International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 4

ISSN: 2454-4248
579 – 584

Client Side Script Phishing Attacks Detection Method using Active Content Popularity Monitoring

P. Malathi
Assistant Professor,
Dhanalakshmi Srinivasan College of Engineering and Technology,
Chennai, Tamil Nadu, India.
malathip2006@gmail.com

Dr. P. Vivekanandan
Professor of Eminence,
Department of Chemical Engineering,
A.C.Tech, Anna University,
Chennai, Tamilnadu, India.
vivek@annauniv.edu

ABSTRACT: The phisher can attack the client side script by means of threatening information which affects the majority of online users in sequence. The malicious users steal a variety of sensitive information from financial organizations in order to run nameless client side script in the phishing attack. In most of the time, the consumer will ignore association script and popup windows which in turn run a set of malicious processes and send the sensitive information to the remote sites. To secure consumers by limiting the client side script, an effective Client Side Script Phishing Attack Detection (CSSPAD) method is proposed to detect the client side script phishing attacks. The proposed methodis based on Active Content Popularity Monitoring (ACPM) and client script classification methods. This method categorizes the client side script according to a mixture of factors like the quantity of information being transferred by the script, the parent information of the script is being accessed. The proposed method computes the active time of the script, amount of data transferred and popularity of the webpage.

KEYWORDS :*Phishing Attacks, Client Scripts, Client Script Categorization, ACPM.*

*****

I. INTRODUCTION

Phishing is an act tied up with computer security. Phisher manipulates search engine results and divert the web users to fake websites to collect their credentials or sensitive information to transfer money from web user account to phisher account. Most of the times, it is very difficult to track phisher as they make use of fake identities and dynamic addresses.Phisher uses several search engine optimization techniques to make their fake websites as a high indexed webpage and rank them as genuine websites.

Web-based attacks are categorized as i) client-side attacks and ii) server-side attacks. Client-side attacks focus on client-side applications which interact with the phished server or run the code which is generated by the phisher. Server-side attacks focus on server-side web application mechanisms and initiate damages using some form of social websites or divert the user to visit phished websites.

In many applications, the scripting languages are used for transferring the information between the users. While transferring the data the cyber-criminals attack the users by means of social networks and other websites.The dynamic nature of the scripting language is to abuse the usage of web applications by using new technologies. This kind of attack is very hard to detect because the code which is used in this technology is very efficient. This leads to develop an effective mechanisms for detecting and mitigating malicious script on the client-side.

II. RELATED WORKS

The information flow control methodology [2] has been proposed for tracking sensitive information flow in the web browser. The information flow will be allowed on the basis of the security class. This method prevents the information leakage and restricts the flow of information from the legitimate channel.

An automated framework was proposed [3] to detect Cross Site Scripting (XSS) attacks at server side based on the notion boundary and policy generation. The server side code generates a notion boundary and identify the features occur during the page generation time to detect XSS attack. This framework suffers from zero false negative and pro of this framework is to detect most common XSS attacks without any modification of client and server entities.

Scriptless timing attack method [4] has been proposed for inhaling user's browsing histories. This method uses the rendering process to protect the users from the leakage of browsing history. The effectiveness of the method is evaluated using six popular browsers, namely,i) Fire Fox ii) Internet Explorer iii) Chrome iv) Safari v) Android built-in browser and vi)Dolphin. This method produces less false positive and false negative.

A new competitive intelligence-based search strategy [5] has been proposed for Search Engine Optimization (SEO) using Graph Structured Search (GSS) techniques. This strategy accommodates user's preference based on a link based ranking evolutionary scheme. It

 ISSN: 2454-4248
579 – 584

combines with Imperialist Competitive Algorithm (ICA) and link based ranking scheme. It handles with large scale data.

The efficient data selection policy [6] was proposed for search engine cache management. This policy handles knapsack problem in storage architecture. It includes three types of storage architectureand they are classified asi) Hard Disk Drives (HDD) ii) Solid Static Disk (SDD) and iii) SSD based hybrid storage architecture for the search engine. This policy is tested by using a greedy algorithm to provide a better result. The efficient data selection policy improves the hit ratio and the information retrieval performance on HDD and SDD.

