International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Evolving Cybersecurity Strategies: Analyzing Trends in Critical Infrastructure Attacks and Defense Mechanisms

Attila Mate Kovacs

Submitted:10/03/2024 **Revised**: 25/04/2024 **Accepted**: 02/05/2024

Abstract: This paper provides a year-based analysis of the incidents directed toward essential infrastructures and their direction. It identifies tactics that may be utilized to avoid such cases. Through the collection of a plethora of historical and incidence data that has been collected and analyzed, the study brings invaluable information to the table as it seeks to inform future cybersecurity policies and measures. To explain more about the approach of cybercriminals over the past six years, the research employs a deception network coupled with a careful study of attack patterns to establish the evolution in the strategies. Utilizing a data-driven instrument that scores a network subject's geographic, organizational, and behavioral factors, the study presents a complex method of categorizing and describing the creation of network anomalies and the interrelations and dynamics of variations within a network. Preliminary observations point to an elevation of the planned and selective attack approximate, underlining the need for progressive approaches to cybersecurity. Applying the findings in this work makes it important to establish proactive approaches to protecting the existing threat vectors and enhancing the reliability of Communicate infrastructures. This paper not only contributes to the theoretical knowledge base of cybersecurity risks but also offers resourceful guidelines in corporate, governmental, and other organizational spheres to strengthen the protection of cyber systems.

Keywords: *Cybersecurity, Critical Infrastructure, Longitudinal Analysis, Attack Vectors, Mitigation Strategies.*

1. INTRODUCTION

Critical infrastructure is defined as the public and private facilities and systems deemed vital to a society and its economy, which a nation relies heavily on for stability. They are the key industries, encompassing energy, healthcare, transportation, and financial industries. Indeed, in the current societies, dependence on these systems and the delicateness of the support systems affect security, economy, and health, among other issues. This infrastructure has grown more entwined through digital networks and, therefore, is vulnerable to cyber threats, making their security a key national interest.

Several characteristics define the evolution of cyber threats: The former attack vectors continue to evolve and adapt to incremental improvements in protection measures adopted by the organizations. As part of the critical infrastructure, these systems/organizations remain open to cyber threats, as recent cases have demonstrated. For instance, the recent cyber-attack on a leading pipeline company in the United States shut pipelines for days, thereby stopping fuel distribution and demonstrating theoretically how an attack on one entity affects national and economic security in practice. Such incidents nicely depict how much today's infrastructure is one system, and any disturbance can influence numerous jurisdictions and sectors.

Óbuda University Doctoral School on Safety and Security Sciences CISA|CISM|OSCP|CEHv11|CCISO|ISO27kCLA

ORCiD: https://orcid.org/0000-0001-5088-5749

In response to these threats, cybersecurity has transitioned from being an issue of technical importance to organizational importance and may occur at the executive level of an organization. This shift brings to an understanding that cybersecurity dangers are not limited to unauthorized data access but catastrophic outcomes when key assets are functional infrastructure. Consequently, preserving these structures is not only an information technology issue but a necessity to sustain a country's readiness.

This longitudinal study aims to discuss in detail the nature and pattern of the cyber incidents that have targeted critical infrastructure for a more prolonged period. Thus, the research focuses on changes in attack vectors, the identity of potential attackers and defenders, and general tendencies that can be observed by analysing the trends in these kinds of incidents. This method is instrumental not only in coming up with ways of how cyber threats have developed but also in predicting other openings and threats that may arise in the future.

This study extends its development beyond preconceptions and examples, as well as individual and discrete cases, to assess the trends, realism, and depth of cyber threats about critical infrastructures systematically and quantitatively. It will also help in advancing the understanding of the type of threats in these crucial sectors and the threats posed towards them. The result derived from this research will be very beneficial in formulating policies or policies that will enhance better cybersecurity strategies or frameworks. Furthermore, the findings can be

used to improve the probability models and early-warning systems to protect key infrastructures against new episodes of cyber attacks.

Finally, in summary, this research will provide relevant and valuable insights into the cybersecurity of critical infrastructure organizations. It will give policymakers, industries, and IT security specialists an understanding of the threats presents in these important systems and the process that should be adopted to avoid these threats. Thus, the study will provide some insightful findings on securing our societies' critical assets from the constant evolution of threats in cyberspace.

1.1. Objective and Scope

This research's primary purpose is to analyse cybercriminals' external threat landscape while employing a data-driven scoring method for deception data from a six-year dataset obtained from a deception network [15, 8, 4]. It provides a detailed understanding of cyber-attack tactics, techniques, trends, patterns, evolution, and behaviour by incorporating geographic, organizational, or any other related information belonging to the network anomalies [3,2,1]. This research also employs cross-validation techniques to bring out the fact that the authors of cyber crimes are gradually changing their strategies, and this calls for constant updates of the models [14],[15],[19],[20].

As for the coverage of cybercrimes analysed in the work, the set of detected incidents in the dataset includes different levels of intrusion attempts, including the most sophisticated APTs [5, 7]. It emphasizes identifying the source of countries' IP addresses and organizations using specific ASNs to detect network abnormalities [4,3].

This research aims to advance context awareness to inform effective and dynamic cybersecurity planning and strategy development and enrich the qualitative analytical cybersecurity scholarship [6, 16, 17]. By providing deeper insight into the nature of cyber threats, this research is intended to contribute to improving measures of comprehending and counteracting cybercriminals' crafted tactics, which in turn should help create more robust and flexible security systems [22,24].

1.2. Research Question and Hypothesis

Aligned with the objectives and scope of this research, the central inquiry is:

RQ1: As for the core patterns, what were the primary types of cyber-attacks during the specified period, and to what extent have the schemes and strategies developed in terms of complexity and selectiveness?

This question is designed to ask about more specific details of the changing nature of threats with more excellent and more accurate threat targeting. The currently ongoing research should identify the arcs of these attacks to underpin future cybersecurity plans [6,15,17].

Based on this question, the following hypothesis is proposed for testing:

H1: At present, new attack types are even more targeted and sophisticated, with some attacks gaining increased attention and becoming more frequent.

