ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

Admonishing and Alerting Malicious Webpages in Real Time

Mrs. L. Paul Jasmine Rani1, V.Naveen Kumar2, S.Arokiya Joseph3, M.Mohammed Aslam4,E.Harish5
1Assistant professor, 2, 3, 4UG scholar, Computer Science and Engineering, Rajalakshmi Institute of Technology, Chennai, India.
Pauljasminerani.l@ritchennai.edu.in[1], Naveenkumarprt97@gmail.com[2], josepharokiya@gmail.com[3], aslamrockers123@gmail.com[4], hhash013@gmail.com [5]

ABSTRACT

The World Wide Web has become an inseparable part of millions of people who use online services e.g. online banking, online shopping, social networking, e-commerce, and store and manage user sensitive information, etc. In fact, it is a popular tool for any class of user over the Internet. Rich Web based applications are available over the World Wide Web to provide such types of services. At the same time, the Web has become an important means for people to interact with each other and do business. This is the positive side of this technology. Unfortunately, the Web has also become a more dangerous place. The popularity of World Wide Web has also attracted intruders and attackers. These intruders abuse the Internet and users by performing illegitimate activity for financial profit. The Web pages that contain such types of attacks or pernicious code are called as malicious Web pages. While the existing approaches are good indicators in detecting malicious Web pages, there are still open issues in malicious Web page detection techniques. In this paper, an extensive survey of existing malicious Web pages detection approaches and features have been mentioned.

Keywords: *Malicious Web pages, Detection, Web Page Features, Machine Learning.*

I.INTRODUCTION

The internet exploits are the challenging issues of the net community. When the client visits the malicious web location the exploit is started through different highlights (lexical, space, way, web substance and hyperlink, etc). To anticipate the client against get to the malevolent websites, a few mechanized examination and discovery strategies have been proposed. Malicious web substance has ended up the essential apparatus utilized by aggressors to perform exploits on the web. As the client gets to a noxious server, the server conveys the exploit to the client as portion of its reaction. A web server that dispatches so-called drive-by-download exploits on web browsers is a common case. As the internet browser demands substance from a web server, the server returns an misuse inserted in the internet pages that permits the server to pick up total control of the client framework. The attackers draw the guest to get to malicious web locals and they take vital data from the client machine or introduce the spyware for abuses.

Phishing is the title of road. It can be characterized as the way of double dealing of an organizations client to communicate with their private data in an unsatisfactory behavior. The reason or objective behind phishing is information, cash or individual data taking through the fake site. Phishing is the act of endeavoring to secure data such as usernames, passwords, and credit card subtle

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

elements (and some of the time, in a roundabout way, cash) by disguising as a dependable substance in a electronic communication.

For this a few relationship between the enlist space rest of the URL are consider too intra URL tenacious is consider which offer assistance to cleaning wish between phishing or non phishing URL. Along with the ever expanding untrustworthiness through phishing tricks, organizations are getting more consideration from their clients with respect to the security of their individual data. [5] discussed about a Secure system to Anonymous Blacklisting. The secure system adds a layer of accountability to any publicly known anonymizing network is proposed. Servers can blacklist misbehaving users while maintaining their privacy and this system shows that how these properties can be attained in a way that is practical, efficient, and sensitive to the needs of both users and services. This work will increase the mainstream acceptance of anonymizing networks such as Tor, which has, thus far, been completely blocked by several services because of users who abuse their anonymity. In future the Nymble system can be extended to support Subnet-based blocking. If a user can obtain multiple addresses, then nymble-based and regular IP-address blocking not supported. In such a situation subnet-based blocking is used. Other resources include email addresses, client puzzles and e-cash, can be used, which could provide more privacy. The system can also enhanced by supporting for varying time periods.

According to the Symantec Web Security report discharged in 2014, malicious Web pages are presently the essential vector for noxious exercises over the Web and a few of the highlights from the risk scene of 2014 are[20]:

1. 91% increment in focused on exploits campaigns in 2013.
2. 62% increment in the number of breaches in 2013.
3. Over 552M personalities were uncovered by means of breaches in 2013.
4. 23 zero-day vulnerabilities discovered.
5. 38% of versatile clients have experienced versatile cybercrime in past 12 months
6. Spam volume dropped to 66% of all email traffic.
7. 1 in 392 emails contain a phishing attacks.
8. Web-based exploits are up 23%.
9. 1 in 8 true blue Websites have a basic vulnerability

This paper is structured as follows, In section 2, we present the Literature Review. Section 3 covers the web attacks taxonomy, focusing on the methods and Web page features used by researchers for classification of Web pages as malicious or benign. In section 4, we present the report of survey. The key principles of malicious Web pages detection in Section 5 and In Section 6, we present our Future Work and In Section 7 is our Conclusion.

