UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE CIÊNCIA DA COMPUTAÇÃO

TATIANE SEQUERRA STIVELMAN

Extracting Cybersecurity Insights from Real-World Incident Data

Work presented in partial fulfillment
of the requirements for the degree of
Bachelor in Computer Science

Advisor: Prof. Dr. Lisandro Zambenedetti Granville
Coadvisor: Dr. Muriel Figueredo Franco

Porto Alegre
Janeiro 2025

ACKNOWLEDGEMENTS

First, I would like to express my heartfelt gratitude to everyone in my support network, especially my partner, Leonardo, my sister, Yasmin, and my dear friend, Valentina. Your unwavering belief in me, even during my toughest moments, meant the world.

I am also deeply thankful to Andréa Goya Tocchetto Osowski, Juliana Tramontina, and Marília Pithan Pereira, the psychiatrists who played a crucial role in helping me maintain my sanity throughout the demanding graduation process.

I would also like to express my gratitude to Axur for providing access to the Polaris data and for their prompt assistance in clarifying my doubts whenever I needed support. Their help was instrumental in the development of my work.

Finally, I want to extend my sincere thanks to my co-advisor, Muriel, and my advisor, Lisandro, for their invaluable guidance and support as I navigated the challenging journey of writing my bachelor's thesis.

ABSTRACT

The increasing reliance on digital platforms and the sophistication of cyberattacks underscore the critical need for proactive and data-driven approaches to cybersecurity. This bachelor thesis examines patterns and vulnerabilities using real-world data collected from Axur's Polaris platform. Through structured analysis and descriptive analytics, the study identifies key threats, such as credential-based exploits (T1078), sector-specific vulnerabilities like public-facing application exploits (T1190), and the targeting strategies of prominent threat actors, including Lazarus Group and APT28. The findings highlight a strong correlation between asset vulnerabilities and attack frequency (Pearson Correlation Coefficient = 0.72), emphasizing the necessity of robust defenses for critical systems like Red Hat Enterprise Linux and Microsoft Windows.

The study contributes practical, sector-specific analysis and a replicable framework for incident analysis, aiming to bridge theoretical research and real-world applications. Future work could extend this research by incorporating predictive analytics and expanding datasets to capture long-term trends and emerging attack vectors. These findings provide actionable insights for enhancing organizational resilience against increasingly sophisticated cyber threats.

Keywords: Cybersecurity. Data-Driven Security. Descriptive Analytics. Incident Data. Threat Analysis.

Extração de Insights de Cibersegurança a Partir de Dados Reais de Incidentes

RESUMO

A crescente dependência de plataformas digitais e a sofisticação dos ataques cibernéticos destacam a necessidade crítica de abordagens proativas e orientadas por dados para a cibersegurança. Este trabalho examina padrões e vulnerabilidades utilizando dados reais coletados da plataforma Polaris da Axur. Por meio de análise estruturada e análises descritivas, o estudo identifica ameaças-chave, como explorações baseadas em credenciais (T1078), vulnerabilidades específicas de setores, como explorações de aplicativos expostos ao público (T1190), e as estratégias de direcionamento de atacantes proeminentes, incluindo o Lazarus Group e o APT28. Os resultados destacam uma forte correlação entre as vulnerabilidades de ativos e a frequência de ataques (Coeficiente de Correlação de Pearson = 0,72), enfatizando a necessidade de defesas robustas para sistemas críticos como Red Hat Enterprise Linux e Microsoft Windows.

O estudo contribui com análises práticas específicas para setores e um framework replicável para análise de incidentes, com o objetivo de conectar a pesquisa teórica às aplicações do mundo real. Trabalhos futuros podem expandir esta pesquisa, incorporando análises preditivas e ampliando os conjuntos de dados para capturar tendências de longo prazo e vetores de ataque emergentes. Esses resultados fornecem insights acionáveis para melhorar a resiliência organizacional contra ameaças cibernéticas cada vez mais sofisticadas.

Palavras-chave: Cibersegurança. Segurança Orientada por Dados. Análise Descritiva. Dados de Incidentes. Análise de Ameaças..

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AI Artificial Intelligence

ML Machine Learning

IoC Indicator of Compromise

APT Advanced Persistent Threat

NGO Non-Governmental Organization

VM Virtual Machine

HTTP Hypertext Transfer Protocol

FTP File Transfer Protocol

DNS Domain Name System

SVR Sluzhba Vneshney Razvedki (Russian foreign intelligence service)

GRU Glavnoye Razvedyvatel'noye Upravleniye (Russian military intelligence service)

RaaS Ransomware-as-a-Service

SQL Structured Query Language

GLM Gordon-Loeb Model

DDoS Distributed Denial of Service

NLP Natural Language Processing

CVaR Cyber Value at Risk

SaaS Software as a Service

API Application Programming Interface

CTI Cyber Threat Intelligence

CONTENTS

1 INTRODUCTION

As our reliance on digital solutions grows, cybersecurity has become a critical focus in safeguarding sensitive information, maintaining operational integrity, and protecting personal and business assets (FRANCO; GRANVILLE; STILLER, 2023). With the shift towards digital in nearly every sector—from finance and healthcare to education and entertainment—vast amounts of data are now stored online, making them vulnerable to cyber threats.

Cyber attacks are increasingly sophisticated, targeting organizations of all sizes and individuals alike, and include tactics such as Distributed Denial of Service (DDoS) attacks and data breaches. These attacks can lead to severe financial losses, disrupt critical infrastructure, and damage reputations (JIN et al., 2022). Effective cybersecurity practices help safeguard personal and business information, ensuring the digital systems we depend on remain secure and resilient against these evolving threats.

In the face of overwhelming data volume and complexity, Artificial Intelligence (AI) has emerged as a powerful tool for enhancing cybersecurity efforts. By leveraging Machine Learning (ML) algorithms and advanced data analysis techniques, AI can detect patterns, anomalies, and early signs of potential threats that may be easily missed by human analysts (ALSHARIF; MISHRA; ALSHEHRI, 2022; FRANCO et al., 2022). As cyber threats evolve, AI's role in cybersecurity becomes essential, aiding in the swift identification and mitigation of vulnerabilities that could otherwise lead to significant security breaches.

In the evolving landscape of cybersecurity, innovative solutions like Axur's Polaris platform have become indispensable in addressing modern cyber threats (Axur Content Team, 2024). Axur is a company that specializes in digital risk protection, offering proactive monitoring and threat detection to safeguard online assets from malicious activities. By providing real-time insights and a comprehensive understanding of digital risks, Axur helps organizations protect their sensitive information, brand reputation, and customer trust in a highly dynamic threat environment.

As part of Axur's catalog, Polaris uses AI and ML to analyze vast amounts of data across the Web, including the Dark Web, to identify and neutralize risks such as brand misuse, credential leaks, and data breaches. By automating threat detection and response, Polaris equips organizations with the agility needed to mitigate potential cyberattacks swiftly and effectively. This reflects the broader trend of utilizing AI to enhance cyberse-

curity, ensuring that businesses and individuals can confidently operate in an increasingly digital and interconnected world.

As Polaris collects data from several differente sources, it is possible to use this data to infer insights about risks and trends of cybersecurity, thus improving the decision-making in companies. Therefore, for this bachelor thesis, we collected 6000 insights generated by Polaris over a six-month period, from 11/03/2024 to 09/09/2024, which represents 33% of all Polaris' data during the period. This dataset included critical information such as the targets of the attacks, methods of detection, recommended protective measures, and key identifiers of the attacks. After collection, the data was enriched, processed, and stored as structured data, enabling efficient classification and querying for deeper analysis.

Using this structured dataset, we identified the instances with the highest number of occurrences and combined them to generate data structures that detailed the frequency of various combinations. This allowed us to uncover correlations between different aspects of the insights, offering valuable perspectives. Such correlations can be instrumental in making strategic decisions, such as selecting high-priority assets for new products or determining the most effective security measures a company should implement to mitigate potential risks.

The rest of this Bachelor's Thesis is organized as follows. Chapter 2 provides an overview of the cybersecurity landscape and introduces key concepts essential for understanding the subsequent chapters; Chapter 3 reviews related works, offering context and insights into existing research in the field; Chapter 4 details the step-by-step process of extracting new insights from Polaris data, outlining the methodology and tools used; Chapter 5 presents the results, accompanied by an in-depth analysis to interpret their significance; Finally, Chapter 6 concludes the thesis with a discussion of the findings and proposes directions for future research.

2 BACKGROUND

2.1 Cybersecurity Landscape Overview

Cybersecurity has been a concern since the early years of the Internet in the 1980s. As the Internet has grown and evolved, so too have the sophistication and frequency of cyberattacks, driving the need for equally advanced cybersecurity measures (G‘UZOROVICH, 2024).

