A Solution for Securing the Information Environment Inspired by Living Organisms and Biology

Seyedali Pourmoafi

School of Physics, Engineering and Computer (SPEC)

University of Hertfordshire

Submitted to the University of Hertfordshire in partial fulfilment of the requirement of the degree of Doctor of Philosophy

April 2023

Acknowledgement

'A journey of a thousand miles begins with a single step.' My PhD is no exception. There are many dreams you see but only a few become reality. Many actors and situations play a major role. Therefore, this dream cannot be realised without giving due credit and acknowledgement. The prime lesson I have learned in this study is that *'Success is a mark of privilege'*.

I am grateful to the many people who supported me in various ways during my Ph.D. journey. First of all, I praise and thank God for his countless blessings and for giving me the strength to accomplish this research during tough times.

This thesis will always be incomplete without expressing gratitude to my supervisor, Dr. Stilianos Vidalis. I would like to express my sincere appreciation to him for his patience, kindness, and invaluable support throughout the entire process. His guidance, tutelage, and motivation were instrumental to the successful completion of this endeavor, and I am forever indebted to him.

Finally, I want to thank my lovely wife, Parisa, and family for their unwavering patience and support over the years. Without their encouragement, love, and understanding, none of this would have been possible. Their presence in my life has been a constant source of inspiration and motivation, and I am deeply grateful for their unending support.

Abstract

Cyber-attacks have become increasingly frequent and severe in recent years, posing a significant threat to both economic and safety domains. In order to mitigate these threats, it is crucial to have a thorough understanding of the various types of attacks, including their behaviour and properties. Bio-cyber operation is a new field of research that draws inspiration from the human immune system, which has found solutions to problems that cybersecurity professionals have been struggling with for decades. By studying the human immune system, cybersecurity can teach valuable lessons on how to detect and deter attacks. To address the issue of sensitive information or data leakage, a "cyber immune" technology can be employed to detect unknown cyber-attacks and provide a powerful defence mechanism. This thesis proposes a methodology for predicting potential vulnerabilities by utilizing natural language processing (NLP) and analysing "Common Vulnerabilities and Exposure" (CVE) reports. Additionally, the implementation of Gradient Sensitivity (GS) and Gradient Sensitivity Input (GI) techniques into our framework represents a significant leap forward in enhancing the model's explainability. The incorporation of GS and GI techniques into our framework means that our models meet the critical needs of cybersecurity applications today. They strike a balance between advanced predictive capabilities and the equally important need for transparency and interpretability. This balance is what will keep driving cybersecurity efforts forward, ensuring that we're not just creating models that can predict threats but also fostering a cybersecurity approach that can effectively act on these predictions.

Finaly, by examining the field of human biology, we can gain significant insight into the bio-cybersecurity domain, which can be applied to developing more effective cybersecurity measures. Our results indicate that by using NLP we can achieve high percentage of accuracy to predict the possible vulnerability (more than 94%) as well as find the interrelationship between different vulnerabilities (more than 53% similarity).

Contents

List of Figures

List of Tables

Abbreviations

ABCD	Ant-Based Cyber Defence.
ACLs	Access Control Lists
AIS	Adaptive Immune System
APT	Advanced Persistent Threats
AI	Artificial Intelligence
BERT	Encoder Representations from Transformers
CAPEC	Common Attack Pattern Enumeration and Classification
CNA	Computer Network Attack
CND	Computer Network Defence
CNE	Computer Network Exploitation
CNO	Computer Network Operations
CSRF	Cross-Site Request Forgery
CVSS	Common Vulnerability Scoring System
DoS	Denial of Service
DDoS	Distributed Denial of Service
E.C.	Educational Complexity
FAIR	FAIR (Factor Analysis of Information Risk)
GIE	Global Information Environment
GS	Gradient Sensitivity
GI	Gradient Sensitivity Input
ICCP	Inter-Control Center Communication Protocol
IDS	Intrusion Detection System
IS	Immune System
InfoSec	Information Security
ISS	Information Security System

