Biosc.Biotech.Res.Comm. Special Issue Vol 13 No 11 (2020) Pp-79-81

BBRC
Bioscience Biotechnology Research Communications

Malicious URL Detection Using Rule Based Optimization Techniques

N. Jayakanthan[1]* and R.M. Anu Varshini[2]
[1]*Assistant Professor (SRG), Department of Computer Applications, Kumaraguru College of Technology,Coimbatore, India
[2]Department of Computer Applications, Kumaraguru College of Technology, Coimbatore, India

ABSTRACT

The suspicious URL cause harms to users dealing with online tractions. The malicious URLs are harmful to society and induvial user. It attacks the victims computer and steal all their confidential information. Such malicious URL to be analyzed identified and blocked. In this paper we prose rule based model DETECTX to detect the malicious URL. This algorithm analyses various features of the URL for classification. The experimental result shows the efficiency of the proposed system.

KEY WORDS: MALICIOUS URL, RULE BASED CLASSIFICATION, URL DETECTION, MALICIOUS, WEB PAGE.

INTRODUCTION

The taint URL attracts the user to visit the suspicious web site. It collects the details. Like users personal information. It spoils the users system to carry out the attack. These kind of attacks are very serious concern in present cyber security era. Hence the users should conscious about cyber attack. The existing approaches using profiles and machine learning. Even though they have some merits but they have major pitfalls also. Most of the current method identifies the traditional pitfalls. But lot of development in malicious URLs and new features are introduced. In United States a report says 15000 malicious attacks are carried out in every second. It is an imbalance situation. The cyber user to be protected from the malicious attacks. So there is need for research works in this area. The proposed approach is a based on Ant Colony Optimization. The taint features of the URL are analyzed and clustered. Based on the features the URL is classified as genuine or malicious. The experimental results show the efficiency of this approach.

Azeez chaudray et all introduced a system which analyzes the lexical attribute of the web page using linear algebra to identify the cyber crime. Babu Kannan et al produce report which analyzes the impact of various cyber vulnerabilities, He reports it will increase in future. N. Jayakanthan et al analyze the various features in malicious URL and identify concern URL is harmful or not. ID3 decision tree algorithm is used for this purpose. N. Jayakanthan et al uses the Graph based classifier to find the taint URL and categorize it into malicious or genuine. N.Jayakanthan et al carried out a research activity to find various malicious activates occurs in web applications. Both staic and dynamic models are explained.

Kzek ymi developed tool URL finder to identify the URL performing criminal activity by analyzing the properties of the URL. Mousavinejad et al identify various malicious

ARTICLE INFORMATION

*Corresponding Author: jayakanthan.n.mca@kct.ac.in
Received 20th Oct 2020 Accepted after revision 5th Dec 2020
Print ISSN: 0974-6455 Online ISSN: 2321-4007 CODEN: BBRCBA

Thomson Reuters ISI Web of Science Clarivate Analytics USA and Crossref Indexed Journal

NAAS Journal Score 2020 (4.31) SJIF: 2020 (7.728)

Online Contents Available at: http//www.bbrc.in/
Doi: http://dx.doi.org/10.21786/bbrc/13.11/18

activities in network control system. Recursive algorithm is used to detect the attack. Musoodjafran et alnarrate a from work which analyze the various factors to URL. To set the URL is malicious or genuine. This approach narrates the various attribute find the malicious URLs. Sheryas analyzes prose linear model to find malicious URLs which detects the cyber security violations in public gathering media.Ying Da et al analyze the URL pattern to segregate the given URL is malicious or not. The authors mine the URL patterns for the classification. The remaining section of the paper is organized as follows. Chapter 2 Material and Methods, the results are given in the chapter 3. Chapter 4 concludes the research work.

MATERIAL AND METHODS

Malicious URL detection is a imperative methodology for the present scenario. The online users are impacted by the malicious attacks. These website steals the user's personal information. Utilize the user's machine for further attack. Hence these attacks to be detected and users to be prevented from these attacks. Lot of research works are carried out in this area. But these methods are having major drawbacks. They are detecting traditional attacks but attackers are launching new attacks. Hence it is essential to carry out new research work in this area. In this research work we proposed rule based approach to detect malicious URL. It compares the occurrence of the various malicious features and detect the given URL is genuine or malicious. List of features are given the following table 1.

Table 1. Features of Malicious URL

SL.No	Features
1	Host name
2	Taint Special Character
3	Cap symbol (^)
4	Dot count (.)