In today’s world, the Internet permeates nearly every aspect of society, shaping how we communicate, work, and access information. It connects us through social media and entertainment while powering critical sectors like healthcare, finance, politics, and military operations. As reliance on digital systems grows, the need for robust cybersecurity has become more urgent than ever (CIUMACENCO DENIS PLEȘCA, 2024).

Suffering a cyber attack can be very damaging, potentially resulting in financial losses, compromising individuals’ privacy and rendering critical resources inaccessible to those who rely on them (FRANCO; GRANVILLE; STILLER, 2023). In 2023 losses connected to cybercrime complaints reached $12.5 billion (Florian Zandt, 2024), a staggering figure that emphasizes the critical importance of strong cybersecurity measures in safeguarding individuals, businesses, and organizations from escalating threats.

Currently cybersecurity functions on multiple levels, blending technical measures, procedural protocols, and user-centered strategies to mitigate risks and safeguard digital assets. Its primary objective is to build strong defenses against diverse threats, including malware, phishing, ransomware, and insider attacks. It also involves identifying system and network vulnerabilities in advance and applying solutions or updates to address them. This adaptability is what enables cybersecurity to effectively counter a wide range of cyber threats, making it both its core purpose and greatest strength (CIUMACENCO DENIS PLEȘCA, 2024).

In this bachelor thesis we used the MITRE ATT&CK framework, a globally accessible knowledge base of adversarial tactics and techniques based on real-world observations. Developed by the MITRE Corporation, this framework provides a structured and detailed taxonomy of how cyber adversaries operate across different stages of an attack, from initial access to data exfiltration (STROM et al., 2018). By mapping cyber threats to the ATT&CK matrix, organizations can better understand potential vulnerabilities in their systems, enhance threat detection capabilities, and implement targeted defenses. Specifi-

cally, in this thesis, the framework was employed to identify and categorize the tactics and techniques used in real-world cybersecurity incidents collected from the Polaris platform.

2.2 Ransomware-as-a-Service (RaaS)

RaaS operates as a structured model that facilitates the widespread deployment of ransomware, allowing individuals with limited technical expertise to conduct sophisticated cyberattacks. By providing ready-made ransomware tools and support in exchange for a share of the ransom, RaaS lowers the entry barrier for cybercriminal activities and has significantly contributed to the proliferation of ransomware incidents (DURAIBI; KAUR; PAWAR, 2023).

A key tactic employed in these attacks is double extortion, where attackers not only encrypt the victim's data but also exfiltrate sensitive information, leveraging the threat of public exposure to compel payment (DURAIBI; KAUR; PAWAR, 2023). This approach intensifies the impact of an attack, increasing the risks of privacy violations, reputational damage, and financial losses. Further advancing this strategy, triple extortion introduces additional pressure points, such as threatening DDoS attacks or targeting third parties, including clients or partners, to force compliance. These multi-layered extortion techniques underscore the evolving complexity and effectiveness of modern ransomware operations.

2.3 Pearson Correlation Coefficient

The Pearson correlation coefficient is a statistical measure used to quantify the strength and direction of the linear relationship between two variables. Represented by the symbol r, it produces a value ranging from -1 to 1. A value of $r = 1$ indicates a perfect positive linear relationship, meaning that as one variable increases, the other also increases proportionally. Conversely, $r = -1$ signifies a perfect negative linear relationship, where one variable decreases as the other increases. A value of $r = 0$ suggests no linear correlation between the variables. The Pearson correlation coefficient is widely used in data analysis to identify patterns and relationships, making it a valuable tool for exploring dependencies and drawing insights.

2.4 Polaris Platform

Among the tools employed to combat cyberattacks we can cite AI, which cybersecurity is increasingly using to stay ahead of evolving cyber threats. AI enhances cybersecurity by enabling the detection of anomalies and threats in real time, often with greater speed and accuracy than traditional methods (FURHAD; NOWROZY; SARKER, 2021).

ML algorithms analyze vast datasets to identify patterns indicative of potential cyberattacks, such as unusual user behavior or network activity. Additionally, AI-powered tools automate routine cybersecurity tasks, such as patch management and threat intelligence, freeing up human experts to focus on complex challenges (FURHAD; NOWROZY; SARKER, 2021).

Polaris, developed by Axur, is an AI-powered Cyber Threat Intelligence (CTI) platform designed to enhance cybersecurity operations. It continuously scans a multitude of sources—including news outlets, forums, reports, and threat feeds—to detect and summarize relevant cyber threats, vulnerabilities, and attack vectors. By correlating this information with an organization's specific attack surface, Polaris delivers curated, actionable insights that enable security teams to swiftly identify, prioritize, and mitigate risks. The platform also facilitates proactive threat hunting, allowing for in-depth investigations into potential threats across various assets (Axur Content Team, 2024).

3 RELATED WORK

The field of cybersecurity has been extensively studied, with research addressing various facets, including data-driven insights, risk analysis, human factors, and economic modeling. This chapter reviews significant works that relate to this thesis, focusing on their methodologies, objectives, and contributions to the field. The aim is to position this research in the context of existing literature and highlight its unique contributions.

3.1 Data-Driven Insights in Cybersecurity

This bachelor's thesis aligns closely with studies that emphasize the extraction and analysis of cybersecurity incident data. For instance, the CASIE framework (SATYAPANICH; FERRARO; FININ, 2020) focuses on automated extraction of cybersecurity events from unstructured text using deep learning techniques. CASIE can identify and label important parts of a cybersecurity event, like who the attacker is, who the victim is, and what tools were used. It organizes this information into a clear and structured format, making it easier for automated systems to analyze and use this data to improve cybersecurity defenses. However, unlike CASIE, which uses only cybersecurity news a data source, this bachelor thesis also uses public cybersecurity reports, threat feeds and cybersecurity groups.

3.2 Economic Models for Cybersecurity Investments

The Gordon-Loeb Model (GLM) (GORDON; LOEB; ZHOU, 2016) provides a theoretical foundation for optimizing cybersecurity investments, balancing the costs and benefits of protection. By segmenting information sets based on value and vulnerability, the model helps organizations allocate budgets efficiently. While the GLM offers a high-level economic perspective, this thesis provides practical, operational insights. Its use of SQL queries and statistical analyses directly addresses patterns in real-world incident data, enabling organizations to understand vulnerabilities and attacks in specific contexts. Unlike the GLM's focus on theoretical cost-benefit analyses, this research emphasizes actionable intelligence derived from observed incidents.

3.3 Behavioral Aspects in Cybersecurity

Understanding human factors in cybersecurity is critical, as demonstrated in the study "Review and insight on the behavioral aspects of cybersecurity" (LAHCEN R.A., 2020). The work "Review and insight on the behavioral aspects of cybersecurity" explores how human behavior and decision-making influence cybersecurity outcomes, applying behavioral and criminological theories to predict and mitigate insider threats. While this thesis solution does not explicitly address human factors, it complements this study by providing technical insights into the patterns and vulnerabilities that human actions might influence. Both works aim to enhance cybersecurity preparedness but from distinct perspectives—behavioral versus technical.

3.4 Language Models for Cybersecurity Trends

The study "What Are the Latest Cybersecurity Trends? A Case Study Grounded in Language Models" (VLADESCU et al., 2021) utilizes Natural Language Processing (NLP) techniques, to analyze trends from news articles. Its temporal analysis of cybersecurity topics provides a macro-level view of evolving threats. In contrast, this bachelor thesis focuses on micro-level patterns within structured incident data, offering specific insights for sectors and assets. While the case study emphasizes topic clustering, this thesis' approach is grounded in practical applications, such as identifying which sectors face the most significant threats or how vulnerabilities correlate with attack frequencies.

3.5 Simulations and Risk Models in Cybersecurity

SIM-Ciber (NUNES et al., 2024) uses Monte Carlo simulations and Bayes' theorem to quantify risks and impacts. SIM-Ciber focuses on simulating financial and technical risks for organizations based on predefined profiles. This thesis differs in its descriptive analytics, which rely on real-world data to uncover trends and correlations. While SIM-Ciber emphasizes predictive modeling, this thesis provides granular insights into existing incidents, making them directly applicable to organizations seeking to understand past and present cybersecurity challenges.

3.6 Complementary Approaches in Cybersecurity Analytics

The SECAdvisor tool (FRANCO et al., 2024) provides recommendations for optimal cybersecurity investments using economic models such as the GLM. It emphasizes resource allocation across segmented organizational assets. Similar to the GLM, SECAdvisor focuses on financial optimization. This bachelor thesis, however, prioritizes descriptive analytics and actionable intelligence derived from incident data, offering insights that organizations can immediately apply to their security strategies. These complementary approaches address different stages of cybersecurity planning: this thesis informs the understanding of risks, while SECAdvisor supports investment decisions.