An Identity Based Secure Distributed Data Storage (IBSDDS) scheme[9] was proposed to handle chosen plain text attack (CPA) and cipher text attack (CCA). The scheme handles with four entities, namely, i) the private key generator, ii) the data owner, iii) the proxy receiver and iv) the receiver. This scheme allows the owner to generate the access permission to extract the file. This scheme will protect the user from the collusion attack.

Network Intrusion Detection and Countermeasure Selection (NICE) [12] is a mechanism which is used to prevent the virtual machine attacks in cloud computing. This mechanism is constructed by using graph-based analytical model and reconfigurable switches. It also investigate the programmability of software switches based solutions and improve the detection rate.

A Self-Organizing Trust Model (SORT)was developed [15] for peer to peer network. It mitigates the service and recommendation based attacks. It prevents the distributed algorithm which enables trust worthiness between peers. This model is based on the previous interaction and recommendations. The results of the model can mitigate the attack on 16 different malicious behavior models. This model produces good peers and able to form trust relationship and isolate malicious peer.

End point elimination technique is based on stream life estimation [16] technique. A set of untrusted connections are detected by this technique and eliminated. Also, it monitors the script. The script includes the following parameters, namely, i) parent of the script ii) lifetime of the script and iii) amount of data transferred. By using the above parameters, the connections are monitored and eliminate the untrusted connection.

An ideal approach has been proposed to identify phishing [17] attacks. This approach extracts source code features, URL features and image features from the website. This approach reduces the features using ant colony optimization algorithm. The naive bayes classifier reduced features are used to categorize the website as legitimate or phished.

A hybrid model is used to detect the phishing website and loss computation[18] using machine learning approach. The loss computation is performed followed by the phishing attack to estimate the expected loss occurred to the stock holders. This model is categorized as i) risk analysis ii) loss corruption and iii) risk mitigation. This model identifies the corrupted Uniform Resource Locator (URL) and reduces the phishing risk for the user.

All the above discussed approaches are difficult to identify the malicious client side script. The proposed ACPM approach handles the popularity of active page and network measure. By using these features the problem of optimization and detecting client script has been done efficiently.

III. THE PROPOSED METHOD

The client side script phishing attack runs in an end user system without the knowledge of the end user and gather the user credentials and send the gathered data to the phished websites. The proposed system mainly focuses on the security of the user to restrict the phishing attack in client side.This detection system has divided into two different stages, namely, i) ACPM and ii) client script classification. The Figure 1 shows the proposed system architecture.

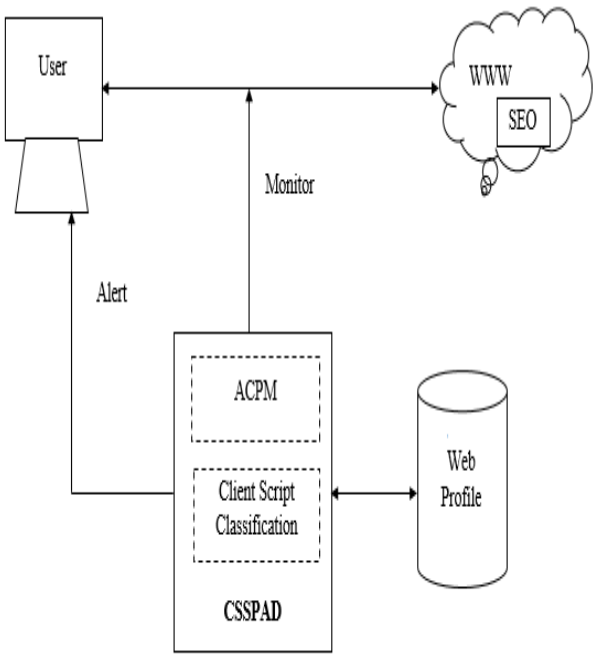

Figure 1: Proposed System Architecture

A. ACPM

The retrieved information is monitored from the client side script. This approach calculates the four different features, namely,i) parent of the script ii) running time of the scriptiii) the amount of data transferred and iv) the popularity of the web page. The computed features will be used in client script classification.

 ISSN: 2454-4248
579 – 584

Algorithm:

Input: Web Log Wl, Client Scrit CS.

Output: Client Script Feature Vector FV.

Step1: start.

