www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 08, August 2024, pp : 550-554

e-ISSN : 2583-1062
Impact Factor : 5.725

SECURING THE BACKBONE: ANALYZING CYBER-ATTACKS ON INFRASTRUCTURE USING THE UNSW-NB15 BENCHMARK

Aderoju Abiola Mayokun[1], Adeosun Olajide Olusegun[2], Akano Victoria Adebimpe[3], Ajewole Fakunle Olaitan[4], Aromavo Pesu Yetunde[5], Atiba Tolulope Janet[6]

[1,2,3,4,5,6]Ladoke Akintola University of Technology , Nigeria.

DOI: https://www.doi.org/10.58257/IJPREMS35754

ABSTRACT

A weak point in the system's security chain could put the entire system in danger and provide attackers with possibilities. Some destructive effects of cyber-attacks could be a compromise on the privacy of people's data, which could render an organization or country's physical systems unworkable, destroy those systems, or provide control of those systems to an outside party. The detrimental impacts of cyber-attacks were also addressed and discussed in detail. This study examines cyber-attacks against infrastructure by utilizing the UNSW-NB15 benchmark network intrusion dataset. A thorough analysis has revealed prevalent instances of common attack types such as Normal, Generic, Exploits, Reconnaissance, Backdoors, Worms, and Shell codes each accompanied by their corresponding event frequencies. Additionally, sub-categories within the top three categories were meticulously examined to ascertain the most widespread attacks. The research findings underscore the prevalent preference among cyber attackers for employing traditional methods, where "normal attacks" comprise over 80% of all documented incidents. Generic attacks and exploit attacks are positioned at 8.6% and 1.7%, respectively. In the domain of Generic attacks, the scrutiny of subcategories highlights Session Initial Protocols (SIP) as notably vulnerable to attacks, eclipsing all others with a ranking of 37%. Meanwhile, within the category of Exploit attacks, subcategory analysis underscores browsers as more prone to attacks than any other subcategories. Moreover, the analysis reveals that office documents are also vulnerable targets for cyber-attacks.

Keywords: Cyber-crime, Cyber-attacks, Generic attacks and Exploit attacks

1. INTRODUCTION

Cyberspace has been designated as a "domain" by the Department of Defense (DoD) and the North Atlantic Treaty Organization (NATO). Glenn Alexander Crowther ranked cyberspace on pace with the air, land, and sea (Crowther, 2017). There has been a surge in cyber crimes worldwide. Cybercrime is a broad term used to describe criminal activities that involve computers and computer networks to commit illegal activities. Individuals of all ages are using various communication channels and devices to commit cyber crimes daily. Cybercrime ranges from identity theft, consumer fraud, and hacking and many more. However, global corporations from various disciplines in the world are coming together to develop new tools to fight against these crimes by developing new theories and strategies to tackle cybercrime.

The basic infrastructure of all cyberattacks is known to be URLs (uniform resource locators) (Rupa et al., 2021). URLs provide a gateway for cybercriminals to gain unauthorized access to computers and computer networks. It is important to be cautious when clicking on links or entering sensitive information on websites, as cybercriminals can use this information to launch attacks such as phishing scams or installing malware on your device.

Detecting cybercrimes can be challenging due to the complex nature of these crimes. The complexity of cybercrimes makes their detection difficult at times. It was mentioned that a large number of cybercrimes might go undetected, making them difficult to track (Vo et al., 2020). However, machine learning models can be used to analyze vast amounts of data and identify patterns that can help in predicting and detecting cybercrimes. By using machine learning algorithms, law enforcement agencies and security professionals can develop more effective strategies and tools to combat cybercrime. A weak point in the system's security chain could put the entire system in danger and provide attackers with possibilities. Some destructive effects of cyber-attacks could be a compromise on the privacy of people's data, which could render an organization or country's physical systems unworkable, destroy those systems, or provide control of those systems to an outside party. In this study cyber-attacks against infrastructure were considered using UNSW-NB 15 benchmark network intrusion dataset. The paper "Detection of Cyber Crime Based on Facial Pattern Enhancement Using Machine Learning and Image Processing Techniques"(Jujjuri et al., 2022) explores the use of machine learning and image processing techniques for the detection of cybercrime, specifically focusing on facial pattern enhancement. The research likely delves into the application of these technologies to identify and prevent cybercrimes, potentially including cyberbullying, phishing, hacking, and other computer-assisted or computer-focused offences. The study may contribute to the development of advanced cybercrime detection methods, leveraging the capabilities of machine learning and image processing.