ITB	Invitation to Bid
KNN	K-Nearest Neighbour
ML	Machine Learning
MITM	Man-in-the-Middle
NIH	National Institutes of Health
NSF	National Science Foundation
NVD	National Vulnerability Database
NLI	Natural Language Interface
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
OPSEC	Operational Security
RFP	Request for Proposal
RFQ	Request for Quotation
SIEM	Security Information and Event Management
SAO	Situational Awareness Operations
TTPs	Tactics, Techniques, and Procedures
IoT	The Internet of Things
TTE	Time to Exploit
CVE	Vulnerabilities and Exposures
VC	Vulnerability Complexity
ZDI	Zero-Day Initiative

Chapter 1

Introduction

1.1 Overview

Cybersecurity, also known as information technology security or computer security, involves protecting networks and computer systems from unauthorised access, data breaches, software and hardware damage, and service disruptions [1]. The definition of security emphasizes the importance of protecting assets, which requires knowing their value. In today's interconnected world, where mergers and acquisitions are common, securing information and data is a complex task that requires expertise in cybersecurity [2]. Cybersecurity strategies aim to protect a user or organization's cyber environment through a set of techniques and methods that safeguard the integrity of data, networks, and programs. This field is becoming increasingly important due to the growing reliance on computer devices and systems, including smartphones, television, and the Internet of Things (IoT) [3].

The emergence of technology and the internet has revolutionized the world in many ways, but it has also introduced new and challenging circumstances. As security measures grow, so too does the hacking world. There are different perspectives on the issue of cybersecurity. Some cloud computing companies argue that their use of the latest encryption technology makes them extremely secure, while cybersecurity experts maintain that these types of attacks can be predicted or detected with the right context [4].

1.1.1 What is Cyber Security?

Cybersecurity, also known as computer security or information technology security, is the practice

of protecting computer systems, networks, devices, and data from theft, damage, or unauthorised access [1]. Cybersecurity measures are put in place to prevent unauthorised access to sensitive information, such as personal and financial data, trade secrets, and government or military information. These measures include hardware and software-based protections, such as firewalls, antivirus software, and encryption, as well as user training and education on safe online practices [2]. Cybersecurity is becoming increasingly important in our digital age, as more of our personal and professional lives are conducted online, and as the number and sophistication of cyber-attacks continue to grow [3]. Cybersecurity refers to the application of technologies, processes, and controls to protect systems or devices connected to the internet, such as networks, data, software, or hardware, from possible cyber vulnerabilities or attacks [4]. The main aim of cybersecurity is to reduce the risk of cyber-attacks and vulnerabilities, and to protect against the unauthorised exploitation of networks, systems, and technologies [5].

1.1.2 Problem Domain

The research hypothesis is that a threat assessment methodology can be proposed to predict, mutate, and address changes in the chaotic and ever-changing environment of cybersecurity risk. However, there is a lack of consensus on the definitions of concepts such as cybersecurity, cyberspace, and other related terms, which can lead to confusion and misinterpretation. The chapter provides an overview of the current definition of cybersecurity in government, industry, academia, and professional contexts. Some definitions of cybersecurity focus on network and communication infrastructures, while others emphasize the importance of users in cybersecurity [9]. This study considers cybersecurity as a continuum of technologies and innovations for ensuring the security of data, information, and networking technologies. The research explores how natural phenomena in complex systems, such as The Human Immune System (HIS), can be used as mechanisms or strategies for adaptive mitigation in complicated cyber environments. The main aim of this research is to help cybersecurity analysts and experts gain insight into the development of attack prediction and detection models from an initial set of scenarios. This process involves analysing the symptoms and behaviour of attacks to deduce a list of possible vulnerabilities (CVE), and then applying different algorithms to learn which scenario corresponds to a successful cyberattack.

This study contributes to the field of cybersecurity by highlighting the critical importance of bio-cybersecurity in the development of information security system (ISS) processes. The main

contribution is the proposal of an objective framework, modelling, developing, and assessing the vulnerability interrelationship and security approach. This framework improves our understanding of the nature of these phenomena and provides a better theoretical and practical basis for data and information systems security. The study focuses on the role of the human immune system in ISS and cybersecurity risk assessment. The findings emphasize the significance of bio-cybersecurity in modern data and information security contexts. The results of this study are expected to have a positive impact on cybersecurity by improving the effectiveness of system security implementation and design.