II. LITERATURE REVIEW

Numerous analyst have proposed diverse strategies for classification and location of malevolent Web pages and location of diverse Webpage exploits like drive-by downloads, noxious JavaScript xploi8nts, cross-site scripting exploits, code infusion exploits, SQL infusion exploits, etc

A Guide approach finding Malicious web pages:

Luca Invernizzi and et al., [29][31] propose a novel approach whose objective is to move forward the viability of the look handle for malicious web pages. they use a seed of known, noxious web pages and extricate characterizing similitudes that these pages share. At that point, they utilize the framework of look motors and the information that they have collected to rapidly distinguish other pages that

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

appear the same characteristics and, in this way, are moreover likely malicious. They have executed this approach in a apparatus, called EVILSEED, and approved it on large-scale datasets. They assumes that EVILSEED can recover a set of candidate web pages that contains a much higher rate of malicious web pages, when compared to arbitrary slithering(and indeed to comes about returned for manually crafted, malevolent inquiries). Hence by utilizing EVILSEED, it is conceivable to make strides the adequacy of the malicious page disclosure prepare. [10] discussed about Enhancement of TCP Throughput using enhanced TCP Reno Scheme. Mobile Ad-Hoc Networks (MANETs) have been an area for active research over the past few years due to their potentially widespread application in military and civilian communications. Based on the analysis, we proposed two simple yet effective ways, namely, TCP Few and ROBUST, to improve the system performance. It was shown via computer simulation that TCP performance can be significantly improved without modifying the basic TCP window or the wireless MAC mechanism. Thus, the TCP window mechanism can still be a viable solution for IEEE 802.11 ad-hoc networks.

Detecting Suspicious URLs in Twitter Stream:

Sulabh.S and et al., [4][17][36] propose, Caution Fowl is strong when securing against conditional redirection, other than existing highlights a few unused highlights named relationship highlights are presented. These highlights offer assistance recognizing noxious and generous URLs in a better way. This ventures works in genuine time. Thus the time taken for discovery is exceptionally less. Comes about appear that Caution Feathered creature is much speedier and productive compared to twitter's discovery framework. But this work cannot dispose of or square when a page is recognized malevolent. This work can be improved in such a way that it can offer assistance distinguishing phishing pages effectively.

Large-Scale Automatic Classification of Phishing pages:

Colin Whittaker [11][30], In these papers they depict our large scale framework for naturally classifying phishing pages which keeps up a wrong positive rate underneath 0.1%. Our classification framework analyzes millions of potential phishing pages every day in a division of the time of a manual survey prepare. By naturally overhauling our boycott with our classifier, we minimize the sum of time that phishing pages can stay dynamic some time recently we ensure our clients from them. Indeed with a culminate classifier and a strong framework, we recognize that our boycott approach keeps us interminably a step behind the attackers. We can as it were distinguish a phishing page after it has been distributed and unmistakable to Web clients for a few time. Be that as it may, we accept that in the event that we can give a boycott total sufficient and rapidly sufficient, we can drive attackers to function at a misfortune and desert this sort of Web wrong doing.

Vaibhav V and et al., [5] propose that, Extortion client utilize social systems are frequently utilized to gather individual data as well as collect money related information of client. Extraction of include set for finding skeptical URL utilizing Bayesian classification in social organizing location [15][29]

Two sorts of irregularity highlights were proposed: space inconsistency and social irregularity highlights. Space irregularity highlights are utilized to recognize conceivable malevolent spaces based on lexical and notoriety components, though social irregularity highlights speak to odd client behaviors in social communications.

Stefan Savage and et al., [12] portrayed an approach for classifying URLs naturally as either

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

malevolent or generous based on administered learning over both lexical and host based highlights. We contend that this approach is complementary to both boycotting which cannot anticipate the status of already concealed URLs and frameworks based on assessing location substance and behavior which require going to possibly unsafe locales.[5] In advance, we appear that with fitting classifiers it is attainable to naturally filter though comprehensive highlight sets(i.e., without requiring space mastery) and recognize the most prescient highlights for classification.