This hypothesis can be explored based on literature chronicling the emergence and advancement of cyber threats [3,7,8] to buttress the idea of steady progress and, thus, the necessity for improvement in cybersecurity. It points out that, like Landman [8], the authors examined different cyber-attacks in detail and identified significant precision changes in their strategies for penetrating the identified systems and networks [13,15]. The hypothesis will be tested through the six-years data extracted from a deception network to understand the rising trends in elaborating the fraudster's modus operandi.

If true, then these research outcomes will further emphasize the necessity for consistent growth and dynamism in cyber defence strategies. They will stress the potential benefits of employing the technology of a deception environment, such as a deception network, decoy systems, and honeypots, to gain intelligence for further protection against specific attacks.

1.4. Significance of the Research

In this study, there are significant implications, as detailed below: To begin with, this piece of work shall offer an extension to the existing literature on cybersecurity and analysis of cybercriminal behaviour by offering a longitudinal and analytical overview of such tendencies during six years [6, 16, 17]. Given this, there is a need for research to be conducted to keep updated with these new developments in the field of cyber defense activities. The research addresses a significant gap in the current literature by identifying major patterns, tendencies, and changes in cyber-attack strategies [3,8,14,15].

Secondly, the practical implications of the study make it a significant undertaking. To address the diverse cybersecurity challenges identified in the six-year deception network dataset, the following considerations are essential: As the world increasingly transitions to digital platforms, introducing new forms of risks, especially with the proliferation of IoT networks, the findings of this study are pivotal. They can guide organizations in adopting security measures that are not only reactive but also predictive and proactive. These measures should enhance organizations' capabilities to detect, respond to, and anticipate cyber threats effectively.

Finally, the necessity of this research can be argued by the fact that defence intelligence agencies turn to deception networks, decoy systems, and honeypots to acquire intelligence against specific systems and networks [13,15]. These expected outcomes foster the utilization of the tools with adequate efficiency in defending organizations, nations, and international systems' cyber threats.

Therefore, this study's theoretical and practical implications suggest that it has significant implications for the academic and practice communities. It facilitates the awareness and analysis of cyber threats and provides a meaningful contribution to enhancing the strategies that target these threats.

2. LITERATURE REVIEW

With the increase in the reliance of critical infrastructure systems on information and communication technologies, the specific critical infrastructures have become vulnerable to the following unique and multiple cybersecurity threats. This systematic literature review aims to analyse the main work of literature in cybersecurity about critical infrastructure and chooses to analyse long-form works to give future insight into the development of cyber threats, measures to protect against them, and implications of such. It also shows the research question this study seeks to answer and the research gaps it seeks to address.

2.1 Existing Research on Cybersecurity Related to Critical Infrastructure

According to the identified research, cybersecurity in critical infrastructure has primarily been based on the sector approach, risks, and opportunities. A fair amount of research has already been published that discusses the cybersecurity of the energy, healthcare, fin-tech, and transportation industries. It elevates the importance of efficient cybersecurity measures for shielding free services from cyber vices that are advanced in modernity and severity [1].

For instance, researchers have paid much attention to the consequences of cyberterrorism targeting the power grid. It is important to note that these breaches not only affect present-day communication but also have long-term economic and security consequences for nations [2]. Similarly, healthcare organizations have realized that data breaches infringe on patients' privacy and disrupt healthcare services, creating significant threats to the public that emanate from the effects of data breaches within the healthcare sector [3].

Learning the nature and behaviour of cybercrimes, the strategies, and the consequences that emanate from them offers a wealth of information to this ever-evolving social vice of cyber warfare, thus laying a solid groundwork, defensive and offensive strategies to counter the menace [2], [3]. The effects of significant information warfare incidents are prodigious, reaching far beyond tangible losses and influencing geopolitical relations and human society as a whole [1]. Hence, studying these mischiefs under a longitudinal lens helps in understanding the operational dynamics, intent, emerging trends, and modus operandi among threat actors. It is vital to improve the protective means of cybersecurity and create strategies and plans aimed at prevention.

To sum up, Figure 1 was generated using data from specified tables representing major cybersecurity incidents for the years 2013-2023 from Tables 1, 2, and 3. As such, the total of 925 events encapsulated in the number provides a succinct summary of the data presented below these tables, presenting a clear picture of the pattern of change of major cybersecurity threats within the last decade. This time perspective becomes a priceless option for quick identification of the vectors of cybersecurity threats.

Table 1: Dataset from previous study[4]

Year	**2013**	**2014**	**2015**	**2016**
DDoS Attacks	5	0	1	1
Malware Attacks	3	2	3	5
Other Forms of Attacks	26	26	30	35
Total	**34**	**28**	**34**	**41**

Table 2: Dataset from Previous Study [4]

Year	**2017**	**2018**	**2019**	**2020**	**2021**
DDoS Attacks	1	2	5	4	8
Malware Attacks	12	25	9	19	21
Other Forms of Attacks	54	82	96	112	92
Total	**67**	**109**	**110**	**135**	**122**

Table 3: New Dataset

Year	**2022**	**2023**
DD0S Attacks	34	19
Malware Attacks	28	40
Other Forms of Attacks	68	56
Total	**130**	**115**

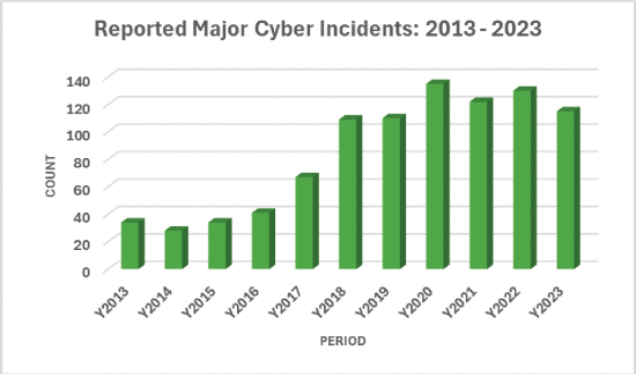

Fig 1: Reported major incidents: 2013 - 2023.

2.1 Related work

Nowadays, there are over 20,000 reported data breaches per year, which cannot be easily counted [1]. It is estimated that a large number of these breaches take place in poorly protected industries such as public services, power utilities, telecommunications networks, transportation systems, oil/gas, and the financial sector. For these reasons, these sectors become vulnerable to targeted military operations during times of conflict with the aim of inducing panic, interrupting means of communication, and ceasing movement. As in the current stage of electronic warfare development, cyberattacks are the ways to cause a detrimental effect or damage to the IAF in pseudonymity and without physical contact. For instance, in healthcare, some firms are synced with government entities, which means their data loss is a crushing blow to the national economy. It makes it difficult for the public to trust service providers, and this could be a blow to the economy, considering that most of these service providers contract to provide services for the government.