1.2 Why is Cyber Security Important?

Cybersecurity is crucial in safeguarding sensitive data, information systems, and networks against a wide array of cyber threats. These threats pose significant risks to personal, organizational, and national security. Recent government publications and scholarly articles have highlighted the evolving landscape of cyber threats alongside the strategic importance of cybersecurity measures. According to the Cybersecurity and Infrastructure Security Agency (CISA) and recent studies, effective cybersecurity operations are pivotal in protecting against and mitigating the impact of cyber-attacks, which increasingly target government, military, and commercial assets. Some of the prevalent cyber threats identified in the latest literature include:

- Malware: Defined as malicious software, malware encompasses viruses, ransomware, worms, and spyware. It represents a significant threat as it can be activated by users inadvertently clicking on malicious links or attachments, leading to unauthorised software installations that compromise security [171].
- Emotet: Highlighted by CISA as a particularly destructive modular banking Trojan, Emotet serves mainly as a dropper or downloader for other banking Trojans, making it one of the costliest malware threats to date [172].
- Denial of Service (DoS) and Distributed Denial of Service (DDoS): These attacks overload a network or system, rendering it incapable of responding to legitimate requests. DDoS attacks are particularly challenging as they originate from multiple sources, complicating mitigation efforts [173]
- Man-in-the-Middle (MITM) Attacks: Occurring when attackers intercept communications between two parties, MITM attacks can lead to significant data breaches, especially when

users connect to unsecured public Wi-Fi networks [174].

- Phishing: Utilizing deceptive communications, typically via email, phishing attacks trick recipients into divulging sensitive information, such as login credentials and financial data [175]
- SQL Injection: This type of attack involves inserting malicious code into servers that use SQL, leading to unauthorised data access or system compromise [176].
- Password Attacks: By obtaining or cracking passwords, attackers can access extensive amounts of confidential information. Techniques include social engineering, which manipulates individuals into violating security protocols [177].

1.3 Research Aim and Objectives

The primary aim of this research is to develop a vulnerability interrelationship model that supports the analysis and reasoning of the critical characteristics of the human immune system within the context of cyber operations in organizations. This research seeks to elucidate the complex interplay between human immune system models and cybersecurity, addressing the critical need for robust defence mechanisms in today's modern technology. To accomplish this aim, this research is grounded in a framework that integrates a variety of theories and methods to explore the applicability and effectiveness of the proposed model. The overarching objective is to enhance organizational resilience to cyber threats through a deeper understanding of the human immune system's principles applied to cybersecurity.

Specific objectives designed to achieve this aim include:

- To examine and synthesize the existing literature on human immune system models and their relevance to cybersecurity, establishing a theoretical foundation for the study.
- To construct a comprehensive vulnerability interrelationship model that embodies the essential features of the human immune system relevant to cyber operations, aiming for a novel approach to cybersecurity.
- To assess the model's utility in identifying and counteracting cyber threats, providing empirical evidence of its effectiveness.
- To explore the dynamics between risk and investment in cybersecurity projects that implement the vulnerability interrelationship model, offering insights into strategic decision-making.

- To offer actionable recommendations for organizations on integrating the vulnerability interrelationship model into their cybersecurity strategies and investment planning, emphasizing practical applications.

1.4 Research Contribution

For many years, the study of human and nature's factors in cyber or system security has been included as an independent and unique discipline known as Bio-cybersecurity. The term "bio-cybersecurity" is a relatively new concept that has emerged in response to the growing threat of cyber-attacks on biological research and development. The first references to bio-cybersecurity can be traced back to the early 2000s, when concerns were raised about the potential for cyber-attacks on critical infrastructure, including biological research facilities. The National Science Foundation (NSF) in the United States recognized the importance of bio-cybersecurity in the early 2000s and began funding research into this area. In 2003, the NSF established the Biosecurity and Bioterrorism: Biodefence Strategy, Practice, and Science Journal, which included articles and research papers on bio-cybersecurity [163].

