IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

A Data Analytics Strategy to the Underground Economy of Cybercrime

K H Sumanth[1] and Mr. Santhosh S G[2]
PG Student, Department of Master of Computer Applications[1]
Associate Professor, Department of Master of Computer Applications[2]
Jawaharlal Nehru New College of Engineering, Shivamogga, India
sumanthkh10@gmail.com and santhoshsgrao@jnnce.ac.in

Abstract: *People, organisations, and governments have tried to discover strategies to fight against the danger of large-scale cyberattacks (such as ransomware and distributed denial of service (DDoS) assaults) and criminality. In 2017, the WannaCry ransomware was to blame for approximately 45,000 strikes across nearly 100 nations. Governments are under pressure to enhance their cybersecurity spending due to the increasing effect of cybercrime. As part of his 2017 budget, U.S. President Barack Obama suggested allocating more than $19 billion on cybersecurity, a more than 35% increase over 2016. As a result, the cybercriminal underground has changed into a brand-new kind of group that runs underground marketplaces and fosters the growth of cybercriminal conspiracies. Cybercrime networks, in contrast, are lateral, diffuse, fluid, and dynamic. Governments, organisations, and people are typically unaware of the threat posed by the development of highly professional network-based cybercrime business models, such as crime ware-as-a-service (CaaS). Despite the quick rise in cyberthreats, little is known about the basics of the field or the approaches that can help practitioners and researchers in information systems who are interested in cybersecurity. Additionally, little is known about Crime-as-a-Service (CaaS), the illegal business model that supports the shadowy world of cybercrime.*

Keywords: Data analytics, Machine learning, Visualization tools, Cybersecurity, Privacy protection

I. INTRODUCTION

Cybersecurity and law enforcement organisations face major obstacles as a result of the underground economy of cybercrime. Cybercriminals engage in illegal acts such hacking, identity theft, fraud, and the selling of stolen data or hacking tools while operating covertly and with sophistication. It is essential to comprehend the dynamics of this shadow economy in order to create efficient defences.Massive cyberattacks and cybercrimes such as ransomware and distributed denial of service (DDoS) assaults have become more and more of a danger, and people, organisations, and governments have struggled to discover effective means to protect against them. In 2017, the WannaCry ransomware was to blame for roughly 45,000 infections across almost 100 nations. Governments have been under pressure to enhance their cybersecurity spending as a result of the increasing effect of cybercrime. Barack Obama, the president of the United States, suggested allocating more than $19 billion on cybersecurity as part of his fiscal year 2017 budget, a rise of more than 35% from 2016.Highly organised criminal gangs carry out widespread cyberattacks (like WannaCry and Petya), and many recent assaults have been carried out by organised or national-level crime groups.

Cybercrime gangs run this black market in order to expand their already-existing illegal businesses. The cybercrime underground is a brand-new subculture that has emerged as a result of this. This underworld is also in charge of running illicit markets and is essential to the spread of cybercrime conspiracies. Because there are no physical venues, organised cybercrime mostly relies on secret underground groups like Hackforums and Crackingzilla to operate and conduct assaults.These variations can be traced to the mafia's organisational structure. However, the networks that are implicated in cybercrime are horizontal, diffuse, fluid, and dynamic rather than vertical. Governments, businesses, and people occasionally overlook the threat posed by the advent of highly professional network-based cybercrime business models like Crimeware-as-a-Service (CaaS) since cyberspace is inherently a network of networks

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

II. RELATED WORK

Here we have selected few key literatures after exhaustive literature survey and listed as below:

Soska, K., & Christin, N. (2015). Measuring the longitudinal evolution of the online anonymous marketplace ecosystem. In Proceedings of the 24th international conference on World Wide Web (pp. 641-651). ACM.

This study focuses on analyzing the evolution and dynamics of online anonymous marketplaces, shedding light on the underground economy of cybercrime. It employs data analytics techniques to track market activity, vendor behavior, and market longevity.

Holt, T. J., &Kilger, M. (2016). The darknet: A digital copyright infringement hub?. Deviant Behavior, 37(10), 1193-1205.

Examining the role of the darknet in facilitating copyright infringement, this research explores the economic aspects of cybercrime in underground markets. It discusses the challenges of investigating and understanding the underground economy, emphasizing the need for data analytics strategies.

Marziale, L., & Fischer, T. (2018). Using data analytics to identify darknet marketplaces. Digital Investigation, 26, S92-S99.

This study presents a data analytics approach to identify and monitor darknet marketplaces involved in illicit activities. It discusses techniques such as web scraping, machine learning, and network analysis to detect and analyze underground cybercrime markets.

Holt, T. J., & Holt, K. D. (2016). Examining the intersection of cybercrime and the underground economy. International Journal of Cyber Criminology, 10(1), 25-45.