III.WEB ATTACKS TAXONOMY

Unused Web-based exploits are coming out each day, this is driving businesses, communities and people to take security issues genuinely. Taking after are a few of the common Web-based exploits against Websites and Web applications. The proposed scientific categorization is outlined in Figure 1[1].

3.1. CLIENT-SIDE ATTACK

Jung-Ying Lai and et al., [1][14][27] proposed that, the browser can show the Internet provides service contents. It is supplied basic HTML form that most of browser is not compatible Web standards and technique. The most of browser to follow these formats such as HTTP, HTTPS, XHTML, XML, and CSS. For example: IE, Firefox, and Safari. The attack interface not only limits in Browser, but also possibly is the short shell code. In addition to the browser, the other clientside manner class with Content. [37]Each HTTP method is used means possible causes Web attack. These attacks sent the server-side that could be the platform or markup language to receive. If the platform or markup language exist vulnerability, it will cause damage in theWebServer.

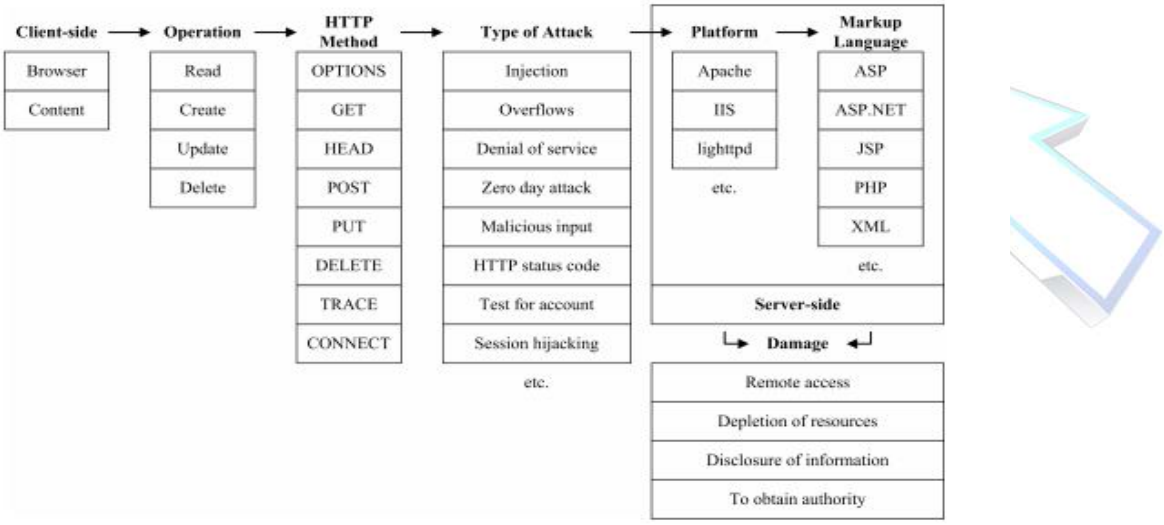


Fig 1: TAXONOMY OF WEB ATTACKS

3.2. DRIVE-BY-DOWNLOAD ATTACK

Dharmaraj rajaram and et al,[2][36][40] proposes that, A drive-by download can be initiated by simply visiting a Web site or viewing an HTML E-mail message. Basically, these attacks are usually downloaded run in the background in a manner that is

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

invisible to the user. In Aprils 2007, analysts at Google found hundreds of thousands of Web pages that started drive-by downloads. One in ten pages was found to be suspect. Sophos analysts in 2008 detailed that they were finding more than 6,000 unused tainted Web pages each day, or around one each 14 seconds [13][39][42].

D.Canali and et al., [14][33][34] potrays that the Drive-by downloads may happen when visiting a website, viewing an e-mail message or by clicking on a deceptive pop-up window: by clicking on the window in the mistaken belief that, for innocuous advertisement pop-up is being dismissed. In such cases, the "supplier" may claim that the user "consented" to the download, although the user was in fact unaware of having started an unwanted or malicious software download.

3.3. CLICKJACKING ATTACK

J.B.Patil and et al., [2] propose that, Click jacking or Click jack attack is a Web vulnerability used by an attacker to collect an infected user's clicks. Using a similar technique, keystrokes can also be hijacked. With carefully created combination of fashion sheets, iframes , and content boxes, a client can be driven to accept they are writing in the watch word to their email or bank account, but are instep writing into an imperceptible outline controlled by the attacker.

Click jacking can be done as follows.