In 2008, the US National Institutes of Health (NIH) issued guidelines on biosecurity, which included recommendations on protecting biological research from cyber-attacks. The guidelines emphasized the need for security measures to protect against unauthorised access, theft, or destruction of sensitive data. Since then, bio-cybersecurity has become an increasingly important area of focus for governments, research institutions, and the private sector. The potential consequences of a successful cyber-attack on biological research or facilities are significant, and the need for robust security measures to prevent such attacks is paramount [164].

This field of study focuses on the nature and characteristics of human relationships from a unified perspective of technology, science, design, engineering, and management of human-compatible systems. It involves natural and artificial products, processes, and living environments. Thus, the relationship between human factors and information security systems is not a new phenomenon.

Inspired by the human immune system, numerous studies have been conducted in computer science, technology, information systems, and information security. Although no research has specifically addressed the role of vulnerability interrelationship in cybersecurity to predict possible attacks by inspiring the human immune system, recent advancements in Information Security System (ISS) procedures have not provided effective security for protecting information and data

[2][3][10][11]. The main motivation behind this research is to identify the interrelationship between vulnerabilities and analyse them using insights from the human immune system. The current literature lacks sufficient and appropriate methodologies and models for analysing cyber risks, highlighting the need for more comprehensive approaches to analysing possible threats in the cyber environment. The contribution of this research is a novel methodology and model that utilizes Natural Language Processing (NLP) to predict vulnerabilities that threaten our systems or devices. The model utilizes concepts from Machine Learning (ML) to analyse Common Vulnerabilities and Exposures (CVE). To protect information systems, this research has illustrated the need for good knowledge and insight into the behaviour of different cyber-attacks, inspiring the human immune system behaviour, and analysing it to create a model for information security. This model can help us decrease the possible cyber threats.

1.5 Research Methodology

The methodology of this study comprises different stages of Investigation, Analysis, and Proposed technique. Initially, we identified the challenges involved in securing the information environment using traditional cybersecurity methods, and we explored the potential benefits of applying bio-cybersecurity approaches. To gain a deeper understanding, we conducted a comprehensive literature review on biology and cybersecurity, as well as natural language processing and machine learning. After establishing the link between biology and cybersecurity, we developed our research design. We introduced a methodology and framework called "Bio Inspired Cybersecurity and Bio-Inspired Cybersecurity Framework" to describe the different phases of our approach. For the technical aspect of our work, we collected and prepared data relevant to our research design. The research design was meticulously developed to align with the study's objectives, drawing inspiration from the resilience and adaptability of living organisms' immune systems. This so-called bio-inspired approach guided the selection of research methods, focusing on the natural language processing (NLP) analysis of Common Vulnerabilities and Exposures (CVE) reports and the implementation of machine learning techniques like Gradient Sensitivity (GS) and Gradient Sensitivity Input (GI). Specifically, we used the CVE dataset, which we pre-processed and formatted for training and testing our models. The overarching methodology combines NLP techniques to parse and analyse CVE reports, identifying patterns and predicting potential vulnerabilities. This quantitative method is complemented by qualitative insights gathered from

expert interviews, providing depth and context to the data-driven findings. To evaluate the effectiveness of our proposed solution, we used a methodology for training and testing models, along with specific performance metrics. We then developed and trained a novel model incorporating bio-cybersecurity security, NLP, and ML techniques, which allowed us to accurately predict vulnerabilities and analyse their interrelationships. Additionally, we generated a vulnerability tree to aid in vulnerability management. Finally, we implemented our work using different machine learning algorithms, such as KNN and NLTK, as well as a popular NLP model named BERT. Through this approach, we were able to achieve high accuracy predictions and demonstrate the effectiveness of our bio-cybersecurity solution for securing the information environment.

The bio-inspired approach is supported by literature suggesting the potential for biological systems to inform cybersecurity strategies [164]. Similarly, each method of this study's choice is thoroughly justified within the context of improving systems protection against cyberattacks by considering the vulnerability interrelationship. The bio-inspired approach is supported by literature suggesting the potential for biological systems to inform cybersecurity strategies. Similarly, the selection of NLP and machine learning techniques have been utilized based on their ability to process and analyse complex data sets, identifying patterns through conventional analysis.