One of the reasons why these 'data disasters' seem to be so common is in finding a way of dealing with the insider threat problem that has long been lurking in the world of cyberspace or when cybercriminals gain access to system firewalls and are able to go for tremendous, long periods without being detected by system administrators. However, that is a misconception that some authors, including the one referred to in [2], argue that security is usually managed as multiple layers or as an onion if you will. They note that one requires an insider threat to transfer between these zones seamlessly, and thus, it is quite difficult to detect if they incorporate innovative threats.

Lately, large-scale information leaks have occurred more often. Some of the most famous cases are the Equifax data breach in the period between May and July 2017, which introduced threats to the personal information of 143 million people in the United States of America and 44 million people in the United Kingdom. Earlier, for example, there were massive hacks at NHS, Ashley Madison, TalkTalk, the Office of Personnel Management, several leaks from the CIA/NSA, and other major leaks that show that these are not unique examples of small organizations. These intrusions, which have been primarily aimed at data stealing, raise questions about the possibility of consequences that stem from sabotage or deliberate manipulations meant to lead to fatal outcomes and increased awareness of information warfare in the contemporary system of interstate relations.

Most crucially, in healthcare, cybercrime jeopardizes lives in the most obvious ways. Many NHS trusts in England and Wales, for instance, register associated with their applications as Web applications connected to a variety of back-end storage solutions and available on both PCs and mobile systems, including medical equipment like patient monitors. An attacker who manages to breach all these systems would disrupt crucial services such as accessing Laboratory results, Radiography, and patients' status updates. Furthermore, the stock of medical devices is also at risk, and their normal functioning can be affected because of the disruption of the networks they employ for communication due to DDoS attacks. One of the main concerns arising from the use of these products is the high traffic of data which flows among these devices, and which can be attacked and raped. " Based on the Verizon VCDB dataset, more than 1200 attacks were specifically plotted to healthcare infrastructure, and the rate is escalating [4]. The WannaCry attack on the NHS was referred to by the National Audit Office as the largest in this particular bracket affecting a healthcare body in the history of the world; unfortunately, the NHS was found to have had a repeat of a known mistake that was not followed strictly, and as a result, the systems were still running on Windows XP even when support for this was cut [5].

In relation to critical infrastructure systems, a crucial concern is that there often needs to be a gap between policy as formulated at the strategic level and the ground realities or practice levels, especially in the NHS, as noted in [5]. Analysing industrial research reports, Thursday pointed out that internal carelessness is responsible for 51% of cyber security threats [6]. Additionally, estimations indicate that possible cyber attacks within the healthcare industry could lead to damage of up to $300bn in the future [7]. When safety has been compromised, the most frequently cited reason is negligence. Hence, the following questions arise as to how these risks can best be prevented. In its report, the National Audit Office highlighted that, as it was only by good fortune, WannaCry ransomware's effect was kept at bay when a cybersecurity analyst accidentally stumbled upon a "phone home" mechanism [5].

This proposed systematic review is developed to focus on the existing research gap where there is little combined research done on the impacts of insider threats on the security of infrastructures, especially healthcare facilities. Systematic reviews are known to provide a structured and bias-free approach to addressing a particular set of questions following the established protocols. They, hence, are an effective tool for demarcating the state of the art in any subject area that is often dogmatic and defective with research. This review seeks to give an independent assessment as compared to a typical literature review in computer science by presenting a detailed critical review of the literature, discussing the current technologies utilized for addressing insider threats, and evaluating the trends together with the effectiveness of the methods used in the insider threats countermeasures.

2.2 Longitudinal Studies in the Cybersecurity Domain

Such investments offer longitudinal exposure to the shifts and patterns of tactics, the identities of cyber attackers, and the efficacy of security safeguards across time. Another paper was a longitudinal study that monitored the number of cyber attacks over ten years to identify patterns in attack approaches and motivations of the hackers. It pointed to the change in a calendar year from 'script kiddies' breaking into systems just for fun and challenging themselves to professional criminals motivated by espionage and moneymaking purposes [4].

One of the key objectives of a longitudinal approach is analysing the evolution of detected malware. Several studies have presented case studies on how malware has found ways to work around security measures, proving that threats continue to innovate to fit into new openings as they are discovered [5]. These studies highlight the ever-changing threat landscape and the current state of the world as a battlefield where cybersecurity defenders and malicious actors are constantly trying to outdo one another.

2.3 Gaps in Current Research

However, some substantial gaps are noticeable and relate to the overall investigation of the various sectors as affected by cyber incidents while focusing on the impact of the incidents across those sectors collectively. Most of the current research could be more extensive in terms of the scope of industries addressed or the range of cybersecurity threats examined, which often focuses on ransomware or phishing risks, among others. Most research done in this area needs to be more cohesive, and no one approaches the problem through a system-level lens that appreciates the interdependencies between the different infrastructure sectors and the resulting knock-on effects of cyber-attacks [6].

However, most of these studies are inclined towards the technical consequences of cybersecurity. At the same time, there needs to be more research that initially derives such a technical knowledge base but then intertwines it with policy issues and socio-economic concerns emanating from cyber events. Such a gap is important because, contrary to the security of critical infrastructure, it is about technical processes; it is also a management of policies and governance, which need a comprehensive strategy [7].

There is also a lack of more preventive and incipient studies that can help in future strategic designs on cybersecurity. In fact, even longitudinal studies, which can otherwise be a rich source of data for the problem at hand, may fail to take advantage of sophisticated analytical tools when forecasting future trends or the efficacy of the implemented measures in the realm of cybersecurity. Such predictive studies are quite helpful in the effective management of future cybersecurity threats and in making necessary policies in advance [8].

2.4 Contribution of This Study

In this research, these deficits will be filled by undertaking a Longitudinal Examination of Cybersecurity in Sectors of Critical Infrastructure Systems. To be more precise, rather than focusing only on specific sectors and justifiable cyber risks and consequences of cyber incidents, the study will use data from many sectors and analyse the interconnections and possible cascading effects, thus providing a broader view of the connected infrastructures' susceptibilities and threats.