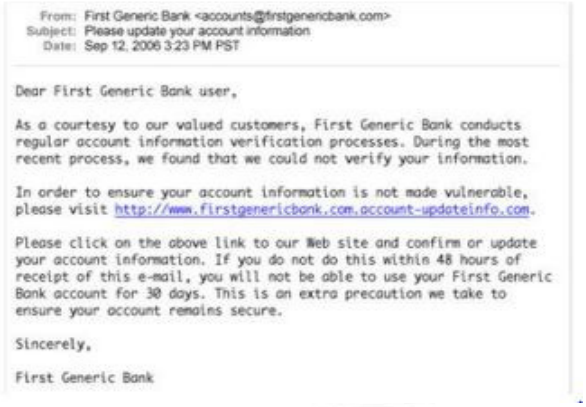
1. A visitor is lured to a vulnerable Web page. Like, "Click to get 1000000 Rs." Or whatever.
2. The vulnerable Web page puts a "Get Rich Now" link with z-index=-1.
3. The vulnerable Web pages puts includes a transparent iframe from the victim domain, say facebook.com and positions it so that "I like it" button is right over the link.

3.4. PHISHING ATTACK

ArindamDasgupta and et al., [5][8][13][15] tells, Phishing is a fraudulent attempt, usually ,made through E-mail, to steal your personal information, appearing to come from legitimate enterprises(e.g. your university, your Internet service provider, your bank). Phishing attack consists of different types. They are malware based phishing, deceptive phishing, hosts file phishing. Using this attack the attackers can able to retrieve the user logins the profiles, mails, secured account etc. Phishing E-mails will continuously tell you to click a interface that takes you to a location where your individual data is asked. The exploit is appeared in figure 2[2] and also the comparision of attack is shown in Table 2.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018


From: First Generic Bank <accounts@firstgenericbank.com>
Subject: Please update your account information
Date: Sep 12, 2006 3:23 PM PST

Dear First Generic Bank user,

As a courtesy to our valued customers, First Generic Bank conducts regular account information verification processes. During the most recent process, we found that we could not verify your information.

In order to ensure your account information is not made vulnerable, please visit http://www.firstgenericbank.com.account-updateinfo.com.

Please click on the above link to our Web site and confirm or update your account information. If you do not do this within 48 hours of receipt of this e-mail, you will not be able to use your First Generic Bank account for 30 days. This is an extra precaution we take to ensure your account remains secure.

Sincerely,

First Generic Bank

Fig 2:Phishing E-mail / Phishing Website Example[2]

3.5. CROSS-SITE SCRIPTING ATTACK

[16] Cross site scripting is an application layer hacking technique. It is also known as XSS attack which is shown in figure 3[1]. In cross-site phishing the attacker infects a legitimate web page with malicious client-side script. When an user visits the webpage the script is downloaded to the browser and executed.

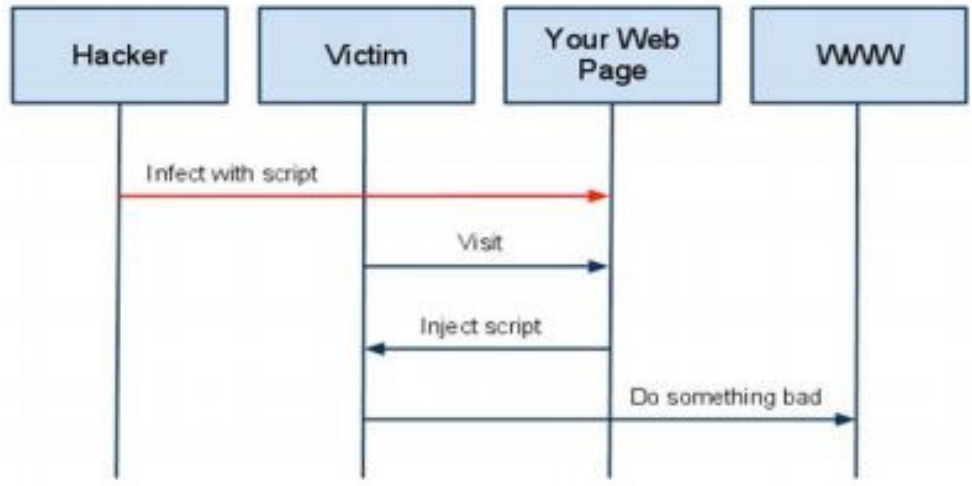


Fig 3: A High Level View of a Typical XSS Attack[1]

[17][28] It consists of two types in XSS attacks. Persistent and non-persistent XSS attack. In Persistent attack the code will be injected and it will be stored in the database. This attack creates more damage then non-persistent attack. In non-persistent exploit the aggressor makes a connect and when the client visits the connect, the created code will executed in the user's browser.