1.6 Research Publications

- Pourmoafi, S. and Vidalis, S., 2021, June. Bio-Cyber Operations Inspired by the Human Immune System. In *European Conference on Cyber Warfare and Security* (pp. 534-XIV). Academic Conferences International Limited.
- Seyedali Pourmoafi, Stilianos Vidalis, Vulnerability Interrelationship Inspire Bio Cybersecurity, *JISET - International Journal of Innovative Science, Engineering & Technology,* Vol. 09 Issue 09, September 2022

Chapter 2

Literature Review

2.1 Overview

This chapter provides insight into more comprehensive understanding of cybersecurity vulnerabilities to enhance our knowledge in this field. We divided this chapter into two main parts:

- Cybersecurity Vulnerabilities Analysis
- Cyber-Immune System (Bio Cybersecurity)

2.2 Cybersecurity Vulnerabilities Analysis

Cybersecurity vulnerabilities refer to weaknesses in a computer system, network, or application that can be exploited by attackers to gain unauthorised access or cause damage to the system or its data. Analysing these vulnerabilities is a critical part of cybersecurity risk management, aimed at identifying, categorizing, prioritizing, and addressing these weaknesses to protect information assets from threats [178]. Here's an overview of the process and key considerations in cybersecurity vulnerabilities analysis:

2.2.1 Identification

The first step in the analysis is to identify all potential vulnerabilities within a system. This process lays the groundwork for a comprehensive cybersecurity strategy, as it allows organizations to understand their exposure to potential threats. Identifying vulnerabilities can be achieved through various means, including:

- **Automated Scanning Tools:** Automated scanning tools play a critical role in detecting vulnerabilities by systematically probing systems, networks, and software for known security weaknesses. These tools maintain databases of known vulnerabilities, often sourced from repositories such as the National Vulnerability Database (NVD) and compare them against the system's components. They are effective in identifying a wide array of vulnerabilities, including missing patches, outdated software versions, and configuration issues. Commonly used tools include Nessus, OpenVAS, and Qualys Guard.
- **Penetration Testing:** Penetration testing, often referred to as ethical hacking, involves simulating cyber-attacks on a system to uncover vulnerabilities. This approach goes beyond automated scanning by employing the creativity and expertise of ethical hackers. They use tactics, techniques, and procedures (TTPs) similar to those used by malicious hackers to identify weaknesses that automated tools may overlook. Penetration testing can be black-box (without prior knowledge of the system), white-box (with full knowledge of the system), or Gray-box (with partial knowledge of the system), providing a realistic assessment of an organization's security posture.
- **Code Review:** A Code review is a detailed analysis of an application's source code to identify security flaws. This process can be manual or automated. Automated tools such as SonarQube or Checkmarx can rapidly analyse large codebases and detect common vulnerabilities like SQL injection, cross-site scripting (XSS), and buffer overflows. Manual reviews, on the other hand, provide a deeper understanding by leveraging the expertise of security professionals to identify complex logic flaws or subtle security issues. Incorporating secure coding practices during development and conducting regular code reviews can significantly reduce vulnerabilities.
- **Configuration Management and Analysis:** Configuration management ensures that systems are configured securely and in accordance with best practices. Misconfigurations are a common source of vulnerabilities, often arising from default settings, unnecessary services, or poorly implemented access controls. Tools like Ansible, Puppet, and Chef can automate configuration management to maintain consistent and secure configurations across systems. Configuration analysis involves auditing system settings and configurations against established security baselines or industry standards like the CIS Benchmarks or NIST guidelines.

2.2.2 Classification

Once vulnerabilities are identified, they are classified to prioritize remediation efforts and allocate resources effectively. This classification is based on their nature and potential impact, which helps organizations understand and address the most serious threats first. The common classifications include:

- **Severity Levels:** Vulnerabilities are assigned severity levels such as critical, high, medium, and low. These levels are determined by various factors including the ease of exploitation, the complexity of the attack required, and the potential impact on the system or data integrity. For instance, a critical vulnerability might allow remote code execution with minimal user interaction, leading to full control of the affected system, while a low severity vulnerability might only allow limited access and require significant user interaction. The Common Vulnerability Scoring System (CVSS) provides a standardized framework to rate the severity of security vulnerabilities and helps in this classification. Such as critical, high, medium, and low, often determined by factors like ease of exploitation and potential impact [61].
- **Type of Vulnerability:** Vulnerabilities are also classified by type, which describes the nature of the weakness in the system. For example, SQL injection, cross-site scripting (XSS), buffer overflows, or misconfigurations.
- **Affected Components:** The affected components of vulnerabilities are identified to specify whether they impact hardware, software, network configurations, or user practices. This classification helps in understanding the scope and potential reach of the vulnerability. For instance, identifying whether vulnerabilities affect hardware, software, network configurations, or user practices [62].