Moreover, this work will couple general technical findings with policy relevance, stressing the requirement to employ operational policies that recognize and assimilate the notion of 'cybersecurity' and its underlying societal consequences. With this approach, the study will provide policy and governance insights for more effective CYBERSECURITY management addressing threats to critical infrastructures.

Moreover, due to the fact that advanced statistical and machine learning methods will be used to analyse big data in relation to cyber security threats, this research study is believed to offer a set of valuable predictive factors that will be able to influence the further development of the comprehensive strategies and pertinent policies in the sphere of cyber security. Such predictions will be essential in a way that it can predict certain areas that may be prone to cyber threats and ways of developing well-fitted defence mechanisms against new forms of threats.

Consequently, this paper ascertains that despite research having offered abundant knowledge on cybersecurity in critical infrastructure, there still needs to be research gaps. To this end, this study presents the following research questions: To what extent does the extant literature effectively capture cybersecurity risk in CI and predict possible future vulnerabilities? To what extent does the

current cybersecurity policy framework sufficiently address the threats that may affect CI in the future?

3. METHODOLOGY

This study adapts a research strategy that best responds to research question one (RQ1)while testing hypothesis one (H1) through the analysis of critical trends and patterns of cyber-attacks for the past six years. The aim was to determine the changes in the complexity and direction of these cyber offenses; a dataset composed of articles that identified cybercriminal activities provided a long-term examination of the overview and plan.

3.1 Data Collection

The data gathered was obtained from the honeypot for the past six years, which established services and systems that cyber attackers often target. From its implementation in October 2022 to its culmination in September 2024, this system logged every form of cyber-attack – including attempts, breach attempts, and actual breaches. This approach offered a rather broad snapshot of data concerning spammers and their operations, including specific strategies and methodologies.

The honeypot applied here was designed to emulate vulnerable systems and networks, draw cyberspace threats toward them, and gather comprehensive data on various attack styles. This decoy network was essential for realizing practical strategies for attackers, which helped me learn more about the escalating forms of cyber threats.

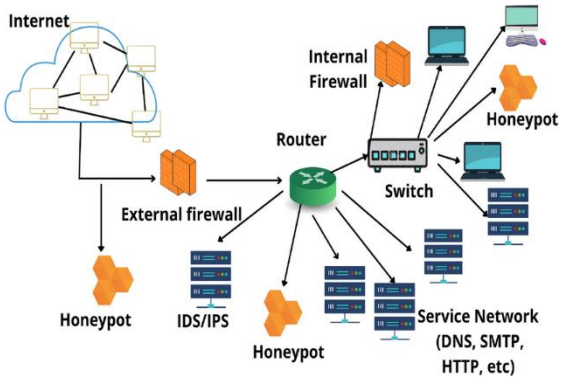

Fig 2: Internal Virtual Honeypot Network

3.2 Data Preprocessing

Preprocessing of the surveyed data consisted of several crucial stages to develop a detailed description of the set. The data preprocessing approach involved cleaning the data to remove or handle irrelevant data points or entries and deal with missing values. This stage also entails organizing the data in a format suitable for analysis and identifying fields of interest that include issue time, originating IP, destination port, service, etc.

Other data preprocessing techniques performed include normalization, where the values for the features are standardized to guarantee uniformity in measurement units, and feature selection, where only essential data features are extracted from the variables. These attributes further help us understand the timing and the approaches taken by offering a more sophisticated level of examination.

Another set of preprocessing operations was data transformation, where categorical metrics were converted into numerical values for easy computations. This step was crucial in ensuring that the data was ready and appropriate for further analysis involving more complex methods.

3.3 Validation

To ensure control, the study used several approaches confirming the study's outcomes. K-means clustering, and outlier analysis were used to find any weird sets of activities, talks, or coordinated propaganda campaigns in the given data set. This analysis was critical as it exposed other subtle methods and patterns malicious actors commonly employ.

Another analysis type of the attack sequence was performed to identify temporal patterns in the attacks, focusing on periodic changes in the attack patterns and frequentation and comparing them with external events or specific threat analytical reports. This step helped validate the discovered patterns of malicious activity based on known indicators of compromise (IOCs) and adversary campaigns.

The data reviewed in the validation stage was also compared with publicly available scientific publications and threat intelligence data. This practice not only guarded against exaggeration or untenable conclusions but also placed the results into the larger cybersecurity analysis ecosystem and highlighted what should be done to counteract the threats and safeguard the networks.

3.4 Visualizing and Reporting

Some of the graphical representations used in the analysis include bar charts, histograms, line charts, and probability plots, among others, to make the data easier to understand as it was in its graphical form. These visualizations were made of attack frequencies over time, attacks' geographical distribution, and interactions between IP addresses and systems.

The final report not only summarized all results and propagated main conclusions but also analysed them, discussed all their implications for further cybersecurity strategies development, and proposed further research directions. Therefore, the report was intended as a scholarly thesis that sought to inform the existing discussion on the consumption of cybersecurity

technology by offering practical suggestions for improving security frameworks.

3.5 Overall Approach

Combining modern quantitative determination techniques with comprehensive data-gathering methodologies, this research offered a profound and prescriptive analysis of cybersecurity threats to different domains. The approach enhanced the research perspectives and advanced the knowledge of the strategies to control and minimize the threats in the information environment. Thus, this approach provided the study with significant opportunities for invaluable information on changes like cyber threats and contributed to enhancing adequate and efficient cybersecurity measures during continuous integration of the digital environment.

4. RESULTS AND DISCUSSION

In cybersecurity, identifying predictable behaviour patterns or motives of attackers is crucial in formulating anticipative strategies to minimize or ward off potential cyber threats. This system certainly involves studying the type of attacks as the connections and interfaces become more complicated and intertwined in the modern world. Still more, it must be noted that honeypot logs are the primary source of information in this context. Far from being a simplified estimate, this work draws from honeypot data revealing daily cyber-attacks from October 2022 to September 2024.