3.6. SQL INJECTION ATTACK

[2][17][5] SQL infusion, more over known as SQLI, is common exploit vector that employments

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

malevolent SQL code for backend database control to get to data that was not aiming to be shown. This data may incorporate any number of things, counting delicate company information, clients records or private client points of interest. It is done through injecting a SQL code in it.

3.7. THIRD-PARTY WEB APP ATTACKS

B.Liang et al.,[7][48][18][2] propose us that, In today's world, the companies and organizations are working under Third Party Applications. Using these applications the attackers can inject their payloads into it and make their attacks possible. For example, if an email contains a third-party apps 1 on 5 can be the injected payload application. In the cloud, or in the case of Web-based apps, this attack would involve a file being shared inbound using a legitimate service like Drop box, Box, Salesforce.com or Google Drive. [26][32][19][47]

3.8. JAVA SCRIPT OBFUSCATION ATTACK

ChaitraliAmrutkar and et al., [19][2][11][14], Obscurity is a method to cover up exploits from inactive location instruments, which utilize marks to co-ordinate against a known noxious string. Confusion causes the appearance of the malevolent string to alter hence side stepping these discovery devices. JavaScript is a energetic client-side Web programming dialect, utilizing this dialect aggressor can create modern exploits by embeddings muddled noxious JavaScript code in a ordinary Web page. The taking after illustration appears the muddling of JavaScript.

The prediction rate of proposed system results are compared against security ontology [41] [25] and tabulated in the Table 1. From the Table 1 it is clear that the proposed system predicts more attacks with the help of the inference process than the existing system.

TABLE 1: COMPARISION OF PREDICTION RATE FOR SOME ATTACKS [25][41]

Web Application Attacks	Proposed System	Existing System
Cross Site Scripting(XSS)	92.3	86.9
SQL Injection(SQLI)	91.05	87.09
Deniel of Services(DOS)	84.56	63.04
Cross Site Request Forgery(CSRF)	84.56	63.04
Content spoofing(CS)	82.57	72.43
Information Leakage(IL)	74.09	70.08
Insufficient Authentication (IAuth)	84.23	66.89
Insufficient Authorization(IAuthr)	83.45	64.23
Brute Force(BF)	84.67	60.56

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

TABLE 2: COMPARISION BETWEEN DIFFERENT TYPES OF PHISHING ATTACKS[54]

Sl.no	FEATURES	CATEGORY	CRITERIA					
			User,s data obtained	Risk	External Dependency	Data retrieval time	Processing Time	Data retrieved size
1	Blacklist[13][22][23][25]		Moderate	Low	Yes	Moderate	Low	Low
2	Lexical [21][22][27][32][35] [37][40][43]	Traditional	Easy	Low	No	Low	Low	Very High
		Advanced	Easy	Low	No	Low	High	Low
3	Host [27][30][43][45][46] [48][49][50]	Unstructured	Easy	Low	No	High	Low	Very High
		Structured	Easy	Low	No	High	Low	Low
4	Content	HTML[28][29][40][51][52] [41][53][43]	Easy	High	No	Depends	Low	High
		JavaScript[28][29][41][43] [51][53]	Easy	High	No	Depends	Low	Moderate
		Visuals[23][40][44][50]	Easy	High	No	Depends	High	High
		Others[21][24][38][49][52]	Easy	High	No	Depends	Low	Low
5	Others	Content-based [16][21][27][40] [46]	Difficult	Low	Yes	High	Low	Low
		Popularity based[8][20][23][28][30] [41][45]	Difficult	Low	Yes	High	Low	Low

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

IV. REPORT

This survey gives an report of existing techniques and approaches of malicious web pages detection. It also consists of features like

blacklist, lexical host, content, and others. We have included all key standards of malicious web pages.

V. KEY PRINCIPLES

[2][13][29], After looking over most of the related work committed to the point of noxious Web pages location, we recognize a set of fundamental

standards that are utilized for development of malicious Web pages discovery framework.

1. There are different exploits on Web pages, Web applications and utilize distinctive set of highlights for modern exploits development against the Net browsers and Web applications.

