e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:04/Issue:11/November-2022 Impact Factor- 6.752 www.irjmets.com

CYBER THREAT DETECTION

Lakshmi Prasanna P.V*1, Prof. Suneetha.A*2

*1Student, Department Of MCA, Sree Vidyanikethan Institute Of Management, Tirupathi,
Andhra Pradesh, India.

*2Professor, Department Of MCA, Sree Vidyanikethan Institute Of Management, Tirupathi,
Andhra Pradesh, India.

DOI : https://www.doi.org/10.56726/IRJMETS31438

ABSTRACT

One of the major challenges in cybersecurity is the provision of an automated and effective cyber-threats detection technique. The proposed technique converts multitude of collected security events to individual event profiles and use a deep learning based detection method for enhanced cyber-threat detection. For this, we developed an AI-SIEM system based on a combination of event profiling for data preprocessing and different artificial neural network methods, including FCNN,CNN and LSTM. The system focuses on discriminating between true positive and false positive alerts, thus helping security analysts to rapidly respond to cyber threats. All experiments in this study are performed by authors using two benchmark datasets collected in the real world. To evaluate the performance comparision with existing methods.

Keywords: Cyber Security, Cyber Attacks, Phishing, Cyber Crime, Network Security, Cyber Security Frameworks, Malware.

I. INTRODUCTION

Check suspicious links by using a mixture of blacklists and deep machine learning by IPQS. Perform a domain phishing check for any URL with the latest IPQS threat data and real-time content analysis.

Our URL using algorithms intelligently match similar indicators from malicious URLs and phishing domains, while ensuring that legitimate URLs are never penalized with false-positives. This approach ensures real-time scanning can identify new threats, even if the malicious URL has never been scanned before such as zero-day malware.

Deploy this URL malware scanner with your SOAR or SIEM applications such as Splunk threat intelligence, Palo Alto, Sumo Logic, Swimlane, IBM QRadar, ThreatConnect, Azure Sentinel and similar security platforms to enrich threat intelligence for malware detection. Lookup domain reputation including parked domain detection, popularity, risk score, malicious links, and similar threat insights.

While most malicious URL checking services rely on Google Safe Browsing, IPQS uses 100% proprietary data and AI algorithms to safely detect phishing links and scan malicious URLs to check URL safety. By performing all URL scanning in-house, IPQS can detect suspicious websites and check website trust with greater accuracy than similar website safety checker services. Quicker detection rates provide support for zero-day phishing links and newly compromised domains used for malware.

The main contributions of our work can be summarized as follows:

Our proposed system aims at converting a large amount of security events to individual event paroles for processing very large scale data. We developed a generalizable security event analysis method by learning normal and threat patterns from a large amount of collected data, considering the frequency of their occurrence. In this study, we specially propose the method to characterize the datasets using the base points in data preprocessing step. This method can reduce the dimensionality space, which is often the main challenge associated with traditional data mining techniques in log analysis. Our event prolong method for applying artificial intelligence techniques, unlike typical sequence-based pattern approaches, provides featured input data to employ various deep-learning techniques.

II. LITERATURE SURVEY

The issue of cyber security is not new but rather has developed more than half-centuary. The arrest of an East German spy in IBM's German by West Germany's police in 1968 was acknowledged as the first case of cyber espionage. In 1983,high school student that was inspired by WarGames movie and called their selves as 414s successfully got inside the unclassified military networks. Ten years ago , “the first real war in cyberspace” attacked Estonia and put the country into “a national security situation”(Hansen and Niessenbaum 2010,p.1168).Now a days, cyber security has been a daily issue that can be found anywhere, from the news that reports spam, scams frauds, and identity theft, to academic articles that discuss cyber warfare, cyber espionage and cyber defense (Dunn-Cavelty, 2010).These significantly bring the issue of cyber security become more important and relevant in recent years. Nevertheless, it remains a complicated task to approach cyber security as merely a simple issue of ‘network security’ or ‘individual security’ as it connects to a larger issue of “the state” ,”society” , “the nation” , and “the economy”(Ibid, p.1155).Our interpretation of cyber security will not be only informed by what we preceive to be the most significant to our daily life, but also by the view of the government and other prominent actors. The interplay of political expression to the variety of cyber threat(Cavelty,2013,p.105) is one of the reasons why is difficult to approach cyber security issue.