Another work that shares similarities with this bachelor's thesis is the RCVaR approach (FRANCO et al., 2023), which focuses on estimating financial risks associated with cybersecurity threats. Both this thesis and RCVaR use real-world data to extract meaningful cybersecurity insights. However, their goals and methodologies differ significantly. While this bachelor's thesis emphasizes descriptive analytics to uncover patterns and correlations in cybersecurity incidents, RCVaR prioritizes the financial dimension by estimating potential monetary losses through metrics like Cyber Value at Risk (CVaR). Additionally, RCVaR incorporates statistical models and cost analysis to help organizations quantify risks and allocate resources effectively. In contrast, this thesis provides technical and operational insights, such as identifying attack patterns and vulnerabilities, making it more suited for addressing sector-specific cybersecurity challenges. Together, these works highlight the diverse ways real-world data can be utilized to tackle cybersecurity issues from both technical and economic perspectives.

3.7 Summary of Related Work

This bachelor's thesis builds on existing research by integrating structured data analysis, correlation techniques, and practical questions into the cybersecurity domain. While CASIE addresses automated event extraction from unstructured text and the GLM, SECAdvisor, and RCVaR focus on economic modeling and financial risk estimation, this thesis distinguishes itself by offering detailed, actionable insights grounded in real-world incident data. Furthermore, it complements behavioral studies by providing technical perspectives on patterns and vulnerabilities influenced by human actions.

Additionally, this thesis fills a gap left by works like SIM-Ciber and language

model studies, which emphasize simulations and macro-level trends, respectively, by focusing on micro-level patterns and sector-specific challenges. By combining descriptive analytics, correlation techniques, and real-world data from multiple sources, this thesis contributes to a comprehensive understanding of cybersecurity challenges and supports more effective decision-making for organizations.

To contextualize this thesis further within the related work, two tables are presented. Table 3.1 provides a comparison of the objectives and some key features of this bachelor's thesis alongside other works, highlighting their use of real-world incident data and focus on micro-level insights. Table 3.2 compares the specific data sources utilized by works that incorporate real-world incident data, showcasing the diversity of approaches and data types leveraged in cybersecurity research.

Table 3.1: Comparison of related work with this bachelor's thesis

Solution	**Objective**	**Real-World Incident Data**	**Focus on Micro-Level Insights**
CASIE	Automated extraction of cybersecurity events using deep learning	Yes	Yes
GLM	Optimization of cybersecurity investment	No	No
SIM-Ciber	Simulations for risk and economic impacts of cyberattacks	No	No
SECAdvisor	Recommendations for optimal cybersecurity investments using economic models	No	No
Language Model Study	Trend analysis using NLP and clustering techniques	Yes	No
Behavioral Study	Human behavioral aspects of cybersecurity	No	No
RCVaR	Estimation of financial risks of cyberattacks using real-world data	Yes	No
This Bachelor's Thesis	Analysis of real-world incidents, correlations, and sector-specific patterns	Yes	Yes

Table 3.2 highlights the variety of data sources leveraged by different works that utilize real-world incident data for cybersecurity research. Notably, while CASIE focuses solely on news as a data source, other works like the Language Model Study and RCVaR incorporate public cybersecurity reports, providing broader perspectives. In contrast, this bachelor's thesis uniquely integrates multiple data sources—news, public cybersecurity

Table 3.2: Comparison of data sources utilized by works that use real-world incident data

Solution	**News**	**Public Cybersecurity Reports**	**Threat Feeds**	**Cybersecurity Groups**
CASIE	Yes	No	No	No
Language Model Study	Yes	Yes	No	No
RCVaR	No	Yes	No	No
This Bachelor's Thesis	Yes	Yes	Yes	Yes

reports, threat feeds, and cybersecurity groups—making it the most comprehensive in terms of data diversity. By combining these sources, this work provides a richer and more holistic analysis of real-world cybersecurity incidents, bridging gaps observed in prior studies. This multidimensional approach ensures a more accurate and detailed understanding of incident patterns, contributing significant value to the field.

4 APPROACH

This bachelor thesis aims to identify cybersecurity patterns and trends, empowering companies to make more informed and strategic decisions in safeguarding their assets. The approach to the proposed solution is shown in Figure 4.1. This solution is divided into three stages: *Data Preparation*, *Data Selection and Combination*, and *Results Analysis*. Throughout this chapter, we will detail and discuss each of these stages.

4.1 Data Preparation

This stage is divided in 3 substeps: *Collect insights as unstructured data*, *Process and enrich data* and *Store as structured data*.

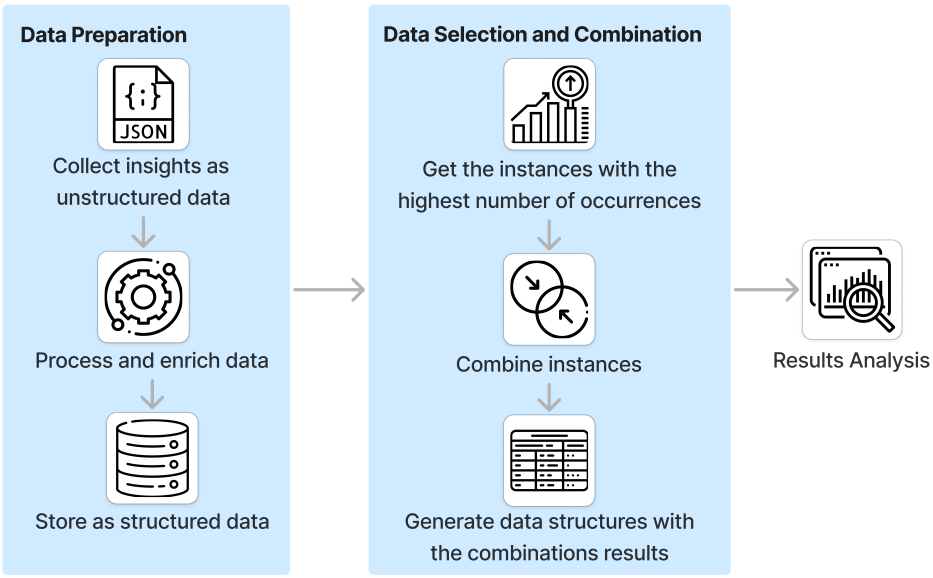
The first substep, *Collect insights as unstructured data*, involves retrieving insights from Polaris using Axur's internal API. These insights, stored in a NoSQL database for optimal handling of unstructured data, were fetched for this bachelor thesis. A total of 6,000 insights were collected, with creation dates ranging from 11/03/2024 to 09/09/2024. Table 4.1 presents a comprehensive list of some of the data elements contained within each insight.

Table 4.1: Insight fields

Field	Description
id	Unique identifier for the insight
createdAt	Timestamp when the insight was created
currentTitle	Title of the insight
currentSummary	Summary of the insight
currentThreatActor	List of involved threat actors
currentMalware	List of associated malware
currentCompanies	List of target companies
currentLocation	List of locations involved
currentIndustries	Affected industries
currentTimeline	Time period when the threat was active
currentMethods	Methods used by the threat actors
currentThreatLevel	Level of threat (e.g., low, medium, high)
currentCourseOfAction	Recommended course of action for each of the targeted assets
currentAttacks	Techniques and tactics used in attacks
currentVulnerabilities	Associated vulnerabilities
currentSources	Sources for the insight
currentKeys	Term keys categorizing the insight

Source: The authors

Figure 4.1: Approach flowchart



In the second substep, *Process and Enrich Data*, the insights were categorized into two main pillars: *Tech*, which encapsulates technical details, and *Business*, which focuses on administrative aspects. Each pillar was further subdivided: the *Tech* pillar was divided into *Identification*, which captures elements for recognizing threats, and *Detection*, which involves mechanisms for identifying security breaches. Meanwhile, the *Business* pillar was split into *Target*, which identifies the intended victims or assets, and *Protection*, which outlines strategies and measures to safeguard against threats.

In the third substep, *Store as structured data*, the enriched insights were migrated to a MySQL database, chosen for its robust capabilities in managing structured data. The database schema was designed to reflect the pillars established in the previous step, with tables dedicated to *Identification*, *Detection*, *Target*, and *Protection*. This structure ensures efficient storage and retrieval of insights, facilitating advanced queries and analysis to support actionable decision-making. The data was splitted in the following way:

- Identification data
  - Threat actors
  - Vulnerabilities

    - Tags
    - Attacks
    - Malware
- Detection data
    - IoCs
    - Sources
- Target data
    - Sectors
    - Organizations
    - Locations
- Protection data
    - Course of action
    - Asset

4.2 Data Selection and Combination

This stage also has 3 substeps, they are, *Get the instances with the highest number of occurrences*, *Combine instances* and *Generate data structures with the combinations results*.