2. Due to the differing nature of Web exploits, diverse exploits utilize distinctive set of Web page highlights to perform the exploits, against the genuine clients. Hence, it gets to be amazingly vital to extricate novel Web page and URL highlights for location of such sorts of progressed Web exploits.

3. JavaScript is the primary dialect chosen by assailants for exploits development due to its energetic nature and client-side execution [51].

4. JavaScript confusion strategies are utilized by assailants to cover up the exploits script from true blue clients and perform the noxious action against them. In this way, it gets to be crucial to chase for present day energetic highlights of JavaScript for area of novel advanced attacks against client-side.

5. To get the tall degree of precision in malevolent Web pages location, it needs expansive sum of preparing information for the appropriate and precise preparing of the framework. This makes a difference in productively identifying the obscure Web page exploits.

6.In spite of the reality that modern Web page highlights can boost the noxious Web page discovery exactness, legitimate determination and preparing of a machine learning models is too all significance[36].

7. Group machine learning calculations have a execution bottleneck on expansive datasets. Online learning calculations donate way better execution affirmation on huge scale learning and higher location rate with effective utilization of computing resources.

VI. CONCLUSION

In this paper, we have performed an extensive survey of existing techniques and approaches for malicious Web pages detection. We have displayed a brief outline of different shapes of Web pages exploits. We have moreover included all the key standards behind the malicious Web pages discovery procedures and calculations.

VII. FUTURE WORK

The World Wide Web has become an inseparable part of millions of people who use online services. For example: Online banking, Online shopping, Social Networking, E-commerce, and store and manage user sensitive information, etc. In fact, it is a popular tool for any class of user over the internet. Rich Web based applications are available over the World Wide Web to provide such types of services. At the same time, the web has become an important means for people to interact with each other and do business. So we are going to create a

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

software which can detect the malicious web pages and also it alerts when you visit the Web pages which are coded by hackers. Our software is based on a mechanism that distinguishes between malicious and benign Web pages. In addition, we find, characterize and report a number of Web pages missed by Google Secure Browsing and Virus Total, but identified by KAYO. Utilizing this component we are going to distinguish the noxious Web pages which are based on URL, HTML and JavaScript. So we can convey nearly 95% precise result to the clients.

REFERENCES:

[1] Chia-Huan Wu, Jung-Ying Lai, Jain-Shing Wu∗ , Shih-Jen Chen+, and Chung-Huang Yang "Designing a Taxonomy of Web Attacks" Graduate Institute of Information and Computer Education National Kaohsiung Normal University, Taiwan Network and Multimedia Institute, Institute for Information Industry

[2] DharmarajRajaramPatil and J. B. Patil "Survey on Malicious Web Pages Detection Techniques" Research Scholar, Department of Computer Engineering, R. C. Patel Institute of Technology, Shirpur (MS), India

[3] Pratik Patil "A Literature Survey of Phishing Attack Technique", Prof. P.R. Devale M Tech Student, Information Technology, BVUCOE, Pune, India

[4] Mr.Sulabh.S, Mr.SivaShankar.S "Survey Paper for WARNINGBIRD: Detecting Suspicious URLs in Twitter Stream".

[5] Christo Ananth, A.Regina Mary, V.Poornima, M.Mariammal, N.Persis Saro Bell, "Secure system to Anonymous Blacklisting", International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST), Volume 1,Issue 4,July 2015,pp:6-9

[6] https://www.researchgate.net/publication/287359077_Survey_on_Malicious_Web_Pages_Detection_Techniques

[7] Ahmed Toumanari ESSI, NadyaElBachir El Moussaid "Web Application Attacks Detection: A Survey and Classification", National School of Applied Sciences, IbnoZohr University, Agadir, Morocco

[8] https://fossbytes.com/unicode-phishing-attack-impossible-detect-browser/

[9] "Malicious JavaScript", Available from, http://aw-snap.info/articles/js-examples.php/, (2015).