Exploring the intersection of cybercrime and the underground economy, this research investigates the economic motives, transactions, and actors involved in cybercriminal activities. It highlights the importance of data analytics in understanding the complex nature of the underground economy.

Ramakrishnan, N., & Upadhyaya, S. (2019). Big data analytics for cybercrime investigation. In Big Data Analytics for Cyber-Physical Systems (pp. 155-175). Springer.

This book chapter provides insights into leveraging big data analytics for cybercrime investigation. It covers various aspects, including data collection, preprocessing, analysis techniques, and visualization methods, emphasizing the need for advanced analytics strategies to combat cybercrime.

Zhang, J., Zhang, Y., & Cheng, Y. (2018). A deep learning framework for cyber threat intelligence in underground forums. IEEE Transactions on Big Data, 4(2), 196-206.

This research paper presents a deep learning framework for cyber threat intelligence, specifically focusing on extracting and analyzing information from underground forums. It demonstrates the effectiveness of advanced analytics techniques in understanding the underground economy of cybercrime

III. METHODOLOGY

- **Defining goal:** Identifying the conceptual range of the investigation is the first stage. This stage specifically describes the context of the analysis, including the objectives and aims. We investigated the restricted community of the cybercrime underground in order to acquire a thorough grasp of the present CaaS research. The suggested approach therefore aims to "investigate the cybercrime underground economy".
- **Identifying source:** Therefore, we secured a malware database from a top worldwide cybersecurity research business and gathered such information from the community itself.We employed a self-developed crawler that can resolve captchas and anti-crawling scripts to acquire the necessary information because cybercriminals frequently use anti-crawling scripts and alter their IP addresses to hide their communications. We gathered 2,672,091 postings advertising CaaS or crimeware from www.hackforums.net, a popular hacking forum with over 578,000 users and more than 40 million posts, between August 2008 and October 2017. Based on their communication histories, pricing, and questions and answers regarding the transactions, we also gathered 16,172 user profiles of vendors and potential purchasers.
- **Data analytics:** In the cybercrime black market, a wide variety of goods are traded, each with a varying level of danger. In this study, we largely concentrated on components essential to hacking. We sorted the communications into the categories after first filtering out all but the ones that posed serious hazards Our

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

categorization approach examines if a given message fits into one of the following five categories to assess its level of danger: Threat, Product/Service, File Extension, Market, and Exclusion. We made use of a vocabulary with 1,191 terms divided into five categories using information from forums, Wikipedia, anti-virus providers, and a cybersecurity research firm.

- **Implementing an application:** Although organisations highlight the steps, they take to combat cybercrime, their general efficacy has not yet been experimentally shown in real-world situations. The application of the suggested CaaS and crimeware definitions, classification model, and analysis framework are shown in the last stage of our approach. The final application demonstrates how our suggested architecture may provide end users with insights by putting all the data analysis techniques discussed in Section IV into practise.

IV. DESIGN AND IMPLEMENTATION

The two different authentication methods utilised in systems or apps are user login and admin login. For frequent users or customers who enter the system to use its features or services, user login is designed. To authenticate and obtain access to their own accounts, it normally entails inputting a special username or email and a corresponding password. On the other hand, admin login is meant for system administrators or other privileged users who have administrative rights. A specific admin username and password are typically required for admin login, which provides access to additional functionalities and administrative controls like managing user accounts, system configurations, and carrying out administrative tasks to guarantee the system's proper operation and security.

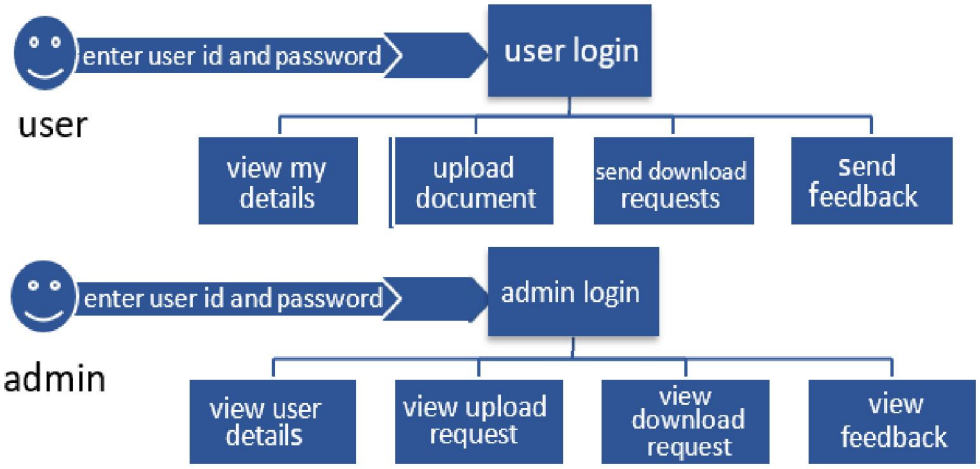

Figure 1: System Design Modules.