In the first substep, *Identify the instances with the highest frequency of occurrence*, we used SQL queries to extract the most frequently occurring instances, providing a richer dataset for further analysis and exploration. From the stored data, only a subset was selected for further analysis—specifically, the entries that appeared most promising in terms of their potential to generate valuable insights. This selection was guided by the project's objectives and the perceived relevance of the data. The remaining data, while not utilized in the current scope, holds potential for future exploration and could contribute to extending or refining the outcomes of this work.

Here are the selected instances:

- Targeted sectors: industry sectors identified as the most frequently targeted based on the results of the SQL queries.
    - Education

  - Finances and Insurance
  - Healthcare
  - Manufacture
  - Power and Chemical
  - Public Sector and NGOs
  - Retail and e-commerce
  - Technology
  - Transport and Logistics
- Targeted organizations: companies identified as the most frequently targeted based on the results of the SQL queries.
  - **Apple:** is a multinational technology company known for designing and manufacturing innovative consumer electronics, software, and services, including the iPhone, Mac, and App Store.
  - **Change Healthcare:** is an American healthcare technology company that provides data and analytics-driven solutions to improve clinical, financial, and operational outcomes in the healthcare industry.
  - **Cisco:** is a technology leader that develops networking hardware, software, and cybersecurity solutions, supporting secure and seamless connectivity worldwide.
  - **Crowdstrike:** is a cybersecurity company that provides cloud-based endpoint protection and threat intelligence to detect, prevent, and respond to cyberattacks.
  - **D-Link:** is a company specializing in creating networking products like routers, switches, and smart home solutions to support reliable connectivity.
  - **Google:** is a global technology company offering Internet services, software, cloud platforms, and hardware products like smartphones and smart home devices.
  - **Microsoft:** is a multinational technology company renowned for developing software, hardware, cloud services, and the widely used Windows operating system.
  - **Red Hat:** is a software company specializing in open-source solutions, including Linux distributions, cloud technologies, and enterprise software.

    - **Snowflake:** is an American cloud-based data storage company

- Targeted countries: countries identified as the most frequently targeted based on the results of the SQL queries.
    - United States
    - United Kingdom
    - Russia
    - Germany
    - France
    - Canada
- Targeted assets: technologies identified as the most frequently targeted based on the results of the SQL queries.
    - **Red Hat Enterprise Linux:** a commercial Linux operating system developed by Red Hat for enterprises.
    - **Microsoft Windows:** is a product line of proprietary graphical operating systems developed and marketed by Microsoft.
    - **VMware ESXi:** is a hypervisor that allows users to create virtual machines (VMs) on a physical server.
    - **Google Chrome:** is a web browser developed by Google.
    - **Apple macOS:** is Apple's operating system designed for Apple computers.
    - **CrowdStrike Falcon:** is a cloud-native platform that monitors what is happening on the computers on which it is installed, looking for signs of cyberattacks.
    - **Linux Kernel:** is the core of the Linux operating system, managing hardware resources and enabling communication between hardware and software applications
    - **Microsof Office:** is a suite of productivity software applications, including Word, Excel, PowerPoint, and Outlook, designed for document creation, data analysis, presentations, and communication
- Attacks: attack tactics mapped by the MITRE ATT&CK framework identified as the most frequently used based on the results of the SQL queries.
    - **T1078 - Valid Accounts:** is a tactic used in penetration testing where an

attacker exploits legitimate user accounts to gain unauthorized access to a system or network. This tactic involves obtaining valid credentials through various means, such as phishing, credential stuffing, or social engineering, and then using those credentials to bypass security controls and access sensitive information or perform malicious actions. The primary goal of this attack is to exploit legitimate access to avoid detection and escalate privileges within the compromised environment (STERNSTEIN STERN SECURITY; MARK WEE; MENACHEM GOLDSTEIN; NETSKOPE; PRAETORIAN; PRASAD SOMASAMUDRAM, 2024).

- **T1071 - Application Layer Protocol:** is a tactic where attackers exploit vulnerabilities or misconfigurations in application layer protocols to gain unauthorized access or disrupt services. This tactic targets protocols such as HTTP, FTP, or DNS, often using techniques like protocol manipulation, command injection, or exploiting known weaknesses in the protocol implementations. Using these common protocols, attackers can blend their traffic with legitimate network activity, making detection more challenging. The primary goal is to bypass security controls, exfiltrate data, or disrupt the normal functioning of the targeted application (MICHAEL, 2024).
- **T1190 - Exploit Public-Facing Application:** is a tactic employed by attackers to target vulnerabilities in applications that are accessible over the Internet. This tactic involves exploiting weaknesses such as unpatched software, misconfigurations, or insecure coding practices to gain unauthorized access, execute arbitrary code, or disrupt the application's functionality. The primary goal is to compromise the application and potentially gain a foothold in the broader network, often leading to data breaches or further malicious activities. Attackers commonly use tools like vulnerability scanners and exploit frameworks to identify and exploit these vulnerabilities (WEIZMAN, 2024).
- **T1566 - Phishing:** is a tactic where attackers trick individuals into disclosing sensitive information, such as credentials or personal data, by masquerading as a trusted entity through emails, websites, or messages. Attackers often craft convincing communications that appear to come from legitimate sources, such as banks, companies, or colleagues, and include links or attachments that lead to fake login pages or malicious downloads. The primary goal is to deceive the victim into providing confidential information, which can then be

used for unauthorized access or further malicious activities (RAVICH CARDINALOPS; OHAD ZAIDENBERG, 2024).

- Threat actors: attackers identified as the most active based on the results of the SQL queries.
    - **Lazarus Group:** also known by aliases such as Hidden Cobra and Zinc, is a highly prolific cyber espionage and cybercrime group believed to be operated by or on behalf of the North Korean government. The group is known for its advanced persistent threat (APT) activities, which include espionage, data theft, and financially motivated cyber attacks (SOCRadar Research, 2021).
    - **Kimsuky:** also known as Velvet Chollima, Black Banshee, or Thallium, is a cyber espionage group believed to be operating from North Korea. The group is known for its targeted attacks against various sectors, including government, military, think tanks, and organizations related to North Korean interests (SOCRadar Research, 2023b).
    - **APT29 or Cozy Bear:** also known by various aliases such as The Dukes, Nobelium, and Yttrium, is an APT group widely believed to be associated with the Russian Foreign Intelligence Service (SVR). This group is renowned for its sophisticated cyber espionage activities, which include long-term compromises of targeted networks, data theft, and intelligence gathering (SOCRadar Research, 2023a).
    - **APT28:** also known as Forest Blizzard, Fancy Bear, STRONTIUM, Sednit, and Unit 26165, is a Russian GRU-linked cyber espionage group active since at least 2010. It targets government, energy, transportation, NGOs, and other sectors globally, focusing on intelligence gathering to support Russian foreign policy (SOCRadar Research, 2024).
- Malwares: malwares identified as the most frequently used based on the results of the SQL queries.
    - **Akira:** is a ransomware that has been operating since March 2023 and has targeted multiple industries, primarily in North America, the UK, and Australia. It functions as a RaaS and exfiltrates data prior to encryption, achieving double extortion (PRADHAN, 2024).
    - **Alphv or BlackCat:** is a type of malware created by Russian-speaking cybercriminals, with links to defunct russian threat actor groups. Active since

2021, BlackCat is written in the Rust programming language, making it difficult to detect and remove, and targets industries such as construction, energy, healthcare, and technology. BlackCat operates as a RaaS model, offering high payouts to affiliates and employing advanced features like customizability for different operating systems, triple extortion tactics (decryption ransom, data exposure threats, and DoS/DDoS attacks), and a public data leak site to pressure victims into paying. (Akamai Technologies, 2024)

- **Black Basta:** is a ransomware group operating as a RaaS, first identified in April 2022. They employ double extortion tactics, demanding payment for both decrypting data and preventing the release of stolen information. Their operations have impacted over 500 organizations across various industries and critical infrastructures in North America, Europe, and Australia. Initial access is typically gained through methods such as phishing and exploitation of known vulnerabilities. Once inside a network, they move laterally to identify critical systems and data before deploying ransomware (PALIWAL, 2024).
- **Blacksuit:** is a ransomware operation that emerged in April/May 2023, operating privately without affiliates or a RaaS model, BlackSuit employs phishing emails with malicious attachments or links, malvertising, and exploits vulnerabilities in public-facing applications to gain initial access. Post-infiltration, the group utilizes tools to establish persistence, escalate privileges, and move laterally within networks (BARRY, 2024).
- **Lockbit**: is a RaaS group, active since 2019, known for its prominence and destructiveness in global ransomware attacks. Operating on a profit-sharing model, it enables affiliates to deploy its ransomware, encrypting victims' critical data and demanding significant ransoms (Flashpoint Intel Team, 2023).
- **Qilin:** is a RaaS affiliate program, using Rust-based ransomware for tailored attacks on Windows, Linux, and ESXi systems. It employs double extortion tactics, encrypting data while exfiltrating sensitive information, demanding ransom for decryption, and threatening to release data even after payment. The ransomware supports various operator-controlled encryption modes for flexibility in attacks (BLEIH, 2024).
- **Ransomhub:** is a RaaS operation that gained prominence in early 2024 on underground cybercrime forums, RansomHub has rapidly built a reputation for its aggressive attacks targeting a wide range of systems, including Windows,

macOS, Linux, and VMware ESXi environments. The ransomware stands out for its advanced encryption techniques, making it a very strong threat (Check Point Team, 2024).