III. MODULE DESCRIPTION

1. System Model

In the first module, we develop the System environment model. Website providers JavaScript agent strings to identify and then redirect users to a specific version. We note that not all static features used in existing techniques differ when measured on mobile and desktop webpages. Websites enable access to a user’s personal information and advanced capabilities of mobile devices through weapons. Existing static analysis techniques do not consider these specific functionalities in their feature set. We argue and later demonstrate that accounting further specific functionalities helps identify new threats specific to the web. For example, the presence of a known ‘bank’ fraud number on a website might indicate that the webpage is a phishing webpage imitating the same bank

2. Malicious Pages

We argue that benign webpage writers take effort to provide good user experience, whereas the goal for malicious webpage authors is to trick users into performing unintentional actions with minimal effort. We therefore examine whether a webpage has no script content and measure the number of no script. Intuitively, a benign webpage writer will have more no script in the code tonsure good experience even for a security savvy user.

3. Identifying relevant static features

We extract static features from a webpage and make predictions about its potential maliciousness. We first discuss the feature set used followed by the collection process of the dataset. Structural and lexical properties of a URL have been used to differentiate between malicious and benign WebPages. However, using only URL features for such differentiation leads to a high false positive rate.

Our data gathering process included accumulating labeled benign and malicious specific webpages. First, we describe an experiment that identifies and defines ‘specific webpage’s. We then conduct the data collection process. We use these crawls specifically because they are closet the publication of the related work, making them as close to equivalent as possible.

4. Detect malicious WebPages

We describe the machine learning techniques we considered to tackle the problem of classifying specific webpages as malicious or benign. We then discuss the strengths and weaknesses of each classification technique, and the process for selecting the best model for proposed technique. We build and evaluate our chosen model for accuracy, false positive rate and true positive rate. Finally, we compare technique to existing techniques and empirically demonstrate the significance of technique features. We note that where automated analysis is possible, we use our full datasets; however, as is commonly done in the research community, we use randomly selected subsets of our data when extensive manual analysis and verification is required.

IV. SYSTEM ANALYSIS

Existing Method:

A popular approach in detecting malicious activity on the web is by leveraging distinguishing features between malicious and benign DNS usage. Both passive DNS monitoring and active DNS probing methods have been used to identify malicious domains. While some of these efforts focused solely on detecting fast flux service networks, another can also detect domains implementing phishing and drive-by-downloads. The best-known non-proprietary content-based approach to detect phishing webpages is Cantina.

Disadvantages:

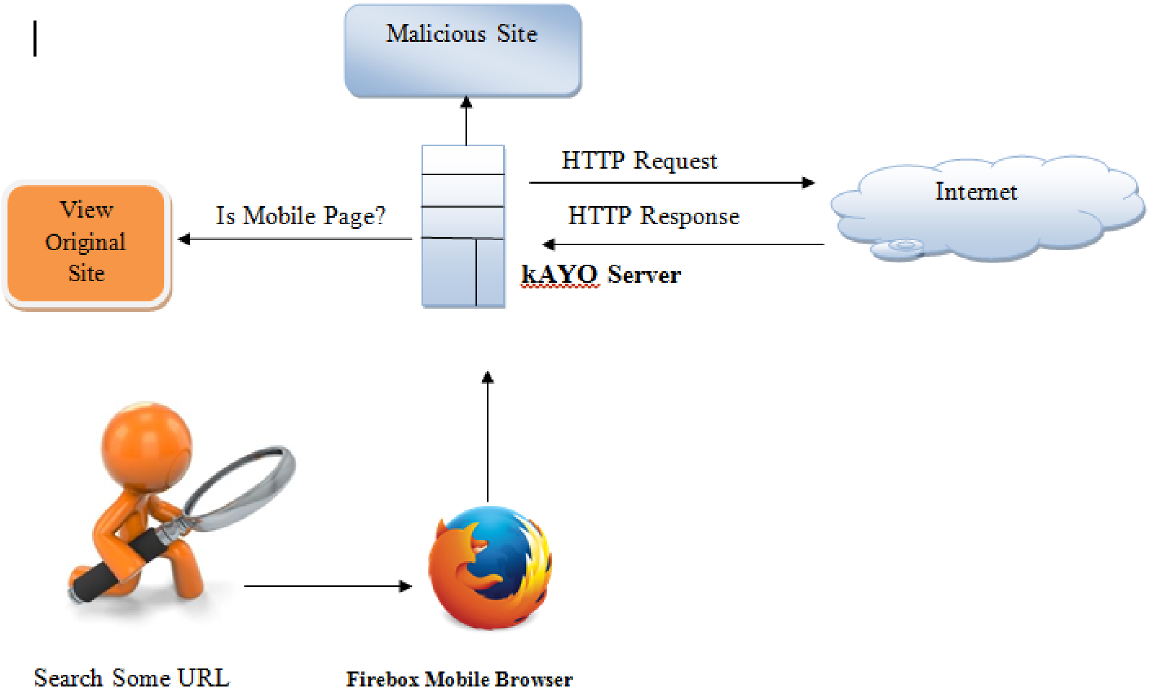
- Existing tools such as Google Safe Browsing are not enabled on the versions of browsers, thereby precluding mobile users.
- DNS based mechanisms do not provide deeper understanding of the specific activity implemented by a webpage or domain.
- Downloading and executing each webpage impacts performance and hinders scalability of dynamic approaches.

Proposed System:

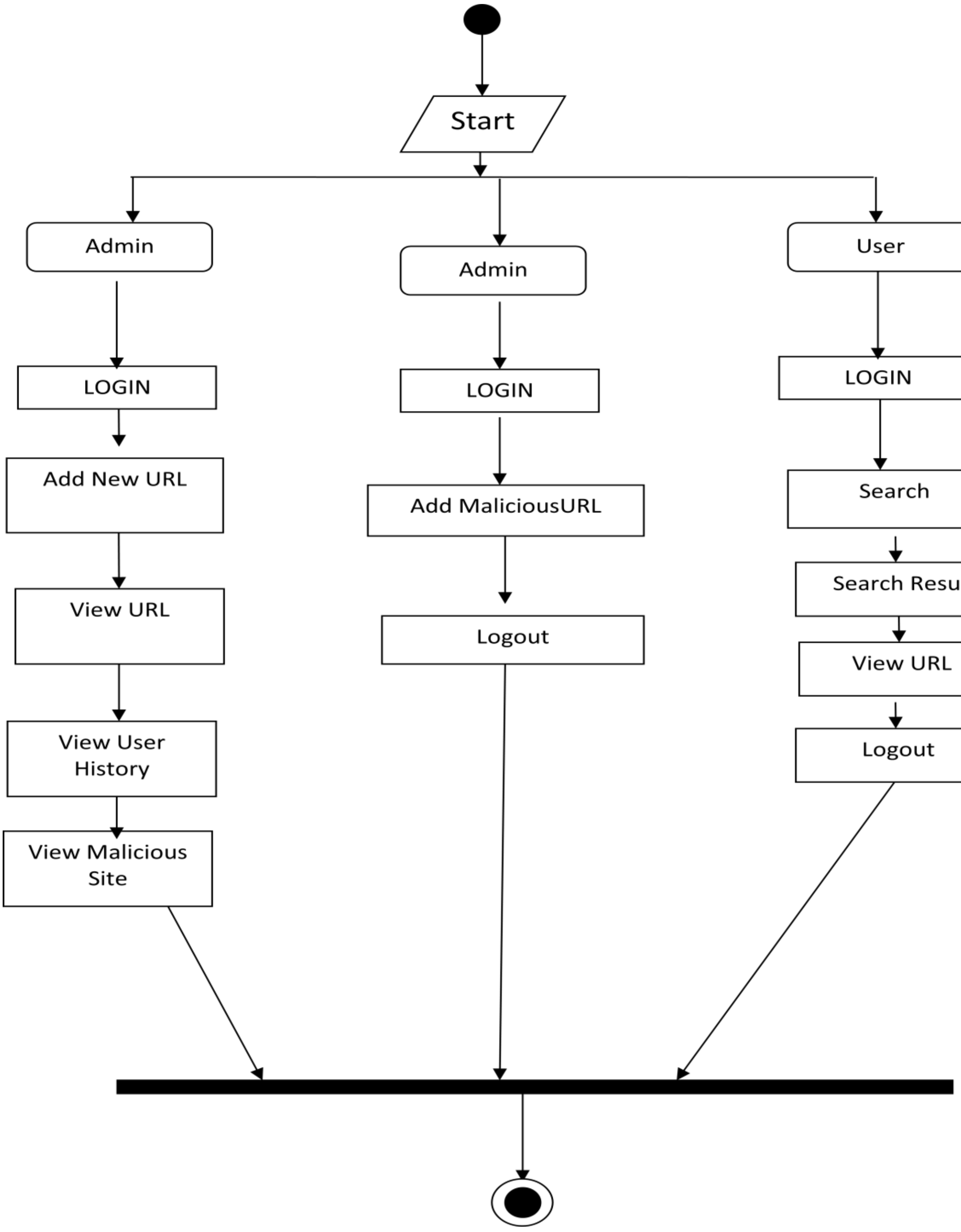
A URL based phishing attack is carried out by sending malicious links, that seems legitimate to the users, and tricking them into clicking on it. In phishing detection, an incoming URL is identified as phishing or not by analysing the different features of the URL and is classified accordingly. Different machine learning algorithms are trained on various datasets of URL features to classify a given URL as phishing or legitimate.