Step 1: User upload: Users upload their data through any media to the database or cloud server.
Step 2: Admin process: Admin response for the request and process data.
Step 3: Decisions to provide access: Using algorithm for decision making and provide access to requested customers.
Step 4: Download access files: Tracking of files and download the requested files.

V. RESULTS AND DISCUSSION

After using a machine learning approach for testing, we discover that gradient boosting classifiers have higher accuracy than other approaches. The gradient boosting algorithm is a machine learning technique that combines a number of ineffective learning models to produce a powerful predictive model. Decision trees are frequently used for gradient enhancement. In order to manage the bias variance trade-off, boosting techniques are necessary. Boosting algorithms are thought to be more effective than bagging algorithms since they regulate both bias and variance in a model, as opposed to bagging algorithms, which solely control for high variance

The integrity, confidentiality, or availability of the data may be compromised by a number of possible assaults that might take place during the loading and processing of the data in a module or dataset. Attackers could flood a module or dataset with too many data requests, resulting in a denial of service and blocking the system from being used by

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

authorised users. Techniques like saturating the system with traffic or taking advantage of resource depletion vulnerabilities can be used to achieve this.

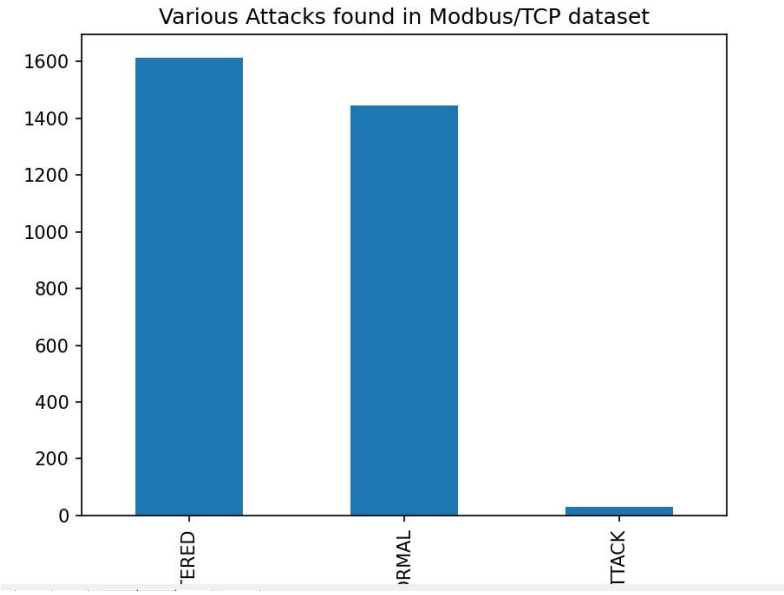

Figure 1: various attack found in dataset

In this case, the CNN algorithm is utilised to foresee a cyberattack. The loaded dataset displays the number of records on a graph. The confusion matrix used by the CNN method, which yields the response attack and the typical predicted class.

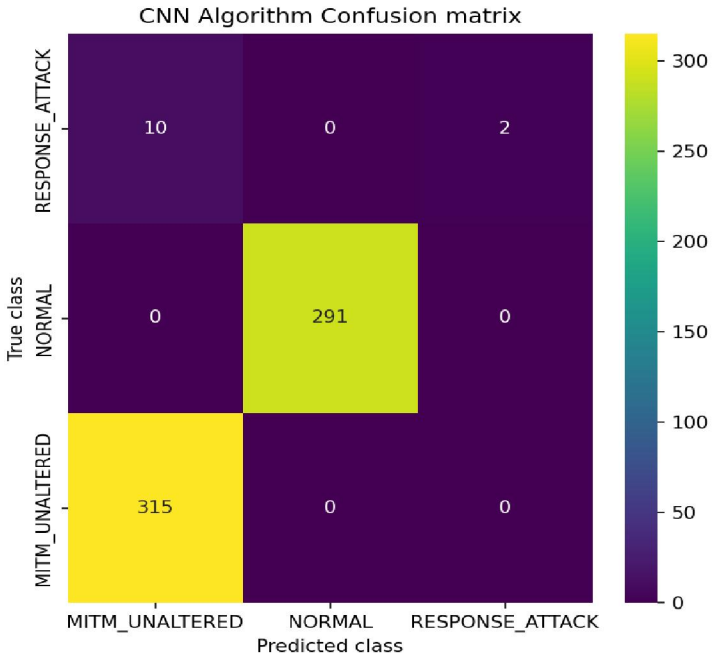

Figure 2: testing a cyberattacks accuracy using the CNN algorithm.