Step2: Identify the parent of the script Pp.

Step3: Analyze the runtime of client script Rt.

Step4: Calculate amount of data transferred Dt.

Step 5: Calculate popularity of the web page Pop.

Step 6: Calculates the feature vector Fv={Pp, Rt,

Dt,Pop}.

Step 7: Stop.

In the above algorithm step 2 identifies the parent process. The step 3 analyses the runtime of the client script to specify the startand end time of the client script. The step 4 calculates the data transfer of the client script in a particular time period. The step 5 calculates the web page popularity that links highest ranked website in the google search engine. The popularity of the feature value varies depending on the web browser used by the user. The above four features generate the feature vector.

B. Client Script Classification

The client script classification is based on the features of the monitored client side script and computes three different measures like popularity measure, information access measure, and network access measure.The above three measures are used to compute the script trustworthy measure (STM).

Algorithm:

Input: Web log Wl, Feature Vector FV.

Output: Class C

Step1: start.

Step2: Compute Popularity Measure PM = $\frac{Number\ of\ results\ of\ same\ domain}{Total\ Number\ of\ links\ returned}$

Step3: Compute Information Access Measure IAM = $\frac{size\ of\ page}{Total\ amount\ of\ data\ transferred}$

Step 4: Compute Network access measure NAM= $\frac{Number\ of\ bytes\ transfered}{Total\ number\ of\ objects\ of\ the\ web\ page}$

Step 5: Compute Script Trust Measure STM = $\frac{IAM}{PM} \times \frac{NAM}{PM}$

Step 6: Classify scripts using STM.

If STM less than or equal Genuine Threshold

Assign Genuine.

Else If STM between the Normal Threshold and

Genuine Threshold

Assign Normal.

Else

Assign Phished.

End

Step 7: Stop.

The popularity measure is calculated by the number of links in the parent process and total number of links returned to the parent process. The information access measure is based on the total number of URLs opened and how much amount of data transferred between web pages. The network access measure computes the ratio between the number of bytes transferred per second and the number of objects present in the webpage. The STM computes the ratio between the information access measure, network access measure and popularity measure. Based on the STM value, the script is classified as legitimate or phished. If STM is less than or equal to genuine threshold then the script is genuine else if STM is between the normal threshold and genuine threshold then the script is normal otherwise the script is phished. A normal threshold value is based on the average bandwidth speed of the data transmission divided by two and genuine threshold value holds the normal threshold divided by two. The threshold values are dynamically changed.

IV. RESULT AND DISCUSSION

The proposed approach is implemented in NetBeans IDE 8.1 with Java (jdk1.8) and the OS is Windows 10 with the network bandwidth load is 2.4 GHz and it is applied totest for its efficiency and accuracy.

 ISSN: 2454-4248
579 – 584

The Figure 2 clearly shows the time complexity of the proposed method with the different methods. It is observed that the proposed CSSPAD method has achieved time complexity of 1.2 seconds. The existing methods such as Identity-Based Secure Distributed Data Storage Schemes (IBSDDS), Network Intrusion Detection and Countermeasure Selection (NICE),Self-Organizing Trust Model (SORT) and End Point has achieved 5.3 seconds, 4.2 seconds, 3.8 seconds, and 1.6 seconds time complexity respectively. From this it is clearly shown that the proposed method has achieved less time complexity compared with the existing methods.

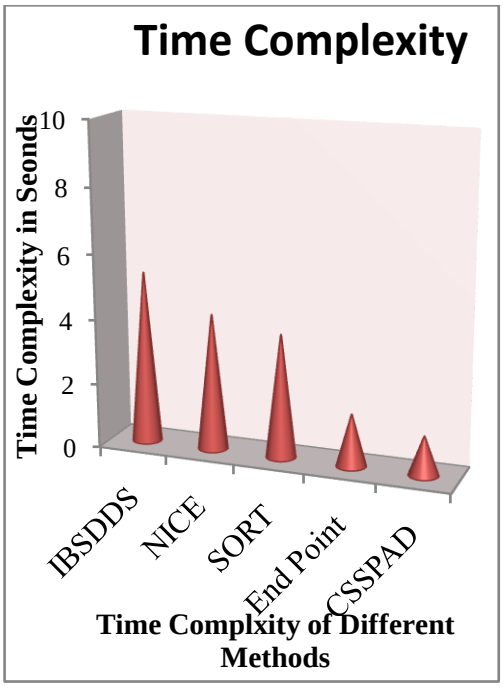

Figure 2: Time Complexity of Different Methods.