In the next substep, *Combine instances*, the data we selected in the previous substep was strategically paired, guided by questions that we aimed to address with this analysis. We then used SQL queries to determine the frequency of occurrence for each paired instance.

We paired **sectors** with the **attacks** to answer the question: *If my company operates in sector X, what types of attacks should I be most concerned about?*. We paired **assets** with their respective **number of attack occurrences and vulnerabilities** to explore questions such as: *Between technologies X and Y, which has more vulnerabilities? Which suffers more attacks? Does a higher number of vulnerabilities correlate with more attacks?*.

We paired **threat actors** with **organizations** to identify potential patterns in their attack strategies. Organizations were selected because they encapsulate a combination of relevant factors for the analysis, including sector, location, market share, and reputation. Lastly, we paired **malwares** with **locations** to investigate whether patterns in malware attacks could be linked to their targeted locations. Locations were chosen because they reflect a blend of critical influences, including geopolitical dynamics, economic conditions, and regional characteristics that are pertinent to the analysis.

The last substep, *Generate data structures with the combinations results*, involved exporting the tables generated by the SQL queries in the previous substep to Google Sheets and creating graphs to facilitate visual analysis. To investigate the relationship between the number of incidents involving an asset and the number of vulnerabilities exploited, the Pearson correlation coefficient was applied. This method was chosen because it quantifies the strength and direction of the linear relationship between two variables, providing a clear measure of how closely the number of incidents is associated with the number of vulnerabilities.

4.3 Results Analysis

The final step consisted in analyzing the output from the previous step to extract insights. This process included identifying correlations, patterns, trends, and outliers, and

conducting further research online to validate the proposed insights. The rationale for these insights was explored across various dimensions, including technical, financial, and political factors, to ensure a comprehensive understanding of the findings.

In the next chapter, we will present the graphs generated from the data combinations and delve into a detailed discussion of the results. We will explore the implications of these findings, providing an interpretation of what they reveal about cybersecurity threats and vulnerabilities. Additionally, we will contextualize the results, linking them to real-world scenarios and broader industry insights to offer a thorough understanding of their significance.

5 RESULTS AND DISCUSSION

In this chapter, we present the analysis of the data obtained, along with possible explanations and insights derived from the data and identified trends.

5.1 Attacks x Sectors

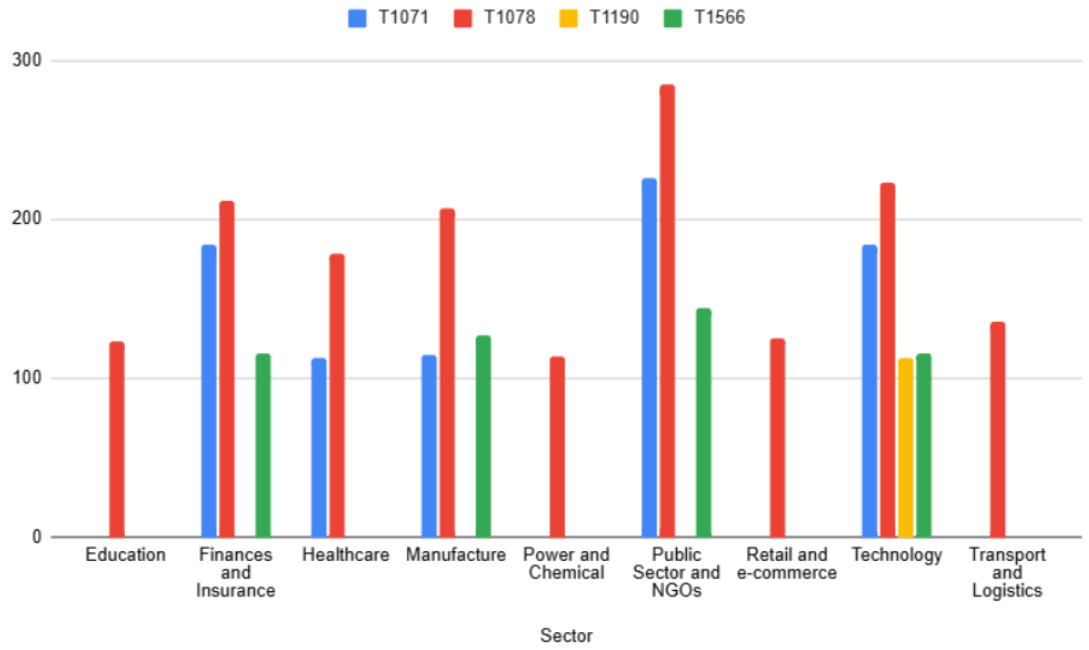
Figure 5.1 shows a graph with the occurrences of the attack codes T1078 - Valid Accounts, T1071 - Application Layer Protocol, T1190 - Exploit Public-Facing Application and T1566 - Phishing, in the sectors Education, Finances and Insurance, Healthcare, Manufacture, Power and Chemical, Public Sector and NGOs, Retail and e-commerce, Technology and Transport and Logistics.

The attack T1078 - Valid Accounts stands out as the one with the highest number of occurrences and the only attack observed across all sectors. This widespread presence can be attributed to several factors. Firstly, valid accounts are fundamental for accessing any system, application, or service, making them critical across all industries, from finance to healthcare to manufacturing. This means any sector can be a victim to this attack. Secondly, attackers can obtain valid accounts through a variety of methods, including phishing, credential stuffing, or exploiting insider threats. These methods are relatively straightforward to execute and can target any sector. Moreover, using legitimate credentials often enables attackers to evade detection by standard security measures, as the activity may appear as normal user behavior.

The Public Sector and NGOs emerge as the most attacked sector. Governments and non-governmental organizations handle vast amounts of sensitive information, including personal data and national security information, which makes them particularly appealing to attackers. These organizations are frequently targeted by hacktivists and politically motivated groups aiming to influence public policy, gather intelligence, or disrupt critical services as a form of protest or leverage. Additionally, Public Sector and NGOs often operate under budgetary constraints, which may lead to weaker cybersecurity defenses compared to well-funded sectors such as finance or technology, further increasing their vulnerability.

In contrast, the attack T1190 - Exploit Public-Facing Application is exclusively observed in the Technology sector. Technology companies often deliver digital products and services accessible over the Internet, such as SaaS platforms, APIs, and cloud-based

Figure 5.1: Occurrences of attack codes by sector



solutions. These offerings inherently expand the attack surface, exposing numerous entry points that attackers can exploit. Compromising a technology company's platform or service not only impacts the company itself but may also provide attackers with indirect access to other sectors that depend on the compromised technology, making T1190 a highly strategic and targeted method of attack.

5.2 Attack Occurrences and Different Vulnerabilities x Assets

Figure 5.2 shows a graph comparing the number of attacks to the assets Red Hat Enterprise Linux, Microsoft Windows, VMware ESXi, Google Chrome, Apple macOs, CrowdStrike Falcon, Linux Kernel and Microsoft Office to the number of different vulnerabilities found on each of them.

The Pearson coefficient of 0.72 indicates a strong correlation between the number of occurrences and the number of vulnerabilities found in each asset. This suggests that assets with more documented vulnerabilities are more likely to experience a higher number of attacks. This relationship is visually represented in Figure 5.3, where a scatter chart plots the occurrences of attacks and the number of vulnerabilities for various assets.

