emails based on features of any email on client side. In case of phishing webpages, the links are embedded in emails sent to the user or any other advertisement. There are various solutions available on the client side as Internet is vast enough to control it. Some blacklist-based applications block the website if domain falls under blacklist. Unlike the blacklist solution for emails that block emails before they reach the SMTP mail server, it blocks the website when browser of client side request for the URL mentioned in the list. Some more solutions like heuristic feature and visual similarities block the webpage only when the browser request for any phishing webpage [82].

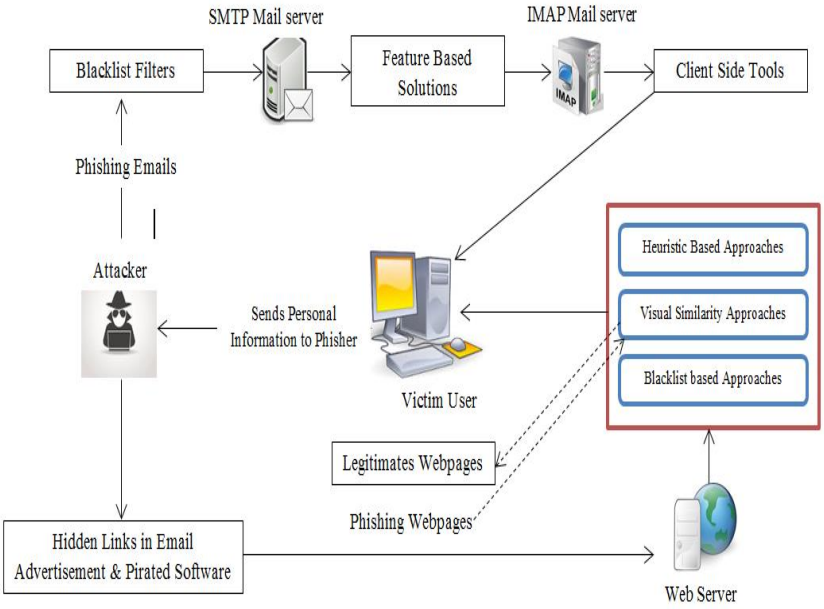

Figure 2 Lifecycle of phishing attacks based on phishing emails and phishing website

 DOI : 10.35291/2454-9150.2019.0228

III. LITERATURE REVIEW

Ms.ShwetaDasharathShirsat(2018) Phishing has accumulated enormously over previous few years and it's become a heavy threat to international security and economy. Existing literature managing the matter of phishing is scarce. Phishing may be a deception technique that uses a mixture of technology and social engineering to accumulate sensitive data like on-line banking passwords, mastercard or checking account details [2]. Phishing will be done through emails and websites to gather lead. Phishers style deceitful web sites that look almost like the legitimate websites and lure the user to go to the malicious website. Therefore, the users should bear in mind of malicious websites to safeguard their sensitive information. But it's terribly tough to tell apart between legitimate and pretend web site particularly for untechnical users what is more, phishing sites are growing speedily. The aim of this paper is to demonstrate phishing detection victimization mathematical logic and deciphering results victimization totally different defuzzification ways [8].

Chuan Pham, Luong A. T. Nguyen (2018) Phishing could be a criminal activity that steals victims' personal data mistreatment dishonourable emails or pretend websites. The word "phishing" is originated from the word "fishing" on-line users are often simply deceived into getting into their personal data as a result of phishing websites are extremely just like real ones. Maliciously, by making phishing sites, "phishers" use variety of techniques to fool their victims, together with email messages, instant messages, forum posts, phone calls, and social networking data Phishing leads to severe economic loss everywhere the globe, and phishing sites also are growing apace in amount and quality. per reports from the Anti-Phishing working party, the quantity of phishing attacks is increasing by five-hitter monthly. Fig. one illustrates the urgency and importance of phishing identification in fashionable society, that relies on a phishing web site report received within the half-moon of 2016 . First, mobile users check their emails and use internet browsers a lot of often than desktop users thence, it's exhausting for users to pick out if associate incoming link is legitimate or not. Third, existing anti-phishing tools (e.g., default plug-ins on internet browsers or native anti-phishing applications) are inefficient in terms of detection (this are going to be analysed concretely later in Section III), and mobile users could also be exposed to phishing attacks once participating in usual behaviors. per the report, mobile users are thrice a lot of probably to submit their login data than desktop users do. Therefore, preventing phishing attacks against terminal users could be a vital issue within the edge of networks [9].