The Figure 3 shows the space complexity of the proposed method with the different methods. It is observed that the proposed CSSPAD method has achieved space complexity of 0.8 kb. The existing methods such as IBSDDS, NICE, SORT and End Point has achieved 4.5 kb, 3.5 kb, 2.7 kb, and 1.2 kb space complexity respectively. From this it is clearly shown that the proposed method has achieved less space complexity compared with the existing methods.

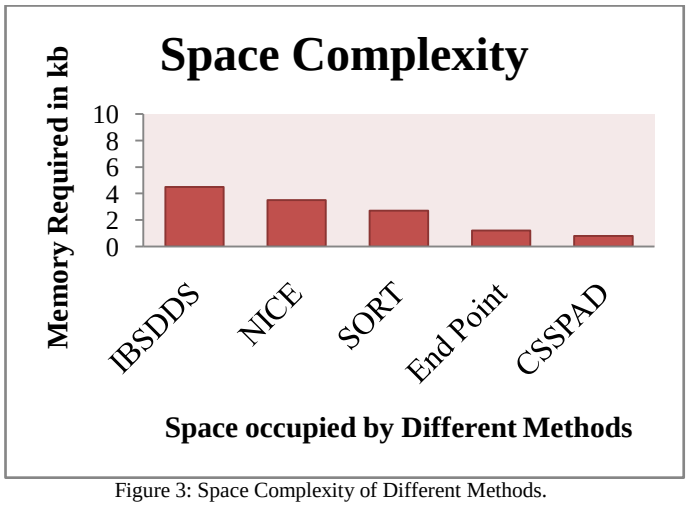

Figure 3: Space Complexity of Different Methods.

The Figure 4 clearly shows the phishing attack detection accuracy of the proposed method with the different methods. It is observed that the proposed CSSPAD method has achieved phishing attack detection accuracy of 97%. The existing approaches such as IBSDDS, NICE, SORT and End Point has achieved 72%, 79%, 82%, and 86% phishing attack detection accuracy respectively. From this it is clearly shown that the proposed method has achieved high phishing attack detection accuracy compared with the existing methods.

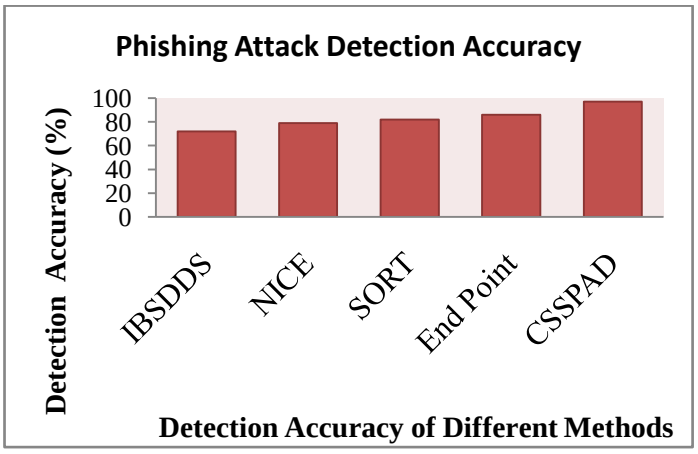

Figure 4: Comparison of phishing attack detection accuracy

The Figure 5 clearly shows the false classification ratio in the proposed method with the different methods. It is observed that the proposed CSSPAD method has achieved false classification ratio 3%. The existing methods such as IBSDDS, NICE, SORT and End Point has achieved 16%, 13%, 8%, and 5% false classification ratio respectively. From this it is concluded that the proposed method has achieved less false classification ratio compared to the existing methods.

