Web guard: Secure Web page access control and URL filtering system

Mrs. C.K. Shruthi[1#] Sneha.J[2#] Selvarasi.A[3#]Vijayalakshmi.M[4#]

[#]Assistant Professor, Department of Information technology, VelTech HiTech, Chennai.

[#]IT Student, Department of Information technology, VelTech HiTech, Chennai.

[1]shruthi@velhightech.com, [2] snehathenu@gmail.com

[3]vh10755_it20@velhightech.com, [4]vh10743_it20@velhightech.com

***Abstract*--**The increasing reliance on digital platforms in office environments has led to a rise in cyber threats and online criminal activities. To address these challenges, this project proposes a comprehensive system for crime detection and malicious URL management within office webpages. The system leverages the power of ensemble machine learning algorithms, including Decision Tree Classifier, Random Forest Classifier, AdaBoost Classifier, KN Neighbors Classifier, SGD Classifier, Extra Tree Classifier, and Gaussian Naive Bayes, to enhance the accuracy and efficiency of the detection process.

Keywords- Url detection, Employee Management

INTRODUCTION

A Digital Forensics Detection og Ongoing Cyber-Attacks ensuring the security and integrity of company operations is paramount. This project is designed to empower organizations to monitor, report, and mitigate internal incidents effectively. Furthermore, the system incorporates a sophisticated Malicious URL Detection and Blocking feature to fortify cybersecurity measures. This web page is mainly used to block the unwanted url to access.

A. GENRAL INTRODUCTION:

In the contemporary digital era, the integration of technology into office management systems has revolutionized the way organizations conduct their daily operations. However, this reliance on digital platforms has also exposed businesses to an escalating threat landscape, particularly in the form of cybercrime. As offices increasingly shift towards web-based applications, the need for robust cybersecurity measures becomes paramount to safeguard sensitive data, ensure business continuity, and protect the integrity of digital infrastructure.

One prevalent and insidious facet of cybercrime is the proliferation of malicious URLs, which pose a significant threat to the security of office webpages. Malicious URLs can be gateways for various cyber threats, including phishing attacks, malware distribution, and unauthorized access to confidential information. Detecting and mitigating such threats in real-time has become a critical aspect of modern office management.

This project endeavors to address these challenges by proposing an innovative approach to cybercrime detection within office management systems, with a specific focus on the identification of malicious URLs. The system employs a diverse set of machine learning algorithms, including Decision Tree Classifier, Random Forest Classifier, AdaBoost Classifier, K.N.Neighbors Classifier, SGD Classifier, Extra Tree Classifier, and Gaussian Naive Bayes. By harnessing the capabilities of these algorithms, the system aims to create a comprehensive and adaptive defense mechanism against cyber threats.

B. PROBLEM BACKGROUND:

DOI: 10.62346/ijcn_q2_v12_no2_24_01

The increasing digitization of office management processes has brought about unprecedented efficiency and convenience, but it has also exposed organizations to a growing array of cybersecurity threats. Cybercriminals exploit vulnerabilities in web-based applications, often using malicious URLs as vectors to compromise the security of office webpages. These threats include phishing attacks, malware distribution, and unauthorized access to sensitive information, jeopardizing the confidentiality, integrity, and availability of crucial data.

Traditional security measures, such as firewalls and antivirus software, are essential but may fall short in providing comprehensive protection against dynamic and evolving cyber threats. Malicious URLs, in particular, can be challenging to detect due to their sophisticated and constantly evolving nature. As a result, there is a pressing need for a proactive and adaptive cybersecurity solution that can identify and mitigate these threats in real-time.