[10] Christo Ananth, Shivamurugan. C.,Ramasubbu. S, "Enhancement of TCP Throughput using enhanced TCP Reno Scheme", International Journal Of Advanced Research Trends In Engineering And Technology (IJARTET), Volume II, Special Issue XXV, April 2015

[11] Large-Scale Automatic Classification of Phishing Pages Colin Whittaker Google Inc. cwhittak@google.com Brian Ryner Google Inc. bryner@google.com MarriaNazif Google Inc. marria@google.com

[12] Geoffrey M. Voelker, Justin Ma, Lawrence K. Saul, Stefan Savage, "Beyond Blacklists: Learning to Detect Malicious Web Sites from Suspicious URLs". Department of Computer Science and Engineering University of California, San Diego

[13] http://searchenterprisedesktop.techtarget.com/definition/driveby-download/, (2015)

[14] D. Canali, M. Cova, C. Kruegel and G. Vigna, "Prophiler: A fast filter for the large-scale detection of malicious Web pages", Proceedings of the 20th International Conference on World Wide Web (WWW), (2011); Hyderabad, India

[15] "What is phishing?", Available from, https://www.phishtank.com/what_is_phishing.php/, (2015).

[16] "XSSAttack(Cross-SiteScriptingAttacks)",Availablefrom,

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

http://www.thegeekstuff.com/2012/02/XSS-attack-examples/, (2015)

[17] "SQL_Injection",Availablefrom,https://technet.microsoft.com/enus/library/ms161953%28v=sql.105%29.aspx/, (2015)

[18] B. Liang, J. Huang, F. Liu, D. Wang, D. Dong and Z. Liang, "Malicious Web Pages Detection Based on Abnormal Visibility Recognition", Proceedings of the International Conference on E-Business and Information System Security, EBISS, (2009); Wuhan

[19] ChaitraliAmrutkar, Young Seuk Kim, and Patrick Traynor, "Detecting Mobile Malicious Webpages in Real Time Senior Member, IEEE

[20] Internet Security Threat Report, Available from http://www.symantec.com/security_response/publications/threatreport.jsp, (2015)

[21] M. Chew, S. Garera, N. Provos, and A. D. Rubin, "A framework for detection and measurement of phishing attacks," in Proceedings of the 2007 ACM workshop on Recurring malcode. ACM, 2007, pp. 1–8

[22] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, "Beyond blacklists: learning to detect malicious web sites from suspicious urls," in Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2009, pp. 1245–1254.

[23] M. Gupta , P. Prakash, M. Kumar, R. R. Kompella "Phishnet: predictive blacklisting to detect phishing attacks," in INFOCOM, 2010 Proceedings IEEE. IEEE, 2010, pp. 1–5

[24] M. Akiyama, B. Sun, T. Yagi, M. Hatada, and T. Mori, "Autoblg: Automatic url blacklist generator using search space expansion and filters," in 2015 IEEE Symposium on Computers and Communication (ISCC). IEEE, 2015, pp. 625–631.

[25] Golnaz Elahi, Eric Yu, and Nicola Zannone, "A Modeling Ontology for Integrating Vulnerabilities into Security Requirements Conceptual Foundations", Lecture Notes in Computer Science, Springer Berlin Heidelberg, pp. 99-114, 2009.

[26] M. Felegyhazi, C. Kreibich, and V. Paxson, "On the potential of proactive domain blacklisting." LEET, vol. 10, pp. 6–6, 2010

[27] "Learning to detect malicious urls," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 2, no. 3, p. 30, 2011

[28] D. Canali, M. Cova, G. Vigna, and C. Kruegel, "Prophiler: a fast filter for the large-scale detection of malicious web pages," in Proceedings of the 20th international conference on World wide web. ACM, 2011, pp. 197–206.

[29] B. Eshete, A. Villafiorita, and K. Weldemariam, "Binspect: Holistic analysis and detection of malicious web pages," in Security and Privacy in Communication Networks. Springer, 2013, pp. 149–166

[30] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, "Identifying suspicious urls: an application of large-scale online learning," in Proceedings of the 26th Annual International Conference on Machine Learning. ACM, 2009, pp. 681–688

[31] T. Finin , P. Kolari, and A. Joshi, "Svms for the blogosphere: Blog identification and

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

splog detection," in AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, 2006, pp. 92–99.

[32] A. Blum, B. Wardman, T. Solorio, and G. Warner, "Lexical feature based phishing url detection using online learning," in Proceedings of the 3rd ACM Workshop on Artificial Intelligence and Security. ACM, 2010, pp. 54–60.

[33] Y.-X. Ding, W. Zhang, Y. Tang, and B. Zhao, "Malicious web page detection based on on-line learning algorithm," in Machine Learning and Cybernetics (ICMLC), 2011 International Conference on, vol. 4. IEEE, 2011, pp. 1914–1919.