This discussion area describes the observations made when applying the structured methodology suggested in this work on such a vast dataset. The following significant processing steps were used in the current work: data preprocessing, exploratory data analysis, anomaly detection, temporal analysis, integrations with external threat intelligence via Open-Source Intelligence (OSINT), and advanced data visualization and reporting schemes. Imposing an advanced machine learning approach and detailed data mining methods, we could identify a rich set of features and correlations from the given data [3,8,11,12,14,16,21].

The results yielded from these analyses are essential in understanding the pattern and striking capabilities that the attacker may employ in an attack. They support hypothesis H1 and explain the critical trends and patterns of cyber-attacks, which means an increase in the level of the worst attack's sophistication and targeted accuracy over time for RQ1. Thus, the analysis revealed standard techniques and designated cybercriminals' tendencies and novel strategies.

As part of this engaged validation, we have ensured that these effects are consistent, robust, and accurate, making the information presented herein precise and reliable [18, 23]. Furthermore, the increase in applicability to the events witnessed during the documented cyber-attacks has also been evaluated in line with existing threat intelligence (OSINT) [3,16].

Sharing these outcomes is intended to contribute to the vast expansion of informational resources in the cybersecurity field. By providing evidence-based findings, this study intends to create better defensive techniques and further protective measures against stealth—and precision-oriented cyber threats. These outcomes provide the framework for the overall approach to cybersecurity, evidencing the importance of quantitative research in enhancing the scale of cyber protection [1,6].

These findings provide an extensive view of the emerging trends, patterns, and methods that cybercriminals have compounded within six years. This includes praising the advancement in resource utilization, specifying the increase in relevance and target choice, thus supporting H1, and answering RQ1 in detail.

4.1 Data Collection and Preprocessing Results

The log data analyzed is from a honeypot source containing over one hundred million entries, logging over six years from October 2016 to September 2022. This provided the study with intensive participation, and the extensive amount of data made available a profound view of the cyber operations' frequency, dispersion, and nature.

4.2 Summary and Descriptive Analysis

The average middle rate of cyber attacks per day over the six years is estimated at 45,741, with a daily variation recorded at 58 788 standard deviations. 5 (Table 1). This fluctuation suggests that the number of daily cyber-attacks may differ significantly, with some several times less than others, while others may contain several times as many. In the present study, the honeypot system recorded the events for 2,191 days of monitoring; the total entries recorded were 100,218,535.

The maximum number of attacks was identified on the same day, a peak of 888203. However, the high standard deviation combines with this max value, which means there were days when the number of cyber-attacks was significantly higher than others, which can be explained by coordinated global cyber-attacks or certain cyber events. Little over two weeks elapsed between events, and it could be described that ‘quiet days’ when no new threats were discovered were scarce, with only seventeen days observed in six years.

That is why the number of daily detected cyber-attacks is distributed unevenly. Hence, that can be explained as 28 447, the median of the daily attacks count, which can be lower than the mean due to positive skewness. This is because there are many days with an average number of

attacks. However, some days contain a significantly higher number, raising the average, which is complete with attacks.

Quartile ranges provide additional insights into the distribution: Further, the first quartile, Q1, indicated that on a given day, the number of attacks was 16,037 or less, and the third quartile, Q3, indicated that on the other quarter of the days, there were attacks of 58,430 and above. 5 attacks. The interquartile range (Q1 – Q3) revealed an impersonal distribution of 42,393. It is evident from Figure 5 that there was an increased spread in the middle 50% of the data.

Table 4: Descriptive Statistics of Unique Daily Cyber-Attack Counts

Descriptive Statistics	
count	2191.000000
mean	45741.001826
std	58788.500082
min	0.000000
25%	16037.000000
50%	28447.000000
75%	58430.500000
max	888203.000000

4.3 Temporal Analysis

According to the temporal analysis, cyber attacks over the period were chosen for the research. The Figure 3 shows this month-by-month variation with intensified activity in July 2017 and October 2019, two periods in which the world experienced increased cyber activity. Starting in late 2019, an upward trend in the monthly attack volume was discovered, and some limited-month attacks in 2021, such as in February and August, peaked at over 3. 2M and 3. 49M, respectively. While the attack rates were slightly declining in late December 2021, 2022, saw the rates rise back up, with May and June registering more than 3 million attacks.

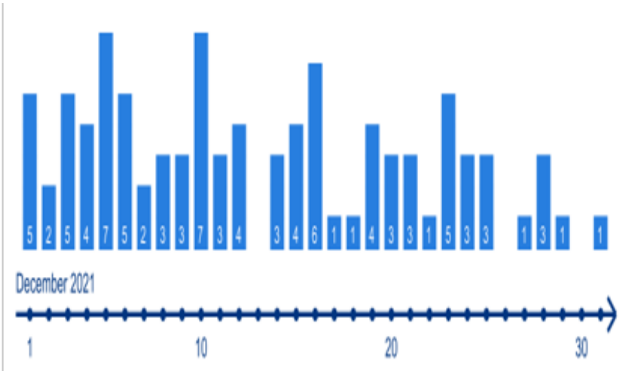

Fig 3: Temporal Distribution of Cyber-Attacks

4.4 Correlation Analysis

The correlational analysis focused on the interactions between different cybersecurity-related factors, which were defined by applying the values of the Pearson correlation coefficient to evaluate the strength and direction of the factors above connections. This process involved 221110 correlation, which was a process of matching the identified threat intelligence sources with the other sources or even matching them with the already known IOCs or the attack campaigns [3,7,16].

The present analysis revealed 8240 correlations reached statistical significance, whereas 2155 correlations did not. The correlation values fluctuated, and values were obtained at approximately 0. All but the last item dipped to a favourable rating of 31, while the final item received a favourable rating of a perfect 1. For instance, data files like malicious-subnet-misp-bro, malicious-subnet-bro.txt, and flows like flow-pcap-malicious-subnet occasionally come in pairs: txt and malicious-subnet-misp-ip-dst.txt revealed a value of 1, which is a perfect positive correlation, meaning that both datasets are either identical or well mirrored. It should be noted, however, that several other pairs also displayed pretty high levels of synchrony, with coefficients of at least 0. 783 to 0.987: However, these imply high statistical relevance of the observed relationships.