C. PROBLEM STATEMENT

The primary challenge addressed by this project is the detection and management of cybercrime within office webpages, with a specific focus on the identification of malicious URLs. The following aspects constitute the core problem statement: *Malicious URL Detection*: The office management system lacks an efficient mechanism for detecting malicious URLs embedded within webpages. Existing security measures may not effectively identify and mitigate the diverse range of threats associated with these URLs. *Dynamic Nature of Cyber Threats*: Cyber threats, including malicious URLs, are dynamic and continually evolving. Traditional security measures may struggle to keep pace with the rapidly changing tactics employed by cybercriminals, necessitating a more adaptive and proactive approach. *False Positives and Negatives:* Current detection systems may generate false positives, leading to unnecessary alarms and potential disruptions in regular operations. Simultaneously, false negatives may result in undetected threats, leaving the office webpage vulnerable to cyber-attacks. *Limited Integration of Machine Learning:* While machine learning algorithms have demonstrated effectiveness in cybersecurity, their integration into office management systems for real-time threat detection remains limited. The project aims to bridge this gap by leveraging a diverse set of machine learning classifiers. *Lack of Ensemble Learning*: Individual machine learning algorithms may have inherent limitations.

D. RESEARCH OBJECTIVE

The objectives of the project are designed to address the identified challenges and contribute to the development of a robust cybercrime detection and management system within office webpages, with a specific emphasis on the identification of malicious URLs. The key objectives include:

(a) Comprehensive Threat Analysis: Conduct a thorough analysis of potential cyber threats within office webpages, considering various aspects of user behavior, access patterns, and anomalous activities.

(b) Malicious URL Detection: Develop and implement a specialized module for the identification of malicious URLs within the office webpage.

(c) Ensemble Learning Approach: Implement ensemble learning techniques to combine the strengths of multiple machine learning classifiers individual algorithm limitations.

(d) Real-time Monitoring and Alerts: Implement a real-time monitoring system to promptly detect and respond to potential security threats.

(e) User-Friendly Interface: Develop an intuitive and user-friendly web interface for the management and monitoring of the cybercrime detection system.

(f) Adaptability to Dynamic Threats: Design the system to be adaptive to the dynamic nature of cyber threats, ensuring that it can evolve and update its detection mechanisms based on emerging threat patterns.

(g) Reduce False Positives and Negatives: Minimize false positives to avoid unnecessary.

(h) Integration with Existing Security Measures: Ensure seamless integration with existing security measures within the office environment, enhancing the overall cybersecurity posture without causing conflicts or redundancies with pre-existing tools and protocols.

DOI: 10.62346/ijcn_q2_v12_no2_24_01

E. SCOPE OF STUDY:

The scope of the project encompasses various aspects related to cybercrime detection and the management of malicious URLs within the context of office webpages. The scope defines the boundaries and focus areas that the project will cover. Here are the key components within the scope of the study:

1.Office Webpage Environment

2.Malicious URL Detection

3.Machine Learning Algorithms

4.Ensemble Learning Techniques

5.Real-Time Monitoring and Alerts

6.User-Friendly Interface

7.Adaptability and Scalability

8.Integration with Existing Security Measures

9.Evaluation and Validation

II. LITERATURE SURVEY

Paper [1] TITLE: The Proactive and Reactive Digital Forensics Investigation Process: A Systematic Literature Review AUTHOR: Cristina-Edina Domokos, Barna Sera , Karoly Simon , Lajos Kovacs ,Tas- Bela Szakacs. PUBLISHER: IEEE YEAR: 2020 CONTEXT: Recent papers have urged the need for new forensic techniques and tools able to investigate anti-forensics methods, and have promoted automation of live investigation. Such techniques and tools are called proactive forensic approaches, i.e., approaches that can deal with digitally investigating an incident while it occurs. To come up with such an approach, a Systematic Literature Review (SLR) was undertaken to identify and map the processes in digital forensics investigation that exist in literature. According to the review, there is only one process that explicitly supports proactive forensics, the multicomponent process [1]. However, this is a very high-level process and cannot be used to introduce automation and to build a proactive forensics system. As a result of our SLR, a derived functional process that can support the implementation of a proactive forensics system is proposed.