2. LITERATURE REVIEW

2.1 Types of Cyber-Attacks

1. Denial of Service DOS: The target site is inundated with bogus requests during a denial-of-service assault. All of the responses take up resources on the site because it has to reply to every request. This often leads to the site being completely shut down and prevents the website from serving users (Syed et al., 2020). DDoS is the process by which an attacker hijacks a large number of devices—possibly thousands—and utilizes them to call upon the features of a target system, such as a website, resulting in a website crash due to an overflow of traffic.
2. Drive-by Attacks: Malicious code is embedded into an untrusted website by a hacker. In a drive-by attack, the script infects a user's machine automatically when the website is visited. The term "drive-by" refers to the notion that all it takes to become infected is for a victim to "drive-by" the website and visit it. Nothing on the website needs to be clicked, and no information needs to be entered before getting data from the website. Drive-by-download assaults are a frequent and sneaky type of cyberattack. The study "Detection and Analysis of Drive-by-Download Attacks" addresses the detection of these attacks. These assaults entail when a user visits an unsecured website or a website with malicious script, malware is silently and automatically installed on their device (Aldwairi et al., 2020).
3. Fuzzers: Fuzzing is a testing technique that involves repeatedly running a program with inputs that are generated automatically and may contain syntax or semantic errors. automatically introduce data that is somewhat random into a system or software while keeping an eye out for anomalies like crashes. Fuzzers come in a variety of forms, including intelligent fuzzers that employ algorithms to identify the most likely targets for an attack. Fuzzing is a technique used by threat actors to discover zero-day exploits as well as for vulnerability detection and software product stability improvement. The main concept of fuzzer, as explained by Deng et al. (2023), is to use large language models (LLMs) as an input generation and mutation engine. By employing this type of testing to help discover potential attack vectors, organizations may proactively address these issues before bad actors can take advantage of them (Deng et al., 2023).
4. Generic attacks: Generic Attacks may be carried out without requiring knowledge of the particulars of a cryptographic primitive, such as a hash function. A general attack might also be a brute-force attack. Common approaches in generic attacks include man-in-the-middle attacks and data injection (Duggan, 2006). When we refer to generic attacks, we mean attacks where the internal variations are assumed to be random (Nachef et al., 2010)
5. Exploit attacks: In contrast to other forms of cyberattacks, exploit attacks concentrate on finding weaknesses in a system or piece of software, while other forms of assaults could include social engineering or other techniques to obtain access. 246 vulnerabilities that were exploited between 2021 and 2022 were examined. Zero-day vulnerabilities accounted for sixteen per cent (153) of the vulnerabilities that were initially exploited. (Jacobs et al., 2020) developed a model to estimate the chance that an attacker will exploit software. The goal of the concept is to assist organizations in more effectively allocating and prioritizing their security resources.

2.2 Adverse Effects of Cyber-Attacks

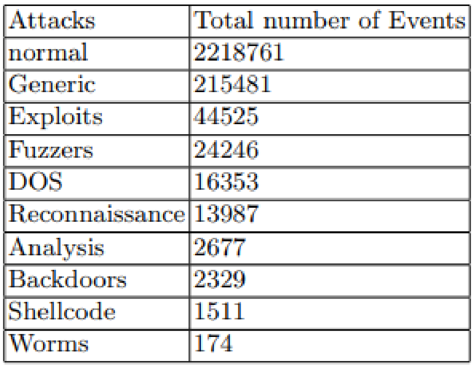
1. Economic setback: Cyber attacks can lead to substantial financial losses for businesses and individuals. This can result from theft of funds, fraud, or the costs associated with recovering from an attack. The May 2021 ransomware attack against Colonial Pipeline led to the shutdown of a major fuel distribution pipeline. Research has examined how cyber security knowledge affects attack detection, and the findings show that having more cyber security knowledge makes it easier to identify hostile events and reduces the likelihood that they would be mistakenly classified as such(Ben-Asher & Gonzalez, 2015).
2. Increased Costs: Cyber-attacks can higher costs for clean-up and recovery. A cyberattack can bankrupt a company with its expenses. Whether it involves paying a ransom, losing data, replacing devices, having to halt operations for a few days, or hiring a security specialist to remove all malware from the system. In the May 2021 ransomware attack against Colonial Pipeline, To get a decryption key and reclaim control of its systems, Colonial Pipeline paid the attackers a ransom of 75 bitcoins or roughly $4.4 million at the time. According to (Furnell et al., 2020), security breaches can come at a high and varied cost; in 2023, the average global cost of a data breach was $4.45 million, up 15% over three years, lost business expenses accounted for over 40% of the average total cost of a data breach.
3. Reputational damage: 95% of the firms surveyed have had several data breaches, according to IBM research from 2023. Numerous companies experienced data loss involving personal clients. One crucial aspect outlined in the taxonomy is Operational Impact, which delineates the tangible repercussions of cyber-attacks on the effective production time within manufacturing facilities. Specifically, Operational Impact measures the loss incurred due to the inability to achieve the anticipated output levels (Espinoza-Zelaya & Moon, 2022).