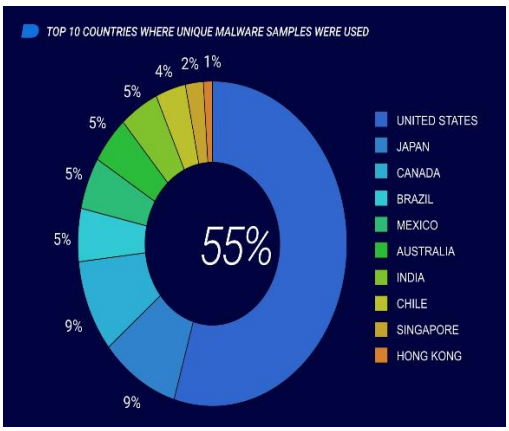

Fig 4: Country Distribution by Percentage

Table 5: Correlation Calculations

Column 1	Column 2	Correla
malicious-subnet-misp-bro.txt	malicious-subnet-misp-ip-dst.txt	
malicious-ip-uceprotect-dnsbl-3.txt	malicious-ip-uceprotect-dnsbl-2.txt	
malicious-ip-firehol-anonymous.txt	malicious-ip-firehol-proxies.txt	0.98733
malicious-ip-dan-torlist-exit-ip.txt	malicious-ip-dan-torlist.txt	0.93326
malicious-ip-blocklist-ssh.txt	malicious-ip-blocklist.txt	0.88054
malicious-subnet-spamhaus-drop.txt	malicious-subnet-snort-pulled-pork.txt	0.84567
malicious-subnet-snort-pulled-pork.txt	malicious-subnet-firehol-spamhaus_drop.txt	0.84567
malicious-ip-firehol-webclient.txt	malicious-ip-firehol-webserver.txt	0.82220
malicious-ip-misp-bro-ipv4.txt	malicious-ip-misp-ip-dst-ipv4.txt	0.78432
malicious-subnet-firehol-anonymous.txt	malicious-subnet-firehol-proxies.txt	0.78302

4.5 Geographic Analysis

The geographical findings in this study showed that attacks are global, with attackers coming from all over the world with representation from six continents and 188 different countries. North America received the highest attack significance, followed by Europe and Asia. The major country under test attacks was the United States, which comprised roughly 47 percent of attacks. It was seen that 611% of the total entries. Other countries in the top fifteen included Russia, China, and the Netherlands; funds from these countries totalled $12 billion. 525%, 8. 188%, and 4. To compare their results with our algorithm, we have calculated that 90% of the entries in this evaluation are exaggerated, followed by 309% of the entries in the second evaluation and 418% of the entries in the third evaluation. Nevertheless, no other country contributed more than 4 percent of the entries into the entrants' pool.

When 20 employees were selected and analysed individually, it was found that they come from the top 20 countries, thus contributing 98 percent of the employment—7% of the total entries indicated that they had utilized a broadband P2P connection. North America, Europe, and Asia were the most involved regions in the continental area, directly affecting over 45%. 017%, 35. 370%, and 22. , respectively, before the increase in the proportion of crossword entries, which constituted 180% of the entries in total. All the South American African Oceania and Antarctic connections contributed only 2. 3% of the total entries.

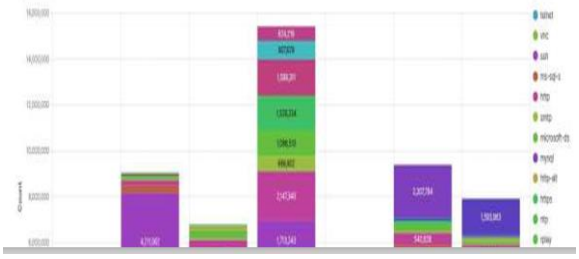

Fig 5: Country Distribution by Percentage

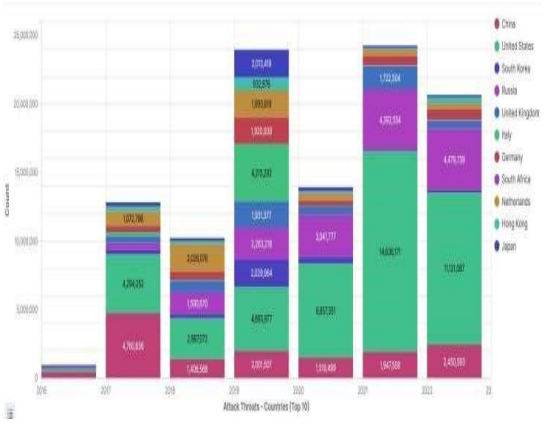

Fig 6: Country Distribution Over Time

4.6 Threat Intelligence Analysis

Examining the new dataset, which consisted of 1,316,585 source IP addresses, the author found several threats with a noticeable distribution. The sources of IPs were dichotomized as the dummy variables: 0, meaning that the system did not contain the source IP address in any Ti databases, and 1, meaning that the system contained the source IP address in a database. The dataset was comprised of seventy-four threat intelligence feeds with the repositories.

Thus, the threat intelligence analysis revealed that 699543 (53. 133%) of the total source IP addresses in the collected list were matched to threat intelligence repositories and received a ''1''median value, whereas 617042 (46. 867%) of entries received the ''0''median value. Although the identifier is not found in the external threat intelligence sources as part of the list in the Detect column, these zero-counts are valuable threats due to being in line with behavioural characteristics related to other malicious conduct identified in the same dataset. The type of analysis further classified the entries into various categories to point out different forms of malicious activities or origins.

Table 6: Threat Analysis

Threat Intelligence Feed or Repository	Matches	Total Source IP
malicious-subnet-uceprotect-dnsbl-3.txt	581,115	44.138%
malicious-subnet-uceprotect-dnsbl-2.txt	300,891	22.854%
malicious-subnet-firehol-webserver.txt	91,983	6.986%
malicious-ip-misp-ip-dst-ipv4.txt	41,701	3.167%
malicious-ip-misp-bro-ipv4.txt	28,080	2.133%

4.7 Source IP Address Analysis

One of them is the analysis of source IP addresses examined in the scope of the study from the point of view of various threat intelligence sources. Over 2,191 days, the SPAMH trapped 100,218,535 entries from 1,316-585 distinct source IP addresses. The IP address is 23. 139. 224. 114 was the month with the highest entry, and the aggregate entry tally reached 2 217 585 (Table IV). When examined more deeply, the authors of entries corresponding to the 20 leading identified sets of unique source IP addresses were 10,835,108 TOP WITH approximately ten pct. The first significant finding was that women constituted 81% of the total entries, indicating a gender bias in the study. About a fourth of these entries have their roots in North America, the second being Europe and third in Asia; the three most populated contributors to these entries are the United States of America, Russia, and China, respectively.