Paper [2] TITLE: Cyber Threat Intelligence from Honeypot Data using Elasticsearch AUTHOR: Zahita Cahyani, Rahmat Nurcahyo, Farizal PUBLISHER: IEEE YEAR: 2020 CONTEXT: —Cyber-attacks are increasing in every aspect of daily life. There are a number of different technologies around to tackle cyber-attacks, such as Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS), firewalls, switches, routers etc., which are active round the clock. These systems generate alerts and prevent cyber-attacks. This is not a straightforward solution however, as IDSs generate a huge volume of alerts that may or may not be accurate: potentially resulting in a large number of false positives. In most cases therefore, these alerts are too many in number to handle. In addition, it is impossible to prevent cyber-attacks simply by using tools. Instead, it requires greater intelligence in order to fully understand an adversary's motive by analysing various types of Indicator of Compromise (IoC). Also, it is important for the IT employees to have enough knowledge to identify true positive attacks and act according to the incident response process. In this paper, we have proposed a new threat intelligence technique which is evaluated by analysing honeypot log data to identify behaviour of attackers to find attack patterns. To achieve this goal, we have deployed a honeypot on an AWS cloud to collect cyber incident log data. The log data is analysed by using elastic search technology namely an ELK (Elasticsearch, Logstash and Kibana) stack.

Paper [3] TITLE: A Real World Study of Personality and Spear phishing Attacks AUTHOR: Trupthi B, Rakshitha Raj R, J B Akshaya, Srilaxmi C P PUBLISHER: IEEE YEAR: 2019 CONTEXT: . Spear-phishing attacks are more sophisticated than regular phishing attacks as they use personal information about their intended victim and present a stronger challenge for detection by both the potential victims as well as email phishing filters. While previous research showed that certain phishing attacks can lure a higher response rate from people with a higher level of the personality trait of Neuroticism, other traits were not explored in this context. The present study included a field-experiment which revealed a number of factors that increase the likelihood of users falling for a phishing attack: the factor that was found to be most

DOI: 10.62346/ijcn_q2_v12_no2_24_01

correlated to the phishing response was users' Conscientiousness personality trait. The study also found gender-based difference in the response, with women more likely to respond to a spear phishing message than men. In addition, this work detected negative correlation between the participants subjective estimate of their own vulnerability to phishing attacks and the likelihood that they will be phished. Put together, the finding suggests that vulnerability to phishing is in part a function of users' personality and that vulnerability is not due to lack of awareness of phishing risks. This implies that real-time response to phishing is hard to predict in advance by the users themselves, and that a targeted approach to defence may increase security effectiveness. Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