 ISSN: 2454-4248
579 – 584

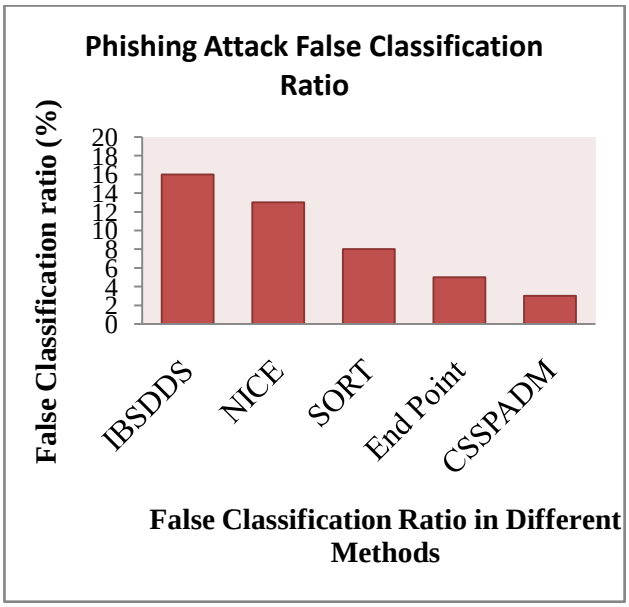

Figure 5: Comparison of false classification ratio

V. CONCLUSION

The proposedCSSPAD method has been used to analyze for the trusted websites and scripts. The proposed method detects the client side script based phishing attack through the search engines.It monitors each and every client request. It has also produced efficient results on time and increases the detection accuracy as well as decrease space complexity. In future it can be utilized to detect and restrict the client side script phishing attacks and to extend the security level of search engines.

REFERENCES

[1] Catherine Bernard,Herve Debarand Salim Benayoune, "Cross-domain vulnerabilities over social networks",Fourth International Conference on Computational Aspects of Social Networks (CASoN), pp. 8-13, 2012.

[2] Bassam Sayed and IssaTraore, "Protection Against Web2.0 Client-sideWebAttacksusingInformationFlowControl", 28th International Conference on Advanced Information Networking and Applications Workshops, pp. 261-268, 2014.

[3] Hossain Shahriar and Mohammad Zulkernine, "S2XS2: A Server Side Approach to Automatically Detect XSS Attacks", Ninth IEEE International Conference on Dependable, Autonomic and Secure Computing 2011.

[4] Bin Liang, Wei You, Liangkun Liu, Wenchang Shi and Mario Heiderich, "Scriptless Timing Attacks on Web Browser Privacy", 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, pp. 112-123, 2014.

[5] Iman Rasekh, "A new Competitive Intelligence-based strategy for Web Page Search",International Conference on Soft Computing and Software Engineering (SCSE 2015),vol. 62, pp. 450-456,2015.

[6] Xinhua Dong, Ruixuan Li, Heng He, XiwuGu, MudarSarem,MeikangQiu and Kequin Li, "An Efficient Data Selection Policy for Search Engine Cache Management", IEEE 17th International Conference on High Performance Computing and Communications (HPCC), pp. 122-127,2015.

[7] Jinguang Han, Willy Susilo, and Yi Mu, "Identity-Based Secure Distributed Data Storage Schemes",IEEE Transactions on Computers, vol. 63, no. 4, pp. 941-953, 2013.

[8] Elhadi M. Shakshuki, Nan Kang, and Tarek R. Sheltami, "Eaack—A Secure Intrusion-Detection System for manets", IEEE Transactions On Industrial Electronics, vol. 60, no. 3, 2013.

[9] Osman Yağan, and Armand M. Makowski, "Modeling the Pairwise Key Predistribution Scheme in the Presence of Unreliable Links"-IEEE Transactions on Information Theory, vol. 59, no. 3, pp. 1740-1760, 2013.

[10] Chun-Jen Chung, Pankaj Khatkar, Tianyi Xing, Jeongkeun Lee, and Dijiang Huang "NICE: Network Intrusion Detection and Countermeasure Selection in Virtual Network Systems", IEEE Transactions on Dependable and Secure Computing, vol. 10, no. 4, pp. 198-211, 2013.

[11] Larry A. Dunning, and Ray Kresman"Privacy Preserving Data Sharing With Anonymous ID Assignment",IEEE Transactions On Forensics And Security, vol. 8, no. 2, pp. 402-413, 2013.