4. Revenue lost: as stated in the National Bureau of Economic Research report of 2018. The average attacked company loses 1.1 per cent of its market value and sees a 3.2 percentage point decline in its annual sales growth rate following a breach of its consumers' data.

3. ANALYSIS AND DISCUSSION

Using the UNSW-NB 15 benchmark network intrusion dataset. The first five highly ranked features are selected for comparison of the event logs and network traffic. This survey provides an overview of the state of the art in detecting different attacks on different servers using Power BI as a tool for comparison.

Table 1. Cyber-Attacks by Total number of Events.



Attacks	Total number of Events
normal	2218761
Generic	215481
Exploits	44525
Fuzzers	24246
DOS	16353
Reconnaissance	13987
Analysis	2677
Backdoors	2329
Shellcode	1511
Worms	174

The total number of attacks detected from the dataset was over 2 million attacks with 10 attack categories. The Categories of attacks compared are Normal, Generic, Exploits, Reconnaissance, Worms, Analysis, Backdoors, Fuzzers, Shellcode, and Dos. All the listed attacks have multiple subcategories and each attack was carried on different servers.

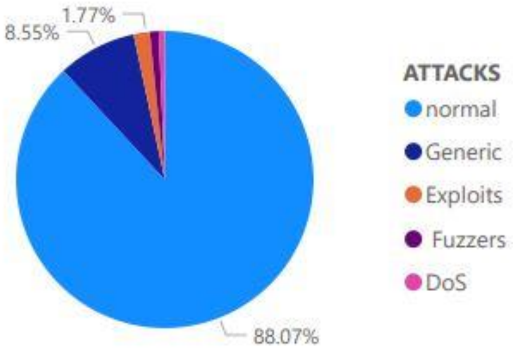


Fig. 1. Sum of the number of events by attacks

Figure 1 shows the top 5 categories of attacks which are Normal, Generic, Exploits, Fuzzers and Dos of which Normal attacks rank the highest compared to the rest of the attacks. This indicates that this is the target of the attacker’s interest. Normal attacks can be in the form of phishing attacks, Malware attacks, Brute force attacks, Drive-by downloads and many more. Most of these attacks exploit the vulnerability without the user’s knowledge, especially through individuals.

Following normal attacks, generic attacks and exploit attacks rank second and third, respectively. The study revealed that generic attacks predominantly target databases, whereas exploit attacks are directed at vulnerabilities within applications or software.

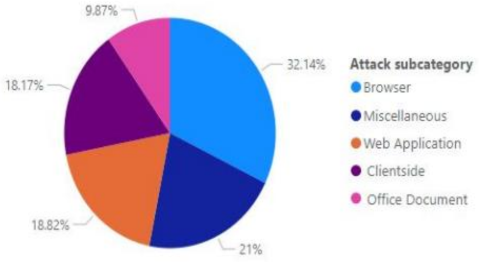


Fig. 2. Sum of the number of events by attacks Sub-category (Exploit)

The frequency of events within Exploit attacks is dominated by the top five attack subcategories: Browser, Miscellaneous, Web Application, Client-side, and Office Documents.

Cyber-attacks targeting web browsers constitute 32.1% of the total incidents. The distinction between cyber attacks targeting browsers and web applications lies in their respective methods of exploitation. In the case of browser attacks, exploits often occur through phishing attempts or browser-based vulnerabilities inherent in the browser software itself. Conversely, attacks on web applications target vulnerabilities within the specific software or services hosted on web servers.

The analysis does not overlook cyber-attacks against office documents, as they account for 9.87% of the total. This underscores the vulnerability of office documents to malicious attacks.

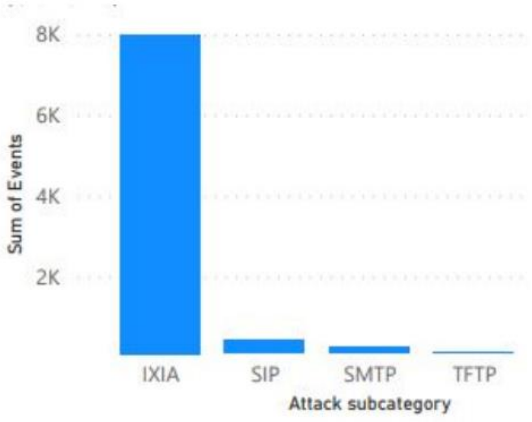


Fig. 3. Sum of the number of events by attacks Sub-category (Generic)

Moreover, figure 2 identifies several subcategories within generic attacks, including IXIA, SIP, HTTP, SMTP, and TFTP. Notably, Session Initial Protocol (SIP) is a communication protocol facilitating various forms of communication, such as audio, video, and messaging, among two or more endpoints (Johnston, 2015).