The assets directly related to operating systems — Red Hat Enterprise Linux,

Figure 5.2: Occurrences of attack and different vulnerabilities found by asset

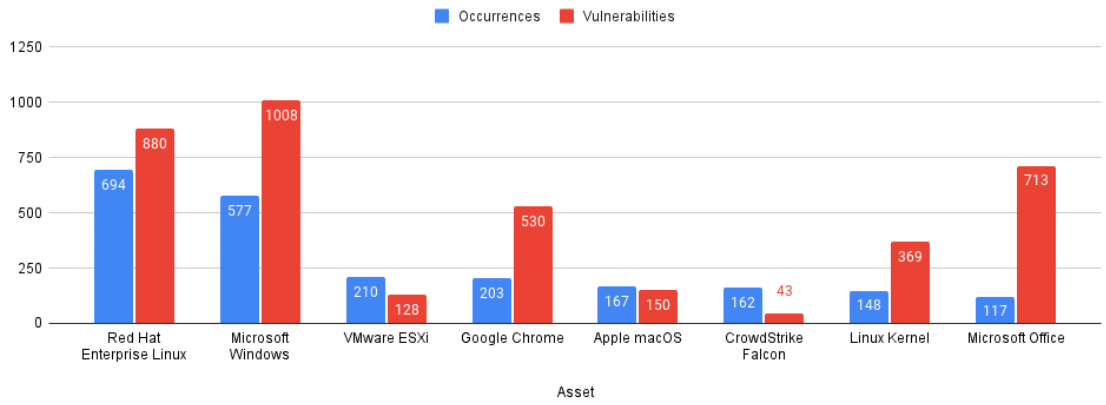
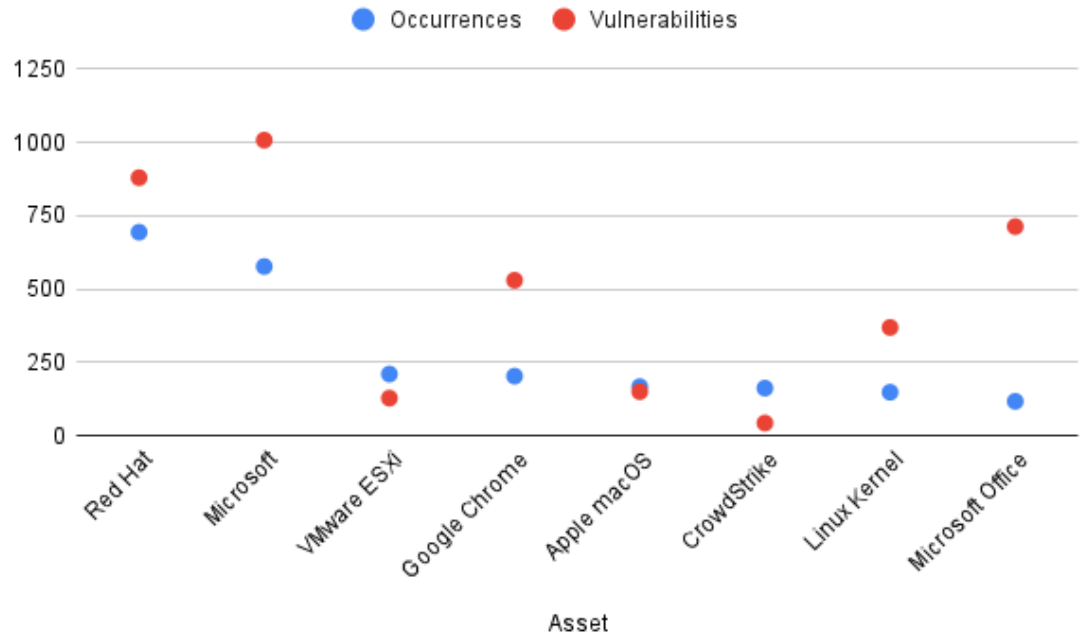


Figure 5.3: Occurrences of attack and different vulnerabilities found by asset - scatter chart



Microsoft Windows, Apple macOS, and the Linux Kernel — make up half of the most frequently targeted assets. This can be because operating systems serve as the backbone of computing, managing hardware resources and providing the foundational platform on which applications run. Their critical role and ubiquitous presence make them prime targets for attackers seeking to exploit vulnerabilities at the core of digital infrastructure.

Red Hat Enterprise Linux has more occurrences than Microsoft Windows, even though Windows has more documented vulnerabilities. This can be attributed to the fact that Red Hat maintains a very active feed of reported incidents and vulnerabilities. This continuous reporting contributes to its higher number of recorded occurrences, reflecting its proactive incident tracking rather than necessarily indicating greater exploitation or risk compared to Windows.

Apple macOS, on the other hand, exhibits significantly fewer occurrences and vulnerabilities compared to its competitors, Linux and Windows. This discrepancy can be explained by several factors. macOS is based on a Unix-like architecture, similar to Linux, but it incorporates additional security features unique to Apple, such as strict application sandboxing and a robust permission system. Furthermore, Apple tightly controls the macOS ecosystem, designing the operating system with a strong emphasis on security and privacy, which can make certain types of vulnerabilities more difficult to exploit. Additionally, macOS has a smaller market share, particularly in enterprise environments. Linux dominates server environments, while Windows is prevalent in both personal and enterprise desktops. Consequently, attackers tend to focus on the most widely used systems, where successful exploits have broader impacts, leading to macOS receiving less attention from threat actors.

Lastly, Microsoft Office demonstrates a significant disparity between its number of occurrences and vulnerabilities. This can be attributed to attackers often prioritizing the exploitation of vulnerabilities in the Windows operating system or commonly networked applications rather than directly targeting Office applications. OS-level exploits generally provide broader system access and greater control over a compromised device, making them more appealing to attackers. This focus reduces the frequency of direct attacks on Office, despite its high vulnerability count.

Figure 5.4: Threat actors by organization

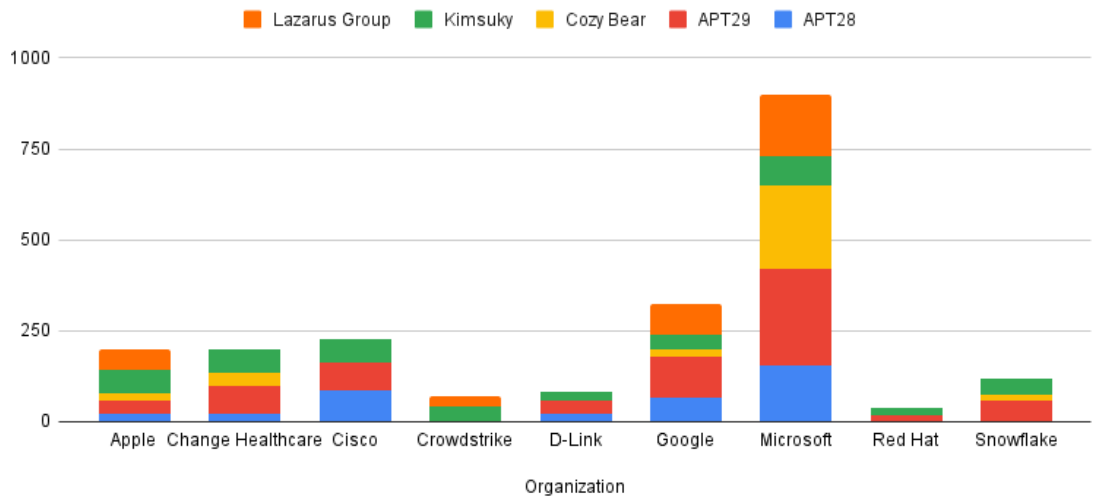
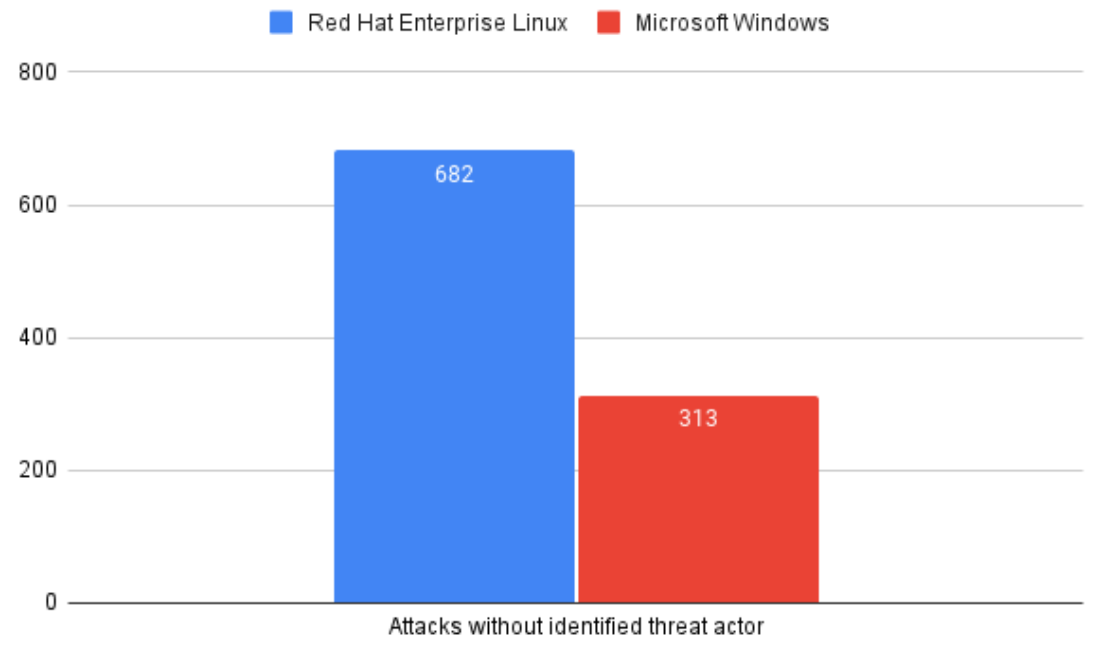