Tianrui Peng, Ian G. Harris (2018) Phishing could be a kind of social engineering attack that focuses on gaining sensitive info by disguising as a trustworthy entity. Electronic communications, like email or text message area unit common platforms for delivering phishing attacks. Phishing has been shown to be a good attack over the years, deceiving a broad vary of individuals. Attackers area unit typically disguised as common social websites, banks, directors from IT departments or common searching websites. These emails could lure users to click on links to initiate malware downloads, or enter personal info into a malicious web site that includes a similar look to a legitimate one. Most automatic phishing email detection approaches admit email data, knowledge related to emails that isn't associated with the linguistics that means of the text message. many approaches examine the URLs contained within the message. There area unit many phishing detection approaches that assess text by sorting out the presence of specific words in every sentence has extensively utilized grammar parsing to infer malicious intent [10].

MuhammetBaykara, ZahitZiyaGürel (2018) Phishing is a form of cybercrime where an attacker imitates a real person / institution by promoting them as an official person or entity through e-mail or other communication mediums. In this type of cyber-attack, the attacker sends malicious links or attachments through phishing e-mails that can perform various functions, including capturing the login credentials or account information of the victim. These e-mails harm victims because of money loss and identity theft. In this study, a software called "Anti Phishing Simulator" was developed, giving information about the detection problem of phishing and how to detect phishing emails. With this software, phishing and spam mails are detected by examining mail contents. Classification of spam words added to the database by Bayesian algorithm is provided [11].

SriendraDeshan Ilangakoon, Abeywardena K.Y(2018) This paper evaluates the background analysis to spot the chance of employing a new vector of social engineering attack employing a psychological thought that so far had been solely employed in selling and promotional campaigns. imperceptible and supraliminal messages are studied by academe with relevancy its ability to influence individual behavior. Social engineering attacks are outlined because the art of manipulating folks into playacting actions or divulging wind. Most of recent social engineering attacks rely upon phishing and spear phishing attacks. This paper explores the chance of distinguishing a correlation between the on top of mentioned psychological ideas and phishing/spear phishing attacks within the domain of cyber security[12].

 DOI : 10.35291/2454-9150.2019.0228

IV. RESULTS TESTING AND EVALUATION

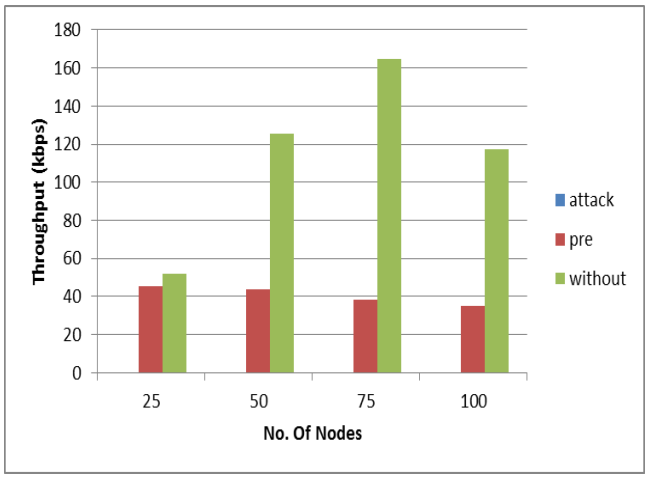

Figure.(3)Throughput under Attack, Prevention and Without Attack

No. of Nodes	Attack	Prevention	Without
25	0	45.50	51.88
50	0	43.75	125.46
75	0	38.50	164.73
100	0	35.0	117.39

Table: (1) Throughputs

4.1.Analysis of Throughput:

Higher value of throughput ensures large number of data packets successfully received at the Destination node. From the above figure it analysed that under replica node attack the Throughput of TAODV is more nearly similar to normal AODV, ascompared to AODV under replica node attack. Figure shows with increasing the number of nodes, throughput of network also increases.

4.1.1 Residual Energy:

Figure shows the Residual Energy under Node Replication attack detection and its prevention through Trust based mechanism i.e. AODV and TAODV for the various node densitsity.

4.1.2 Analysis of Residual Energy: It is the total amount of remaining energy by the nodes after the completion of Communication or simulation. If a node is having 100% energy initially and having 70% energy after the simulation than the energy consumption by that node is 30%.The unit of it will be in ms.

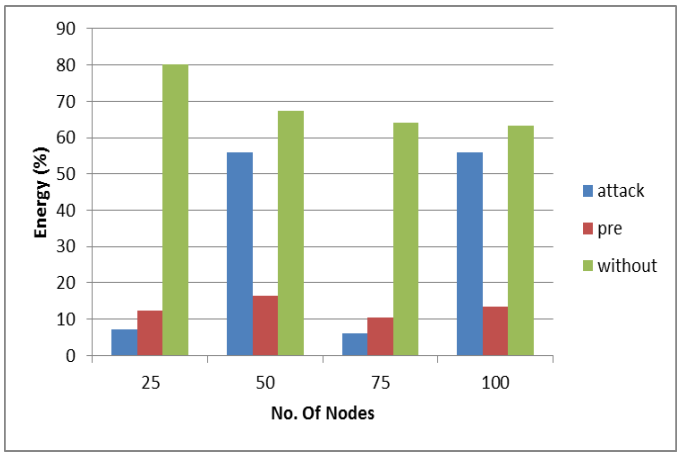

Figure (4) Residual Energy under Attack, Prevention, Without attack

No of Nodes	Attack	Prevention	Without
25	7.069349	12.45165	80.19328
50	56.06935	16.45165	67.27762
75	6.069349	10.45165	64.1895
100	56.06935	13.45165	63.30197