[34] A. Reddy, S. Yadav, A. K. K. Reddy, and S. Ranjan, "Detecting algorithmically generated malicious domain names," in Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. ACM, 2010, pp. 48–61

[35] M. Faloutsos , A. Le, A. Markopoulou "Phishdef: Url names say it all," in INFOCOM, 2011 Proceedings IEEE. IEEE, 2011, pp. 191– 195.

[36] H.-K. Pao, Y.-L. Chou, and Y.-J. Lee, "Malicious url detection based on kolmogorov complexity estimation," in Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01. IEEE Computer Society, 2012, pp. 380–387.

[37] T. Engel, S. Marchal, J. Franc¸ois, R. State, "Phishscore: Hacking phishers' minds," in Network and Service Management (CNSM), 2014 10th International Conference on. IEEE, 2014, pp. 46–54.

[38] "Phishstorm: Detecting phishing with streaming analytics," Network and Service Management, IEEE Transactions on, vol. 11, no. 4, pp. 458–471, 2014

[39] W. Chu, B. B. Zhu, F. Xue, X. Guan, and Z. Cai, "Protect sensitive sites from phishing attacks using features extractable from inaccessible phishing urls," in Communications (ICC), 2013 IEEE International Conference on. IEEE, 2013, pp. 1990–1994

[40] H. Choi, B. B. Zhu, and H. Lee, "Detecting malicious web links and identifying their attack types," in Proceedings of the 2nd USENIX conference on Web application development. USENIX Association, 2011, pp. 11–11.

[41] International Journal of Computer Applications (0975 – 8887) Volume 116 – No. 21, April 2015 42 Prediction and Classification of Web Application Attacks using Vulnerability Ontology P. Salini Pondicherry Engineering College Puducherry, India

[42] T. W. Chow, H. Zhang, G. Liu and W. Liu, "Textual and visual content-based anti-phishing: a bayesian approach," IEEE Transactions on Neural Networks, vol. 22, no. 10, pp. 1532–1546, 2011

[43] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song, "Design and evaluation of a real-time url spam filtering service," in Security and Privacy (SP), 2011 IEEE Symposium on. IEEE, 2011, pp. 447–462.

[44] D. K. McGrath and M. Gupta, "Behind phishing: An examination of phisher modi operandi." LEET, vol. 8, p. 4, 2008.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 2, January 2018

[45] M. Kuyama, Y. Kakizaki, and R. Sasaki, "Method for detecting a malicious domain by using whois and dns features," in The Third International Conference on Digital Security and Forensics (DigitalSec2016), 2016, p. 74

[46] D. Chiba, K. Tobe, T. Mori, and S. Goto, "Detecting malicious websites by learning ip address features," in Applications and the Internet (SAINT), 2012 IEEE/IPSJ 12th International Symposium on. IEEE, 2012, pp. 29–39.

[47] T. Holz, C. Gorecki, F. Freiling, and K. Rieck, "Detection and mitigation of fast-flux service networks," in Proceedings of the 15th Annual Network and Distributed System Security Symposium (NDSS08), 2008.

[48] C. Whittaker, B. Ryner, and M. Nazif, "Large-scale automatic classification of phishing pages." in NDSS, vol. 10, 2010

[49] L. Bilge, E. Kirda, C. Kruegel, and M. Balduzzi, "Exposure: Finding malicious domains using passive dns analysis." in NDSS, 2011.

[50] M.-S. Lin, C.-Y. Chiu, Y.-J. Lee, and H.-K. Pao, "Malicious urlfilteringa big data application," in Big Data, 2013 IEEE International Conference on. IEEE, 2013, pp. 589–596.

[51] D. Dong, B. Liang, J. Huang, F. Liu, D. Wang, and Z. Liang, "Malicious web pages detection based on abnormal visibility recognition," in E-Business and Information System Security, 2009. EBISS'09. International Conference on. IEEE, 2009, pp. 1–5.

[52] C. Seifert, I. Welch, and P. Komisarczuk, "Identification of malicious web pages with static heuristics," in Telecommunication Networks and Applications Conference, 2008. ATNAC 2008. Australasian. IEEE, 2008, pp. 91–96.

[53] Y.-T. Hou, Y. Chang, T. Chen, C.-S. Laih, and C.-M. Chen, "Malicious web content detection by machine learning," Expert Systems with Applications, vol. 37, no. 1, pp. 55–60, 2010.

[54] Malicious URL Detection using Machine Learning: A Survey Doyen Sahoo, Chenghao Liu, and Steven C.H. Hoi