Table 1 gives the features of malicious URL. The Host name belongs to malicious repository then the URL is declared as malicious. If URL contain a taint special character like @ character then it is malicious. If cap symbol occurs then it is declared as malicious. If dot count > 5 then the URL is suspicious. The algorithm DETECX of the proposed approach is given below:

Algorithm DETECX(URL)
Input : Uniform Resource Locator (URL)
Output : Taint or legal
T_R – Taint Feature set
Status – Variable store the result.
Step 1 : Set Status = Legal
Step 2 : Initialize T_R = NULL
Step 3: If Host name (HS) € Malicious Repository then
Set Status = Taint
$T_R = T_R \cup HS$
Step 4: If URL contain taint special character (Ts)
Set Status = Taint
$T_R = T_R \cup T_S$
Step 5: If URL contain cap symbol (^)
Set Status = Taint
$T_R = T_R \cup ^$
Step 6: If Dot Count (Dc) > 5 then
Set Status = Taint
$T_R = T_R \cup D_c$
If Status = Taint then Display the URL is Taint
Display List of malicious features
Else
Display the URL is legal

The proposed rule based algorithm analyze the features if any malicious activity found the URL is declared as malicious otherwise the URL is declared as genuine.

Figure 1: Performance Analysis

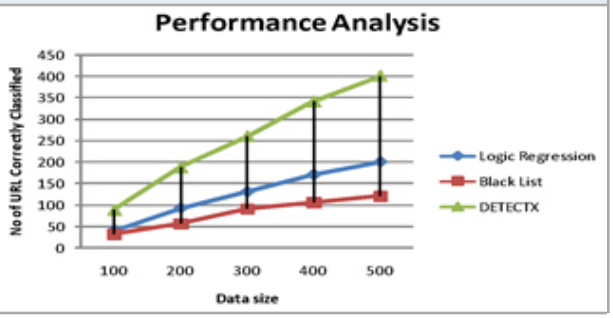

Table 2. Result Analysis

	Logic Regression	Black list	DETECTX
Number of URL Correctly Classified	200	120	400
Success ratio	40	24	80
Failure ratio	60	76	20

RESULTS

The proposed approach is implementing as Java class. It analyzes all features listed in table1. The malicious URL is collected from Phishtank and genuine URLs are collected from DOMZ. In total 500 URLs are collected 250 URLs are legal and 250 URLs are taint. 200 URLs are used for training and 300 URLs are used for testing. To check the efficiency of the proposed approach the testing and training set are kept as disjoint. The proposed approach is compared with other approaches using Logic Regression and Block list based approach the experimental results are given figure 1.

The experimental results shows our approach is efficiently classify the URLs than other approached the classification accuracy is shown in the following table2. The proposed approach is compared with logic regression and black listed based approaches. The success rate of the proposed

system is 80% which is higher than other approaches. It shows the efficiency of the proposed approach.

CONCLUSION

The cyber attacks are the significant threat to the society. It impact the user by accessing the personal information. It essential to prevent such attacks. In this paper a naval framework DETECTX is proposed to detect malicious URL. Rule based clustering algorithm is used for this purpose. Our approach is compared with Logic regression and Blacklist based models. The experimental results show the efficiency of the proposed approach.

REFERENCES

Azeez chaudray(2019) Cyber Crime identification, Ireland Journal of Cyber Law, pp.1021-1057.

Babu Kannan(2017) Impact of Security Vulnerability – An Analysis, Proceeding of 6th International Conclave, Delhi, India.

N.Jayakanthan and A.V.Ramani(2016) A Feature Based Framework to Detect Malicious URLs, International Journal of Control Theory and Applications pp.1327-1340.

N. Jayakanthan and A.V.Ramani(2017) Graph based Classifier to Detect Malicious URL, International Journal of Mechanical and Production Engineering Research and Development, pp. 223-234.

N.Jayakanthan and M.Manikantan, "Malicious Attack Detector(2017), International Journal of Advance Research and Innovative Ideas in Education.Issue .

Kzek ymi, and Nova berk(2017) URL Finder- An Impertive tools, Proceeding of the 7th International Conference on Computer Science, Amristsar, India.

Kumar.S (2014) Malicious Web Attacks Detectio", Journal of Cyber Security.

Mousavinejad,Fuwen Yang,Qing-Long Han and Ljubo Vlaci,A(2018) Novel Cyber Attack Detection Method in Networked Control Systems, IEEE Transactions on Cybernetics.

Musoodjafran,(2018) URL : is Indicator to identify cyber attack, Proceedings of International Conference on Computing Technology, Chennai, India, August.

Sheryas v(2018), Detecting Cyber Attack in Public Sites, International Journal of Technology.

Suji T(2019) Malicious URL Indication International Journal of Engineering.

Tiwan v and Suki S(2016) Phising URLs a Over view.

Ying Da, HuangKai XuJian Pei(2013)

Malicious URL detection by dynamically mining patterns without pre-defined elements, Journal of World Wide Web.

Zanvu and Miam(2015) Phishing Attack Classification, Journal of Computer Science.