For the purpose of predicting a cyberattack, the LSTM algorithm is applied here.A graph displaying the number of records has been loaded with the dataset. The response attack and typical predicted class are provided by the confusion matrix of the LSTM algorithm.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

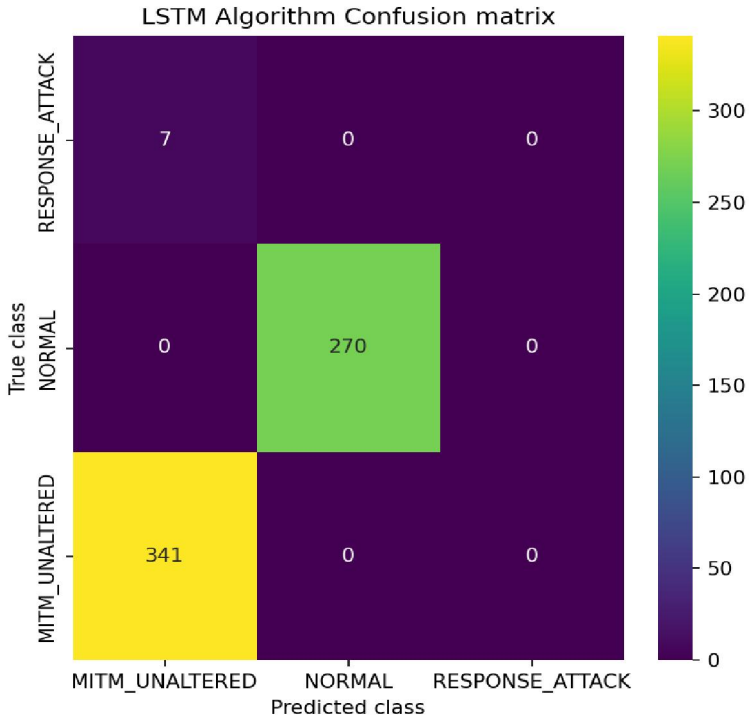

Figure 3: testing a cyberattacks accuracy using the LSTM algorithm.

5.1 User Interface

A user-friendly and secure platform for system access is provided by the user interface of the web page for the data analytics approach on the black market of cybercrime. Typically, the login page has two input fields: one for the username or email and one for the password. The login page makes sure that only approved users may access and use the system, protecting the security and privacy of data**.**

VI. CONCLUSION

This work makes numerous contributions to the DSR literature in a larger IS setting. It adds to the design artefacts, foundations, and processes in this field since it adopts a DSR approach. First, by developing illustrative front-end apps, proved the practical applicability of our design artefacts (the suggested framework and categorization model). As a result, we have suggested two artefacts: a categorization model and a framework for data analysis. Additionally, using sample applications, we evaluated our classification model's accuracy ex ante and its implementation ex post. These four sample applications show the variety of potential practical applications that are open to future researchers and practitioners, which is consistent with the initiation viewpoint of DSR. Our study has largely concentrated on CaaS and crimeware from a RAT viewpoint, as opposed to other studies that provided broader analyses of a wide spectrum of cybercrime. Based on definitions taken from academic and business practise literature, we have also proposed sets of definitions for various CaaS (phishing, brute force attack, DDoS attack, spamming, crypting, and VPN services) and crime ware (drive-by download, botnets, exploits, ransomware, rootkits, Trojans, crypters, and proxies).

REFERENCES

[1]. Anderson, R., Barton, C., Böhme, R., Clayton, R., Van Eeten, M. J., Levi, M., ... & Savage, S. (2012). Measuring the cost of cybercrime. In WEIS (Vol. 12, pp. 1-22).

[2]. Christin, N. (2012). Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. In Proceedings of the 22nd international conference on World Wide Web (pp. 213-224).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

[3]. Eckert, C., &Söllner, M. (2018). Detecting and preventing insider threats: A literature review and integrative model. Computers & Security, 77, 326-350.

[4]. Holt, T. J., &Kilger, M. (2016). The darknet: A digital copyright infringement hub?. Deviant Behavior, 37(10), 1193-1205.

[5]. Marziale, L., & Fischer, T. (2018). Using data analytics to identify darknet marketplaces. Digital Investigation, 26, S92-S99.

[6]. Ramakrishnan, N., & Upadhyaya, S. (2019). Big data analytics for cybercrime investigation. In Big Data Analytics for Cyber-Physical Systems (pp. 155-175). Springer.

[7]. Soska, K., & Christin, N. (2015). Measuring the longitudinal evolution of the online anonymous marketplace ecosystem. In Proceedings of the 24th international conference on World Wide Web (pp. 641-651). ACM.

[8]. Zhang, J., Zhang, Y., & Cheng, Y. (2018). A deep learning framework for cyber threat intelligence in underground forums. IEEE Transactions on Big Data, 4(2), 196-206.