Figure 5.5: Attacks without identified threat actor



5.3 Threat Actors x Organizations

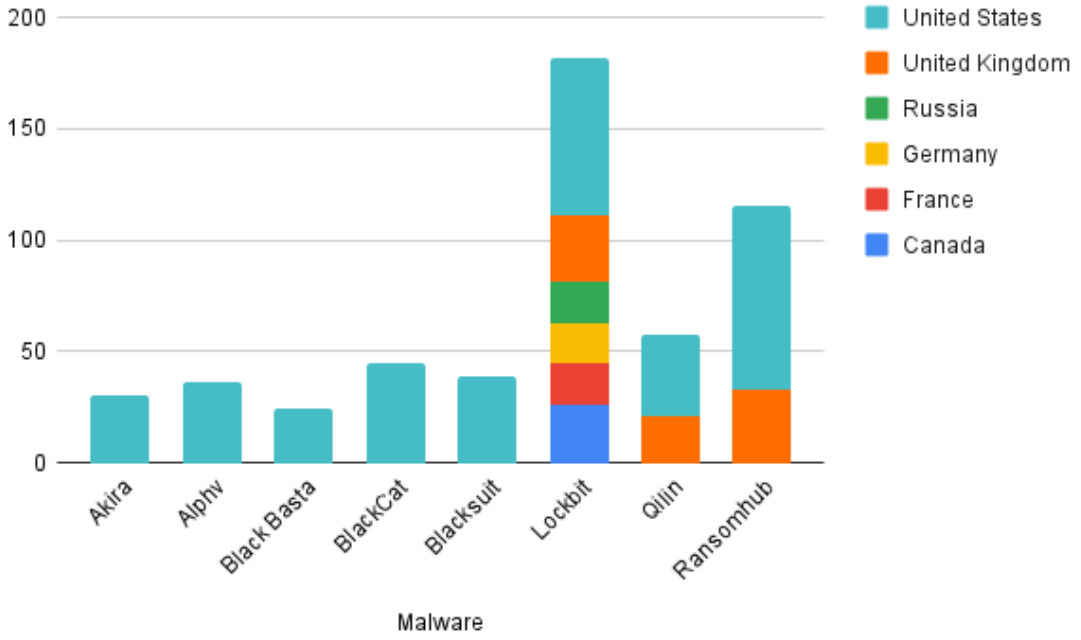
Figure 5.4 shows the number of attacks from APT28, APT29, Cozy Bear, Kimsuky and Lazarus Group threat actors that targeted the organizations Apple, Change Healthcare, Cisco, Crowdstrike, D-Link, Google, Microsoft, Red Hat, Snowflake.

Red Hat Enterprise Linux experiences a higher volume of attacks overall compared to Microsoft Windows. However, as shown in Figure 5.5, Red Hat Enterprise Linux has fewer attacks attributed to specific threat actors. One explanation is that 682 attacks on Red Hat were recorded without an identified threat actor, compared to 313 for Microsoft Windows. This disparity maybe reflects the differing nature of attacks targeting these systems. Red Hat Enterprise Linux, widely used in critical infrastructure and server environments, is often subjected to automated and opportunistic attacks leveraging generic exploits, which are less likely to be traced to specific actors. In contrast, Microsoft Windows is frequently targeted by advanced, highly focused attacks from well-known groups like APT28 and Lazarus Group, resulting in better attribution and documentation. Additionally, Red Hat's active reporting community may contribute to the discrepancy, as the abundance of contributors may lead to inconsistent identification of threat actors due to varying expertise and resources.

Apple, Google, and Microsoft stand out as the only companies targeted by all identified threat actors. This is likely because these tech giants are among the largest in the world, managing vast amounts of sensitive data and serving millions of users globally. Their prominence, influence, and the critical nature of their services make them highly attractive targets for threat actors aiming to compromise impactful systems.

On the other hand, Kimsuky is the only threat actor that targets all companies across the dataset. This can be attributed to Kimsuky's association with North Korean state-sponsored cyber activities, which are often focused on cyber-espionage campaigns. Kimsuky targets a wide range of industries—including technology, healthcare, telecommunications, and government—to gather intelligence and steal intellectual property. Known for its adaptability, Kimsuky frequently modifies its tactics and attack vectors. By targeting a broad spectrum of organizations, the group increases its chances of exploiting security weaknesses across different sectors, maximizing its likelihood of success regardless of the industry.

Figure 5.6: Malware by country



5.4 Malwares x Countries

Figure 5.6 shows the number of Akira, Alphv, Black Basta, BlackCat, Blacksuit, Lockbit, Qilin and Ransomhub, malware attacks that targeted companies from Canada, France, Germany, Russia, United Kingdom and United States.

All of the malware with the highest number of occurrences are ransomware, driven by their high profitability, accessibility through RaaS, and versatility in targeting a broad range of victims across industries. Modern tactics, such as double and triple extortion, amplify their impact by increasing pressure on victims and improving the success rate of attacks.

The United States is by far the most attacked country, standing out as the only nation targeted by all malware types in the dataset. This can be attributed to several factors. The U.S. is home to many of the world's largest and most influential companies, including technology giants, financial institutions, and defense contractors. These organizations hold valuable intellectual property and sensitive data, making them prime targets for attackers seeking financial gain or strategic advantage. Furthermore, attackers often focus on regions with the highest potential for financial returns. U.S. organizations are known to pay some of the highest ransoms in ransomware attacks, which further incentivizes

financially motivated cybercriminals to target the country.

LockBit is unique in that it is the only malware that targets all countries in the dataset. This is likely due to its highly effective RaaS model, which allows affiliates to launch attacks without requiring advanced technical expertise. Its use of robust and fast encryption techniques ensures that victims cannot recover their data without paying the ransom, while its advanced evasion capabilities, such as disabling antivirus software and exploiting remote desktop protocol vulnerabilities, make it difficult to detect and stop. Its widespread adoption by a large affiliate network, coupled with its adaptability for targeting diverse industries has solidified its status as one of the most widely used ransomware families globally. These factors, combined with its frequent updates and customizable features, have made LockBit a powerful and attractive tool for cybercriminals.

6 CONCLUSIONS

This bachelor's thesis presented an in-depth analysis of cybersecurity incidents using real-world data collected by Axur's Polaris platform. By using structured data analysis, correlation techniques, and descriptive analytics, this research has offered actionable insights into the patterns, trends, and vulnerabilities prevalent in today's cybersecurity landscape.

6.1 Summary of Key Findings

- **Cross-Sector Attack Analysis:** the analysis identified that attacks leveraging valid accounts (T1078) are the most pervasive, affecting all sectors examined. This underscores the universal vulnerability to credential-based exploits, which often evade traditional security measures.
- **Sector-Specific Threats:** certain attack vectors, such as the exploitation of public-facing applications (T1190), were found to be unique to the technology sector. This highlights the critical need for targeted cybersecurity measures that account for the specific vulnerabilities of different industries.
- **Asset Vulnerabilities:** a strong correlation ($r = 0.72$) was identified between the number of documented vulnerabilities in an asset and the frequency of attacks targeting it. Operating systems such as Red Hat Enterprise Linux and Microsoft Windows emerged as prime targets due to their foundational role in digital infrastructure.
- **Threat Actor Strategies:** prominent threat actors, including Lazarus Group, APT28, and Kimsuky, exhibited distinct targeting patterns, often focusing on high-value organizations like Apple, Google, and Microsoft. This reflects the strategic nature of their operations, aimed at maximizing impact.
- **Global Malware Trends:** the prevalence of ransomware across all analyzed malware types underscores the financial and operational risks posed by RaaS models.

6.2 Contributions to the Field

We can state three main outcomes of this bachelor's thesis as contributions to the field of cybersecurity. These contributions are supported by the analysis of the results obtained using the defined approach. First, by utilizing actual incident data, this study bridges the gap between theoretical models and practical applications, providing a grounded perspective on cybersecurity challenges. Second, the findings support the development of tailored strategies for industries, enabling organizations to allocate resources effectively and enhance their defenses against prevalent threats. Finally, the integration of structured data and statistical techniques offers a replicable framework for analyzing cybersecurity incidents, which can be adopted by other researchers and practitioners.