III. SYSTEM DESIGN

A. Block Diagram

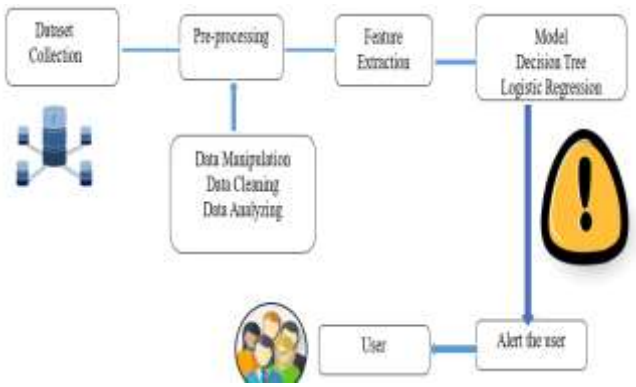

Figure: 3.A Block Diagram

B. Architecture diagram

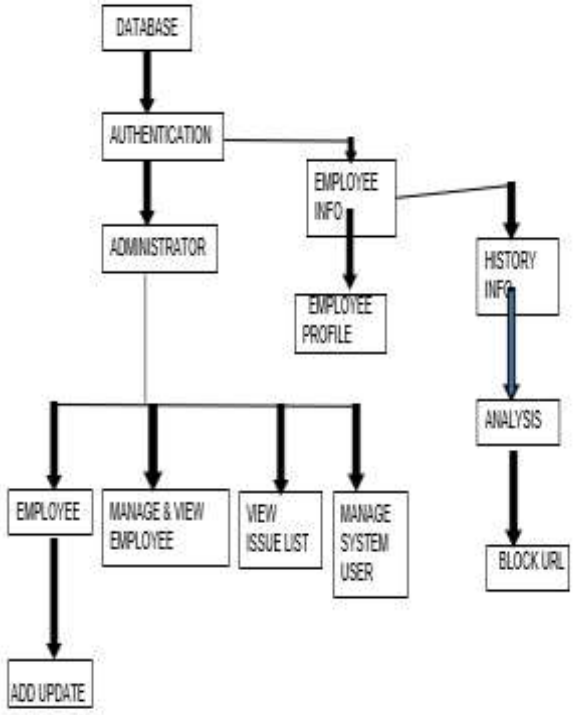

Figure: 3.B Architecture Diagram

IV. RESULT

A. Login Page

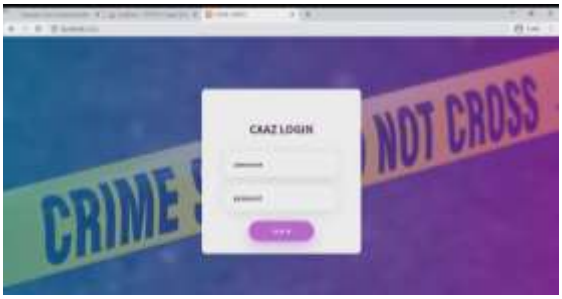

Figure: 4.A Login page

B. Home Page

Figure: 4.B Home Page

DOI: 10.62346/ijcn_q2_v12_no2_24_01

C. Employee Edit

Figure: 4.C Employee Edit

D. View and Manage list of Employee

Figure: 4.D View and Manage list of Employee

E. Database of Employee

Figure: 4.E Database of Employee

F.URL Detection using Algorithm

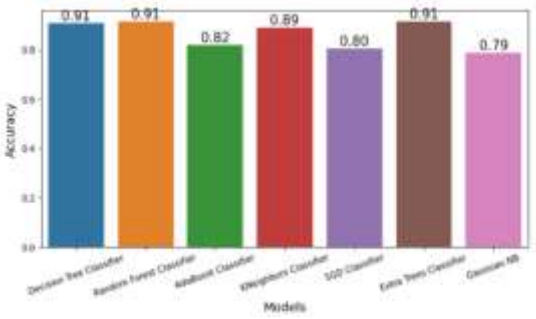

Figure: 4.F URL Detection using Algorithm

V. CONCLUSION

Digital Forensics Detection of ongoing Cyber Attacks presents a comprehensive solution for cybercrime detection in office webpages, focusing on malicious URL identification. Leveraging a diverse set of machine learning algorithms and ensemble learning techniques, the system achieves robust threat analysis. The real-time monitoring and alerting system ensure swift response to potential security threats, reducing false positives and negatives. The user-friendly interface empowers administrators for effective system management. The project's adaptability to dynamic threats and seamless integration with existing security measures provide a resilient and scalable cybersecurity solution. Successful implementation promises enhanced security for digital workspaces, safeguarding sensitive data and ensuring business continuity. The project's outcomes contribute to the evolving field of cybercrime detection and fortify organizations against the ever-changing landscape of cyber threats.

REFERENCES

1. Awojide, Simon,I.M. Omogbhemhe,O.S. Awe, andT.S. Babatope, " Towards the digitalization of Restaurant Business Process for Food Ordering in Nigeria Private University the Design Perspective. A Study of Samuel Adegboyega University Edo State Nigeria," Int.J. Sci. R
2. O.I. Mike and A. Simon, "Towards the Design Perspective," vol. 8, no. 2, pp. 1175 – 1178, 2017.
3. Adithya.R.,A. Singh,S. Pathan, andV. Kanade, "Online Food Ordering System," Int.J. Comput.Appl., vol. 180, no. 6, pp. 22 – 24, 2017.
4. Varsha Chavan, Priya Jadhav, Snehal Korade, Priyanka Teli," enforcing Customizable Online Food Ordering System Using Web Grounded operation", International Journal of Innovative wisdom, Engineering Technology (IJISET) 2015.
5. Patel, Mayurkumar," Online Food Order System for caffs" (2015). Specialized Library. Paper 219.
6. Ali Abdalah Alalwan, "Mobile Food Ordering Apps an Empirical Study Of The Factors Affecting clientE-Satisfaction And Continued Intention To Exercise, " International Journal of Information Management,vol. 50, February 2020.
7. Carsten Hirschberg, Alexander Rajko, Thomas Schumacher, and Martin Wrulich, "The changing request for food delivery, "pp. 1- 6, November 2016.