Table 7: Top 10 Source IP Addresses

Source IP	Count	Percentage
23.139.224.114	2,217,585	2.215%
162.142.125.128	1,045,622	1.044%
100.27.42.150	758,851	0.758%
100.27.42.187	754,386	0.754%
100.27.42.157	693,224	0.693%
64.227.110.98	687,625	0.688%
92.63.197.18	677,060	0.677%
143.110.156.7	580,346	0.580%
161.35.232.85	569,259	0.569%
93.115.29.34	531,990	0.532%

4.8 Destination Ports Analysis

The log file analysis found that out of 65,535 destination IP port numbers available, 65,533 were targeted by the attackers during the period under consideration. The top three ports with the highest frequency of attacks were port 5900, associated with the VNC server; port 8, identified with ICMP Echo Requests; and port 22, linked to SSH, contributing to 15. 883%, 10. 223% and 4. 53% of total entries, respectively, for Liberal Arts, while 459% of total entries, respectively, for Science & Health. These ports are typically linked to remote control services and diagnostics tools, meaning that the attackers, while driving the intrusion activity, prefer targeting those applications associated with remote access and diagnosis of the network.

The total entries that these Twenty ports cumulatively presented were Sixty-six point seven one one seven (66,741,317), representing approximately 66. 52% of the total. A trend analysis focusing on the traffic flow in the network for the years 2016, 2017, 2018, 2019, 2020, 2021, and 2022 exposed the level of traffic flow in some of the ports over the years. For some ports, there was an increase in traffic flow, while others saw a decline in flow, and some ports experienced very high traffic flow only for some years.

By incorporating the honeypot data from the last two years, this analysis offers a comprehensive overview of the cyber threats and their global distribution, stresses the targeting of the attacks, and proves the applicability of honeypot data as the key to avoiding and overcoming the existing and emerging threats.

5. CONCLUSION

5.1 Summary of Significant Findings

The emphasis in the research on long-term approaches to critical infrastructure breaches has provided essential data that has improved the understanding of cybersecurity threats. Firstly, the research showed that such events occurred at various times and presented an oscillation, which confirms that cyber attacks on critical infrastructures happen with some frequency and can be associated with certain events and events around the world or in a particular region. Secondly, the authors examined the impact level of cyber disruptions on the critical infrastructure sectors and found out that sectors such as energy and telecommunication are highly vulnerable. These sectors are considered as essential as they provide support to other vital services, which means the consequences could affect various sectors in a chain reaction, depending on what could happen. Finally, in terms of advanced threat patterns, the study noted that the global threat landscape features sophisticated cyber threats, such as state-sponsored cyber threats and advanced persistent threats, which are becoming more sophisticated and constantly changing to account for new vulnerabilities in the systems of critical infrastructures.

5.2 Implications for Future Cybersecurity Policies and Practices

These recommendations imply that they afford the ability to inform and update cybersecurity policies and practices. Measuring the frequency and intensity of focused temporal tendencies in cyber threats allows policy decision-makers, as well as technical IT security practitioners, to devise ways to counter or even anticipate threats, especially during high-risk times. This predictive capability requires the deployment of adaptive security mechanisms that can be easily tweaked in accordance with the real-time threat evaluation. Moreover, the highlighted vulnerabilities in sectors reveal the need for specially focused cybersecurity frameworks to make the standards discovered mandatory for ensuring security in sectors belonging to the list. For instance, improved security measures, including forced multi-factor authentications and more frequent probing of system weaknesses, could be put in place to increase the armour of these integral sectors. Also, the updating and dynamic environment characteristic of cyber threats mentioned in the study implies that existing polio must be periodically subjected to modification to consider state actors and APTs. It could involve international relations in the pronouncement of a coordinated strategy or plan against cyber threats, which is a call for cooperation, sharing of information, and defence coalition.

5.3 Recommendations for Further Research

Considering the evolving cybersecurity landscape, the study proposes several directions for further research that could substantiate and expand upon the current findings. There is a critical need for ongoing research into the specific modalities and impacts of cyber attacks on different sectors of critical infrastructure. Such studies could employ a mix of quantitative and qualitative methodologies to provide a more comprehensive understanding of the vulnerabilities and threats specific to each sector. Additionally, future research should explore the effectiveness of newly implemented cybersecurity policies and technologies, evaluating their real-world efficacy in preventing and mitigating cyber incidents. This could involve longitudinal studies that track the success rates of various cybersecurity measures over time, providing data-driven insights into the most effective strategies for protecting critical infrastructure.

Moreover, considering the international implications of cybersecurity, it is imperative to conduct comparative studies that examine the cybersecurity frameworks and incident response strategies of different countries. Such comparative analyses could identify best practices and foster a greater exchange of knowledge and resources among nations, which is crucial in the fight against global cyber threats. Another promising area of research involves the development and application of advanced artificial intelligence (AI) technologies in cybersecurity. AI can potentially revolutionize the field by enhancing threat detection capabilities and automating responses to security breaches. However, the implications of AI in cybersecurity, including ethical considerations and the risk of AI being exploited by malicious actors, must be thoroughly investigated.

In conclusion, this study not only highlights critical insights into the cybersecurity of critical infrastructures but also sets the stage for a series of actions and further inquiries that are essential for advancing our defensive capabilities in an increasingly digital world. As cyber threats continue to evolve in complexity and scale, it is paramount that research and policy adaptation move at a commensurate pace, ensuring robust protection for the vital systems that sustain modern societies.

6. ACKNOWLEDGMENTS

I would like to extend my gratitude to those who provided informal support and insights that enhanced the quality of this research. While this study did not receive direct funding or formal organizational support, the invaluable discussions and critiques from colleagues and peers in the cybersecurity community played a crucial role in refining the methodologies and interpretations presented.