6.3 Future Work

Looking forward, several directions for expanding this research have been identified. One key opportunity lies in examining the unused portions of the dataset. This untapped data could provide additional insights into lesser-known aspects of the cybersecurity landscape, contributing to a more comprehensive understanding of existing and emerging threats. Another possibility involves extending the analysis beyond the current six-month timeframe. By incorporating data over longer periods, future studies could reveal trends that evolve over time and identify threats that may only emerge in the long term. Finally, integrating predictive analytics and machine learning models offers immense potential for enhancing the ability to forecast and mitigate future cybersecurity risks. These advanced techniques could enable organizations to proactively identify vulnerabilities and adapt to an ever-changing threat environment.

6.4 Final Remarks

In conclusion, this bachelor's thesis demonstrates the potential of data-driven approaches to enhance our understanding of cybersecurity threats. By uncovering patterns and correlations within real-world data, it provides actionable intelligence that can inform strategic decision-making and bolster organizational resilience. As the digital landscape continues to evolve, such research will remain vital in the ongoing effort to protect critical

assets and infrastructure from increasingly sophisticated cyber threats.

The insights gained from this research highlight the urgent need for proactive cybersecurity measures. Organizations must prioritize the adoption of adaptive strategies and continual investment in innovative solutions. By doing so, they can stay ahead of emerging threats and contribute to building a safer digital environment for all.

REFERENCES

Akamai Technologies. **What Is BlackCat Ransomware?** 2024. <https://www.akamai.com/glossary/what-is-blackcat-ransomware>.

ALSHARIF, M.; MISHRA, S.; ALSHEHRI, M. Impact of Human Vulnerabilities on Cybersecurity. **Computer Systems Science and Engineering**, v. 40, n. 3, p. 1153–1166, 2022.

Axur Content Team. **Axur introduces Polaris: the first AI-powered Threat Intel Analyst**. 2024. <https://blog.axur.com/en-us/axur-introduces-polaris-the-first-ai-powered-threat-intel-analyst>.

BARRY, C. **BlackSuit ransomware: 8 years, 6 names, 1 cybercrime syndicate**. 2024. <https://blog.barracuda.com/2024/10/29/blacksuit-ransomware--8-years--6-names--1-cybercrime-syndicate>.

BLEIH, A. **Qilin Ransomware: Get the 2024 Lowdown**. 2024. <https://cyberint.com/blog/research/qilin-ransomware/>.

Check Point Team. **June 2024's Most Wanted Malware: RansomHub Takes Top Spot as Most Prevalent Ransomware Group in Wake of LockBit3 Decline**. 2024. <https://blog.checkpoint.com/research/june-2024s-most-wanted-malware-ransomhub-takes-top-spot-as-most-prevalent-ransomware-group-in-wa
>.

CIUMACENCO DENIS PLEȘCA, I. P. Cybersecurity: from past threats to present solutions. In: **Technical Scientific Conference of Undergraduate, Master and PhD Students**. [S.l.]: Universitatea Tehnică a Moldovei, 2024. v. 2, p. 736–740.

DURAIBI, S.; KAUR, C.; PAWAR, A. Cyber extortion unveiled: The evolution, tactics, challenges, and future of ransomware. In: **2023 International Conference on Computational Science and Computational Intelligence (CSCI)**. [S.l.: s.n.], 2023. p. 861–867.

Flashpoint Intel Team. **LockBit Ransomware: Inside the World's Most Active Ransomware Group [Updated]**. 2023. <https://flashpoint.io/blog/lockbit/>.

Florian Zandt. **How Much Money Is Lost to Cybercrime?** 2024. <https://www.statista.com/chart/32341/worldwide-reported-losses-connected-to-cybercrime/>.

FRANCO, M. et al. Secadvisor: A tool for cybersecurity planning using economic models. In: . [S.l.: s.n.], 2024. p. 554–569.

FRANCO, M. F.; GRANVILLE, L. Z.; STILLER, B. CyberTEA: A Technical and Economic Approach for Cybersecurity Planning and Investment. In: **36th IEEE/IFIP Network Operations and Management Symposium (NOMS 2023)**. Miami, USA: [s.n.], 2023. p. 1–6.

FRANCO, M. F. et al. **RCVaR: an Economic Approach to Estimate Cyberattacks Costs using Data from Industry Reports**. 2023. Available from Internet: <https://arxiv.org/abs/2307.11140>.

FRANCO, M. F. et al. SecRiskAI: a Machine Learning-Based Approach for Cybersecurity Risk Prediction in Businesses. In: **24th IEEE International Conference on Business Informatics (CBI 2022)**. Amsterdan, Netherlands: IEEE, 2022. p. 1–10.

FURHAD, M.; NOWROZY, R.; SARKER, I. AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling and Research Directions. **SN Computer Science**, v. 2, May 2021.

GORDON, L.; LOEB, M.; ZHOU, L. Investing in Cybersecurity: Insights from the Gordon-Loeb Model. **Journal of Information Security**, v. 07, p. 49–59, 01 2016.

G'UZOROVICH, E. A. The Evolution of Cybersecurity: Safeguarding the Digital Era. **Synergy: Cross-Disciplinary Journal of Digital Investigation (2995-4827)**, v. 2, n. 4, p. 111–117, apr 2024.

JIN, X. et al. Factors Affecting Reputational Damage to Organisations Due to Cyberattacks. **Informatics**, v. 9, n. 1, 2022. ISSN 2227-9709.

LAHCEN R.A., C. B. M. R. e. a. M. Review and insight on the behavioral aspects of cybersecurity. **Cybersecurity**, v. 3, n. 1, p. 10, April 2020. ISSN 2523-3246.

MICHAEL, D. **Application Layer Protocol**. 2024. <https://attack.mitre.org/techniques/T1071/>.

NUNES, J. et al. Sim-ciber: Uma solução baseada em simulações probabilísticas para quantificação de riscos e impactos de ciberataques utilizando relatórios estatísticos. In: . [S.l.: s.n.], 2024. p. 570–585.

PALIWAL, A. **Black Basta Ransomware: What You Need to Know**. 2024. <https://blog.qualys.com/vulnerabilities-threat-research/2024/09/19/black-basta-ransomware-what-you-need-to-know>.

PRADHAN, A. **Threat Brief: Understanding Akira Ransomware**. 2024. <https://blog.qualys.com/vulnerabilities-threat-research/2024/10/02/threat-brief-understanding-akira-ransomware>.

RAVICH CARDINALOPS; OHAD ZAIDENBERG, o. P. W. S. C. C. O. L. I. L. **Phishing**. 2024. <https://attack.mitre.org/techniques/T1566/>.

SATYAPANICH, T.; FERRARO, F.; FININ, T. CASIE: Extracting Cybersecurity Event Information from Text. **Proceedings of the AAAI Conference on Artificial Intelligence**, v. 34, n. 05, p. 8749–8757, April 2020.

SOCRadar Research. **APT Profile: Who is Lazarus Group?** 2021. <https://socradar.io/apt-profile-who-is-lazarus-group/>.

SOCRadar Research. **APT Profile: Cozy Bear / APT29**. 2023. <https://socradar.io/apt-profile-cozy-bear-apt29/>.

SOCRadar Research. **APT Profile: Kimsuky**. 2023. <https://socradar.io/apt-profile-kimsuky/>.

SOCRadar Research. **APT28 Deploys 'GooseEgg' in Attacks Exploiting the Windows Print Spooler Vulnerability, CVE-2022-38028**. 2024. <https://socradar.io/apt28-deploys-gooseegg-in-attacks-exploiting-the-windows-print-spooler-vulnerability-cve-2022-38028/>.

STERNSTEIN STERN SECURITY; MARK WEE; MENACHEM GOLDSTEIN; NETSKOPE; PRAETORIAN; PRASAD SOMASAMUDRAM, M. S. S. M. S. U. F. M. Y. W. A. D. R. T. J. **Valid Accounts**. 2024. <https://attack.mitre.org/techniques/T1078/>.

STROM, B. E. et al. Mitre att&ck: Design and philosophy. In: **Technical report**. [S.l.]: The MITRE Corporation, 2018.

VLADESCU, C. et al. What are the latest cybersecurity trends? a case study grounded in language models. In: **2021 23rd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)**. [S.l.: s.n.], 2021. p. 140–146.

WEIZMAN, A. D. R. T. P. Y. **Exploit Public-Facing Application**. 2024. <https://attack.mitre.org/techniques/T1190/>.