8. Dávid Földes, Csaba Csiszár, "Model of Information System For Combined Lift- Sourcing Service, " Proc. Smart metropolises Symp Prague(SCSP 2017) – IEEE, May 2017.
9. Heriyanto, Trisno. GoFood Jajal Fitur Baru. 'Review' Makanan dan Layanan Penjual. 2019.
10. Reily, Michael. Tersebar ke 178 Kota, GrabFood Klaim Pengirim Tumbuh 10 Kali Lipat. 2019.
11. Efcom. Adults' Media Use and Attitudes Report 2014. http://bit.ly/1sy0H4A, 2014.
12. F. Enos, S. Benus, R. L. Cautin, M. Graciarena, J. Hirschberg, and E. Shriberg. Personality Factors in Human Deception Detection: Comparing Human to Machine Performance. INTERSPEECH - ISLP, 2006.
13. P. Finn and M. Jakobsson. Designing and Conducting Phishing Experiments. IEEE Technology and Society Magazine, Special Issue on Usability and Security, 2007.
14. T. Halevi, J. Lewis, and N. Memon. Phishing, personality traits and facebook. CoRR, abs/1301.7643, 2013.
15. Y. A. Hamburger and E. Ben-Artzi. The relationship between extraversion and neuroticism and the different uses of the Internet. Computers in Human Behavior, 16(4):441–449, July 2000.
16. C. A. Hill and E. A. O'hara. A Cognitive Theory of Trust. Minnesota Legal Studies Research Paper No. 05-51, 2005.
17. J. Hirsh, S. Kang, and G. Bodenhausen. Personalized persuasion: tailoring persuasive appeals to recipients' personality traits. Psychological Science, 23(6):578 – 81, 2012.
18. C. K. Johann Schrammel and M. Tschelig. Personality Traits, Usage Patterns and Information Disclosure in Online Communities. Proceedings of HCI, September 2009.
19. D. Kahneman and A. Tversky. Prospect Theory: An Analysis of Decision under Risk. Econometrica, March 1979.
20. P. Kumaraguru, A. Acquisti, and L. F. Cranor. Trust modelling for online transactions: A phishing scenario. In Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services, PST '06, pages 11:1–11:9, New York, NY, USA, 2006. ACM.
21. R. R. McCrae and O. P. John. An Introduction to the Five-Factor Model and Its Applications. Journal of Personality, 60(2):175–215, June 1992.
22. M. Mehroof and M. D. Griffiths. Online gaming addiction: the role of sensation seeking, self-control, neuroticism, aggression, state anxiety, and trait anxiety. Cyberpsychol Behavior Social Networks, 13(3):313–316, 2010.
23. T. Minkus and N. Memon. Leveraging Personalization to Facilitate Privacy. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2448026, 2014.
24. Proofpoint. Spear Phishing Statistics: 2012 Findings from Microsoft TechEd, RSA Security Conference Surveys. http://blog.proofpoint.com/2012/07/spear-phishing-statistics- 2012-findings-from-teched-rsa-security-conference- surveys.html, 2012.
25. A. Ramey, J. Klingler, and G. E. Hollibaugh. More than a Feeling: Personality and Congressional Behaviour. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2405140, 2014.
26. M. N. Riaz, M. Akram, and N. Batool. Personality types as predictors of decision-making styles. Journal of Behavioural Sciences, 22(2):99, 2012.
27. B. Roberts, O. Chernyshenko, S. Stark, and L. Goldberg. The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. Personnel Psychology, 58:103–139, 2005.
28. F. Roesner, B. T. Gill, and T. Kohno. Sex, lies, or kittens? investigating the use of snapchat's self-destructing messages. Financial Crypto, 2014.
29. S. Rothmann and E. P. Coetzer. The Big Five Personality Dimensions and Job Performance. Journal of Industrial Psychology, 29(1):68 – 74, 2003.
30. A. Rustichini, C. G. Deyoung, J. Anderson, and S. Burks. Toward the integration of personality theory and decision theory. In University of Minnesota, Mimeo, 2011.