Table(2) Residual Energy

V. CONCLUSION AND FUTURE WORK

Lack of awareness on phishing education makes the attacksuccessful. Even with the help of few indicators used by thebrowser such as pad lock identification, lock icon, and siteidentity button, the user still cannot identify the attack. Webspoofing attack is not easy to detect. Even with the newestsecurity prevention method, these attacks still occur. Themain aim of this study is to help the users especially todifferentiate between the legitimate and phishing web pagesby using URL as an indicator. Finding of this researchdemonstrates its ability to identify the fake webpages basedon their URLs.As a conclusion, the most important way to protect theuser from phishing attack is the education awareness.Internet users must be aware of all security tips which aregiven by experts. Every user should also be trained not toblindly follow the links to websites where they have to entertheir sensitive information. It is essential to check the URLbefore entering the website.

There are a few limitations in this work. The accuracyof this heuristic-based depends on the discriminativefeatures that may help in distinguishing the type of websitewhether it is a legitimate or phishing site. This study onlychecks the validity of Universal Resource Locator (URLs)based on a few characteristics for detecting phishing attack.Future works of this study will include the automaticdetection of the web page and the compatibility of theapplication with the web browser. Additional work also canbe done by adding some other characteristics todistinguishing the fake web pages from the legitimate webpages. PhishChecker

 DOI : 10.35291/2454-9150.2019.0228

application also can be upgraded intothe web phone application in detecting phishing on themobile platform.

REFERENCES

[1] Ludl, C., McAllister, S., Kirda, E., & Kruegel, C. (2007). Onthe effectiveness of techniques to detect phishing sites.In Detection of Intrusions and Malware, and VulnerabilityBAssessment (pp. 20-39). Springer Berlin Heidelberg.

[2] Anti-Phishing Working Group Phishing, (2014). Anti-Phishing Working Group Phishing Trends Report. [Online]Available at: https://apwg.org/ [Accessed 30 Mar. 2015].

[3] McAfee Labs Threats Report: February 2015. Retrievedfrom http://www.mcafee.com/us/resources/reports/rpquarterly-threat-q42014.pdf.

[4] Jagatic, T. N., Johnson, N. A., Jakobsson, M., &Menczer, F.(2007). Social phishing. Communications of the ACM,50(10), 94-100.

[5] Why HTTPS and SSL are not secure as you think (2014,March 12) .Retrieved fromhttp://scottiestech.info/2014/03/12/why-https-and-ssl-arenot-as-secure-as-you-think.

[6] Zhang, Y., Hong, J. I., &Cranor, L. F. (2007, May).Cantina: a content-based approach to detecting phishing websites. In Proceedings of the 16th international conference onWorld Wide Web (pp. 639-648). ACM.

[7] Yasin Sönmez, TürkerTuncer,"Phishing Web Sites Features Classification Based on Extreme Learning Machine",978-1-5386-344@IEEE2018

[8] Ms.ShwetaDasharathShirsat "DEMONSTRATING DIFFERENT PHISHING ATTACKS USING FUZZY LOGIC",Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018) IEEE Xplore Compliant - Part Number: CFP18BAC-ART; ISBN:978-1-5386-1974-2.

[9] Chuan Pham, Luong A. T. Nguyen, "Phishing-Aware: A Neuro-Fuzzy Approach for Anti-Phishing on Fog Networks",1932-4537 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

[10] Tianrui Peng, Ian G. Harris, "Detecting Phishing Attacks Using Natural Language Processing and Machine Learning", 0-7695-6360-0/18/$31.00 ©2018 IEEE DOI 10.1109/ICSC.2018.00056.

[11] MuhammetBaykara, ZahitZiyaGürel ,"Detection of phishing attacks" 978-1-5386-3449-3/18/$31.00 ©2018 IEEE.

[12] SriendraDeshan Ilangakoon, Abeywardena K.Y, "The Use of Subliminal and Supraliminal Messages in Phishing and Spear Phishing based Social Engineering Attacks; Feasibility Study", 978-1-5386-5495-8/18/$31.00 ©2018 IEEE.

 DOI : 10.35291/2454-9150.2019.0228