7. AUTHOR CONTRIBUTIONS

As the sole author of this study, I was responsible for every aspect of the research. This included the initial conceptualization of the study's scope and objectives, the design and implementation of the methodology, the rigorous analysis of data, and the writing and revising of the final manuscript. My comprehensive involvement ensured that the research adhered to the highest standards of academic rigor and integrity.

8. CONFLICTS OF INTEREST

There are no conflicts of interest to declare. I maintained complete autonomy throughout the research process, ensuring that all findings and conclusions drawn are the result of unbiased analyses and discussions. This independence also extends to the absence of any financial or commercial influences that could potentially affect the integrity of the research.

References

[1] Makrakis, Georgios Michail, Constantinos Kolias, Georgios Kambourakis, Craig Rieger, and Jacob Benjamin. "Industrial and critical infrastructure security: Technical analysis of real-life security incidents." Ieee Access 9 (2021): 165295-165325.

[2] Li, Ye, Yu Tu, Qi Fan, Changyin Dong, and Wei Wang. "Influence of cyber-attacks on longitudinal safety of connected and automated vehicles."

Accident Analysis & Prevention 121 (2018): 148-156.

[3] Madnick, Benjamin, Keman Huang, and Stuart Madnick. "The evolution of global cybersecurity norms in the digital age: A longitudinal study of the cybersecurity norm development process." Information Security Journal: A Global Perspective 33, no. 3 (2024): 204-225.

[4] Kuypers, Marshall A., Thomas Maillart, and Elisabeth Paté-Cornell. "An empirical analysis of cyber security incidents at a large organization." Department of Management Science and Engineering, Stanford University, School of Information, UC Berkeley 30 (2016).

[5] Falowo, Olufunsho I., Murat Ozer, Chengcheng Li, and Jacques Bou Abdo. "Evolving Malware & DDoS Attacks: Decadal Longitudinal Study." IEEE Access (2024).

[6] Edwards, Benjamin, Steven Hofmeyr, Stephanie Forrest, and Michel Van Eeten. "Analyzing and modeling longitudinal security data: Promise and pitfalls." In Proceedings of the 31st Annual Computer Security Applications Conference, pp. 391-400. 2015.

[7] Walker-Roberts, Steven, Mohammad Hammoudeh, and Ali Dehghantanha. "A systematic review of the availability and efficacy of countermeasures to internal threats in healthcare critical infrastructure." IEEE Access 6 (2018): 25167-25177.

[8] Moore, Erik, Steven Fulton, and Dan Likarish. "Evaluating a multi agency cyber security training program using pre-post event assessment and longitudinal analysis." In Information Security Education for a Global Digital Society: 10th IFIP WG 11.8 World Conference, WISE 10, Rome, Italy, May 29-31, 2017, Proceedings 10, pp. 147-156. Springer International Publishing, 2017.

[9] Farokhnia Hamedani, M. Essays on Cybersecurity and Information Privacy. ProQuest Dissertations Publishing, University of South Florida, 2023. 30421027.

[10] Rosa, F. R. Global Internet Interconnection Infrastructure: Materiality, Concealment, and Surveillance in Contemporary Communication. ProQuest Dissertations Publishing, American University, 2019. 13902857.

[11] Adewopo, V. Exploring Open Source Intelligence for Cyber Threat Prediction. ProQuest Dissertations Publishing, University of Cincinnati, 2021. 28890231.

[12] Cho, S. Tackling Network-Level Adversaries Using Models and Empirical Observations. ProQuest Dissertations Publishing, State University of New York at Stony Brook, 2021. 28718487.

[13] [13] Muoi, T. D. Handling Network Attacks Exploiting Routing Information Asymmetries. ProQuest Dissertations Publishing, National University of Singapore (Singapore), 2022. 29352339.

[14] Li, G. An Empirical Analysis on Threat Intelligence: Data Characteristics and Real-World Uses. ProQuest Dissertations Publishing, University of California, San Diego, 2020. 27955013.

[15] Hillis, J. S. Enterprise Advanced Persistent Threat Group Identification and Technique Discovery. ProQuest Dissertations Publishing, Marymount University, 2023. 30484790.

[16] Alsarhan, H. F. Real-Time Machine Learning-based Intrusion Detection System (IDS) for Internet of Things (IoT) Networks. ProQuest Dissertations Publishing, The George Washington University, 2023. 30000678.

[17] Al-Haija, Q. A.; Krichen, M.; Elhaija, W. A. Machine-Learning-Based Darknet Traffic Detection System for IoT Applications. Electronics, 11(4), 556. DOI: 10.3390/electronics11040556.

[18] Luitel, A. A Framework for Modeling Data Breach Risk Using Machine Learning Models for High-Dimensional Panel Data. ProQuest Dissertations Publishing, The George Washington University, 2022. 28865998.

[19] Ongun, T. Resilient Machine Learning Methods for Cyber-Attack Detection. ProQuest Dissertations Publishing, Northeastern University, 2023. 30418436.

[20] Mengidis, N.; Panagiotou, P.; Tsikrika, T.; Vrochidis, S.; Kompatsiaris, I. Host-based Intrusion Detection Using Signaturebased and AI-driven Anomaly Detection Methods. Information & Security, 50(1), 37-48. DOI: 10.11610/isij.5016.

[21] Panagiotou, P.; Mengidis, N.; Tsikrika, T.; Vrochidis, S.; Kompatsiaris, I. An in Depth Analysis of Open Source Tools: Host Intrusion Detection System, Intrusion Detection System, and Honeypots, and How They Can Protect a SME's Network.

[22] ProQuest Dissertations Publishing, Utica College, 2019. 22622076.

[23] Butt, S. M.; Reaiche, C. Cognitive Analysis of Intrusion Detection System. Journal of Siberian Federal University. Engineering & Technologies, 15(1), 102-120. DOI: 10.17516/1999-494X-0377.

[24] Barron, T. Addressing the Imbalance between Attackers and Defenders Using Cyber Deception. ProQuest Dissertations Publishing, State University of New York at Stony Brook, 2020. 28091212.

[25] Bobish, M. Sharing Cyber Threat Information Between the United States' Public and Private Sectors. ProQuest Dissertations Publishing, Utica University, 2023. 30488959.

[26] Alowaisheq, E. Security Traffic Analysis Through the Lenses Of: Defenders, Attackers, and Bystanders. ProQuest Dissertations Publishing, Indiana University, 2020. 28259642.