Advantages:

- Detects a number of malicious mobile webpages not precisely detected by existing techniques such as Virus Total and Google Safe Browsing.
- The results of our experiments demonstrate the need for mobile specific techniques for detecting malicious webpages.
- To the best of our technique that detects mobile specific malicious webpages by static analysis.



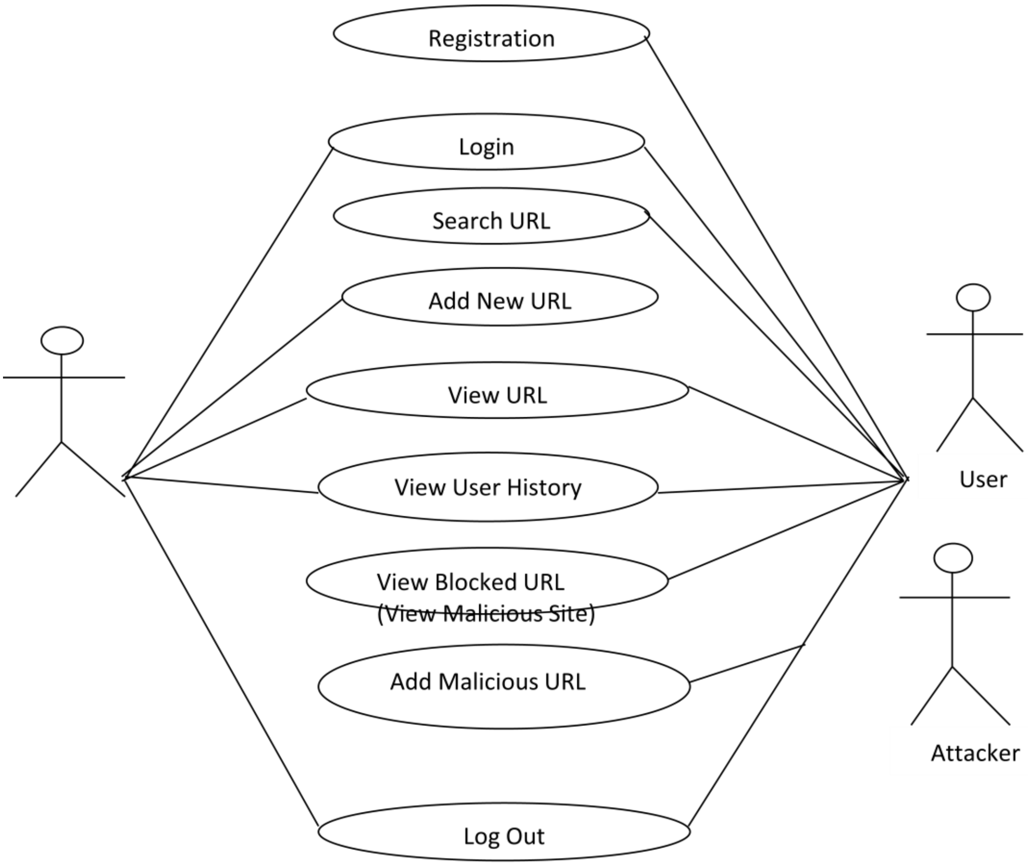
System Architecture



Activity Diagram

USE CASE DIAGRAM

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.



Use case diagram

V. FUTURE SCOPE

Further present a multiparty access control mechanism over the cipher text, which allows the data co-owners to append their access policies to the cipher text. Besides, we provide three policy aggregation strategies including full permit, owner priority and majority permit to solve the problem of privacy conflicts. In the future, we will enhance our scheme by supporting keyword search over the cipher text.

VI. CONCLUSION

Phishing detection is now an area of great interest among the researchers due to its significance in protecting privacy and providing security. There are many methods that perform phishing detection by classification of websites using trained machine learning models. URL based analysis increases the speed of detection. Furthermore, by applying feature selection algorithms and dimensionality reduction techniques, we can reduce the number of features and remove irrelevant data. There are many machine learning algorithms that perform classification with good performance measures. In this paper, we have done a study of the process of phishing detection and the phishing detection schemes in the recent research literature. This will serve as a guide for new researchers to understand the process and to develop more accurate phishing detection systems.

VII. REFERENCES

[1] K. Graves, Ceh: Official certified ethical hacker review guide: Exam 312-50. John Wiley & Sons, 2007.

[2] R. Christopher, “Port scanning techniques and the defense against them,” SANS Institute, 2001.

[3] M. Baykara, R. Das¸, and I. Karado ˘gan, “Bilgi g ¨uvenli ˘gi sistemlerinde kullanilan arac¸larin incelenmesi,” in 1st International Symposium on Digital Forensics and Security (ISDFS13), 2013, pp. 231–239.

[4] RashmiT V. “Predicting the System Failures Using Machine Learning Algorithms”. International Journal of Advanced Scientific Innovation, vol. 1, no. 1, Dec. 2020, doi:10.5281/zenodo.4641686.

[5] S. Robertson, E. V. Siegel, M. Miller, and S. J. Stolfo, "Surveillance detection in high bandwidth environments," in DARPA Information Survivability Conference and Exposition, 2003. Proceedings, vol. 1. IEEE, 2003, pp. 130–138.

[6] K. Ibrahimi and M. Ouaddane, "Management of intrusion detection systems based-kdd99: Analysis with lda and pca," in Wireless Networks and Mobile Communications (WINCOM), 2017 International Conference on. IEEE, 2017, pp. 1–6.

[7] Girish L, Rao SKN (2020) "Quantifying sensitivity and performance degradation of virtual machines using machine learning.",Journal of Computational and Theoretical Nanoscience, Volume 17, Numbers 9-10, September/October 2020, pp.4055-4060(6) https://doi.org/10.1166/jctn.2020.9019

[8] L. Sun, T. Anthony, H. Z. Xia, J. Chen, X. Huang, and Y. Zhang, "Detection and classification of malicious patterns in network traffic using benford's law," in Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2017. IEEE, 2017, pp. 864–872.

[9] S. M. Almansob and S. S. Lomte, "Addressing challenges for intrusion detection system using naive bayes and pca algorithm," in Convergence in Technology (I2CT), 2017 2nd International Conference for. IEEE, 2017, pp. 565–568.