SMTP, or Simple Mail Transfer Protocol, serves as the primary standard for sending emails (Karim et al., 2019). Despite SMTP servers being susceptible to attacks, the dataset indicates a lower frequency of SMTP incidents compared to IXIA. The top three subcategories under generic attacks encompass critical protocols essential for network communication. Notably, IXIA stands out as the most targeted protocol among them, highlighting its significance in facilitating communication within networks.

4. CONCLUSION

The research findings highlight that the majority of cyber-attackers are inclined towards deploying conventional methods, with "normal attacks" constituting over 80% of all recorded attacks. These results underscore the paramount importance of individual cybersecurity within the cyber realm, as attackers predominantly focus on targeting individual users.

It has been noted that (Alawida et al., 2022) identified 15 distinct categories of cyberattacks, with hacking emerging as the predominant technique. Although this study uses a different dataset, its conclusions are parallel to those of Alawida al et al. (2022). Moreover, the analysis unveils that office documents are also susceptible targets for cyber-attacks.

5. REFERENCES

[1] Alawida, M., Omolara, A. E., Abiodun, O. I., & Al-Rajab, M. (2022). A deeper look into cybersecurity issues in the wake of Covid-19: A survey. Journal of King Saud University-Computer and Information Sciences.

[2] Aldwairi, M., Hasan, M., & Balbahaith, Z. (2020). Detection of drive-by download attacks using machine learning approach. In Cognitive analytics: Concepts, methodologies, tools, and applications (pp. 1598–1611). IGI Global.

[3] Ben-Asher, N., & Gonzalez, C. (2015). Effects of cyber security knowledge on attack detection. Computers in Human Behavior, 48, 51–61.

[4] Crowther, G. A. (2017). The cyber domain. The Cyber Defense Review, 2(3), 63–78.

[5] Deng, Y., Xia, C. S., Peng, H., Yang, C., & Zhang, L. (2023). Large language models are zero-shot fuzzers: Fuzzing deep-learning libraries via large language models. Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis, 423–435.

[6] Duggan, D. P. (2006). Generic attack approaches for industrial control systems. Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

[7] Espinoza-Zelaya, C., & Moon, Y. (2022). Taxonomy of Severity of Cyber-Attacks in Cyber-Manufacturing Systems. ASME International Mechanical Engineering Congress and Exposition, 86649, V02BT02A018.

[8] Furnell, S., Heyburn, H., Whitehead, A., & Shah, J. N. (2020). Understanding the full cost of cyber security breaches. Computer Fraud & Security, 2020(12), 6–12.

[9] Jacobs, J., Romanosky, S., Adjerid, I., & Baker, W. (2020). Improving vulnerability remediation through better exploit prediction. Journal of Cybersecurity, 6(1), tyaa015.

[10] Johnston, A. B. (2015). SIP: understanding the session initiation protocol. Artech House.

[11] Jujjuri, R., Tripathi, A. K., Chandrika, V. S., Majji, S., Prathap, B. R., & Patnala, T. R. (2022). Detection of Cyber Crime Based on Facial Pattern Enhancement Using Machine Learning and Image Processing Techniques. In Using Computational Intelligence for the Dark Web and Illicit Behavior Detection (pp. 150–165). IGI Global.

[12] Karim, A., Azam, S., Shanmugam, B., Kannoorpatti, K., & Alazab, M. (2019). A comprehensive survey for intelligent spam email detection. IEEE Access, 7, 168261–168295.

[13] Nachef, V., Patarin, J., & Treger, J. (2010). Generic attacks on misty schemes. Progress in Cryptology–LATINCRYPT 2010: First International Conference on Cryptology and Information Security in Latin America, Puebla, Mexico, August 8-11, 2010, Proceedings 1, 222–240.

[14] Rupa, C., Srivastava, G., Bhattacharya, S., Reddy, P., & Gadekallu, T. R. (2021). A machine learning driven threat intelligence system for malicious URL detection. Proceedings of the 16th International Conference on Availability, Reliability and Security, 1–7.

[15] Syed, N. F., Baig, Z., Ibrahim, A., & Valli, C. (2020). Denial of service attack detection through machine learning for the IoT. Journal of Information and Telecommunication, 4(4), 482–503.

[16] Vo, T., Sharma, R., Kumar, R., Son, L. H., Pham, B. T., Tien Bui, D., Priyadarshini, I., Sarkar, M., & Le, T. (2020). Crime rate detection using social media of different crime locations and Twitter part-of-speech tagger with Brown clustering. Journal of Intelligent & Fuzzy Systems, 38(4), 4287–4299.