[12] Guilin Wang, Jiangshan Yu, and Qi Xie, "Security Analysis of a Single Sign-On Mechanism for Distributed Computer Networks", IEEE Transactions on Industrial Informatics, vol. 9, no. 1, pp.294-302, 2013.

[13] Ahmet Burak Can, and Bharat Bhargava,"SORT: A Self-organizing Trust Model for Peer-to-Peer Systems", IEEE Transactions on Dependable And Secure Computing, vol. 10, no. 1, pp. 14-27,2013.

[14] P.Malathi and P.Vivekanandan, "Stream Life Estimation Based Endpoint Elimination Technique To Restrict Client Side Scripts For Secure Computing", Asian Journal of Information Technology, vol. 15, no. 19,pp.3846-3851,2016.

[15] R.Priya, "An Ideal Approach for Detection of Phishing Attacks using Naïve Bayes Classifier", International Journal of Computer Trends and Technology (IJCTT), vol. 40, no 2, pp. 84-87,2016.

[16] Baidyanath Biswas and ArunabhaMukhopadhyay, "Phishing Detection and Loss Computation Hybrid Model: A Machine-learning Approach", Information Systems Audit and Control Association (ISACA) Journal, vol. 1, pp. 1-8, 2017.

[17] Sangho Lee, and Jong Kim, "Warningbird: A Near Real-time Detection System for Suspicious urls in Twitter Stream"-IEEE Transactions On Dependable And Secure Computing, vol.10, no. 3, pp. 183-195, 2013.

[18] SedaGurses and Claudia Diaz "Two tales of privacy in online social networks", IEEE Security and Privacy, vol. 11, no. 3, pp. 29-37,2013.

[19] Indrajeet Singh, Michael Butkiewicz, Harsha V. Madhyastha, Srikanth,V. Krishnamurthy, SateeshAddepalli "Twitsper: Tweeting Privately" IEEE Security and Privacy, vol. 11, no. 3, pp. 46-50, 2013.

[20] NgangbamHerojit Singh and, A.Kayalvizhi, M.Tech. "Combining Cryptographic Primitives to Prevent Jamming Attacks in Wireless Networks", IEEE International conference on Information Communication and Embedded Systems (ICICES), 2013.

[21] EnginKirda and Christopher Kruegel, "Protecting Users Against Phishing Attacks with antiphish",Proceedings on 29th Annual International Computer Software and Applications Conference (COMPSAC'05), vol. 1, pp. 517-524, 2005.

 ISSN: 2454-4248
579 – 584

[22] Andronicous, "Classification of Phishing Email Using Random Forest Machine Learning Technique", Journal of Applied Mathematics, 2014.
[23] Layton, "Automatically determining phishing campaigns using the USCAP methodology", IEEE Conference on ecrime Researchers Summit (ecrime), 2010.
[24] Haijun Chang, "Textual and visual content-based anti-phishing: a Bayesian approach", IEEE Transactions on Neural Networks, vol. 22, no. 10, pp. 1532-1546, 2011.
[25] Rafgul Islam, "A multi-tier phishing detection and filtering approach", Journal of Network and computer applications, vol. 36, no. 1, pp.324–335, 2013.

BIOGRAPHIES

Ms. MALATHI received her B.Sc degree in Computer Science from Nehru Memorial College, Puthanampatti, India, in 2003 and received her M.C.A degree in Master ofComputer Applications from Dr.NavalarNedunchezhiyan College of Engineering, Toludur, India, in 2006. She is working as Assistant Professor in Dhanalakshmi Srinivasan College of Engineering and Technology from 2006 andpursuing the part time Ph.D degree in Science and Humanities from the Anna University, Chennai. Her research interest includes internet security, Data Mining.

DR VIVEKANANDAN PERIYASAMY received his Master of Science in Applied Mathematics from Madras University in 1978 and Doctor of Philosophy from Anna University in 1987. Also, he obtained his postgraduate degree in Master of Engineering in Computer Science and Engineering from Anna University in 1995. He is working as Professor of Eminence, Department of ChemicalEngineering in Anna University from 1978. He visited Singapore, Malaysia, German, Bangladesh, Sultanate of Oman and USA for presenting research papers and Chairing sessions. He has published more than 129 research papers in national journals, international journals and conferences. His area of research are Neural Network, Internet Security and Software Reliability.