2.2.3 Evaluation

The evaluation stage in vulnerability management is a critical process where each identified vulnerability is assessed to understand its potential impact on the organization's assets. This step involves several key components:

- **Exploitability:** This refers to how easy it is for an attacker to exploit a given vulnerability. Several factors contribute to exploitability, including the complexity of the exploit, the level of access required, and whether an exploit is publicly available. Vulnerabilities for

which reliable, easy-to-use exploits exist in public databases such as Exploit-DB or that are included in exploitation frameworks like Metasploit are considered highly exploitable. Other factors influencing exploitability include the need for user interaction (such as clicking a link or downloading a file) and the availability of information or tools that facilitate the attack [53].

- **Scope of Impact:** Another vital classification concerns the scope of the impact, which determines how much of the system can be compromised if the vulnerability is exploited. This includes local impacts that affect only a single user's data or settings, and network impacts that could compromise entire systems or networks.
- **Risk Assessment:** Risk assessment combines the factors of exploitability and scope of impact to estimate the overall risk posed by each vulnerability [52].

2.2.4 Mitigation

The final step in vulnerability management is addressing the identified vulnerabilities to mitigate risks effectively. This step is crucial as it involves implementing specific actions designed to prevent attackers from exploiting known weaknesses in the system. Various methods are employed to achieve this, including:

- **Patching:** Patching is one of the most direct and effective methods for addressing vulnerabilities. This involves applying updates or patches released by software vendors to fix security flaws. Regular patch management ensures that all systems and applications are up-to-date and protected against known vulnerabilities. Organizations should prioritize patches based on the severity and exploitability of the vulnerabilities they address, applying critical patches immediately to prevent potential breaches [61].
- **Configuration Changes:** Many vulnerabilities arise from improper or insecure configurations of systems and software. Adjusting these settings according to security best practices can significantly enhance system security. This includes disabling unnecessary services, enforcing the use of strong passwords, setting up appropriate user permissions, and enabling security features that are disabled by default. Regularly reviewing and updating configurations to align with the latest security guidelines helps maintain a robust defence against potential attacks [74].
- **Network Segmentation:** Network segmentation involves dividing a network into smaller, controlled segments to limit the spread of attacks and reduce the attack surface. By isolating

critical systems and data, organizations can ensure that an attack on one segment does not compromise the entire network. Segmentation is particularly useful in larger organizations where different departments may have varying security needs and exposure levels. Effective segmentation is often accompanied by strict access controls and monitoring to ensure that traffic between segments is legitimate and does not pose a security risk [74].

- **Education and Awareness:** Human error is a significant factor in many security breaches. Educating and raising awareness among staff about cybersecurity risks and best practices is essential. Regular training sessions should be conducted to inform employees about the latest security threats, such as phishing, social engineering, and ransomware attacks. Employees should also be trained on the importance of using strong passwords, recognizing suspicious emails or links, and safely handling sensitive data. An informed and vigilant workforce can act as the first line of defence against cyber threats [61].

2.2.5 Tools and Frameworks

The cybersecurity vulnerabilities analysis process is supported by a variety of tools and frameworks designed to enhance the detection, assessment, and mitigation of security risks. These tools and frameworks range from automated scanners to comprehensive databases and scoring systems. Here is an overview of some key resources used in the field [53]:

- **Vulnerability Scanners:** Vulnerability scanners are automated tools designed to discover vulnerabilities in networks, systems, and applications. They scan for known security issues and generate reports detailing vulnerabilities that need to be addressed. Tools like Nessus, OpenVAS, and Qualys can automatically detect vulnerabilities [74].
- **Security Information and Event Management (SIEM):** SIEM systems provide real-time analysis of security alerts generated by applications and network hardware. These systems collect and aggregate log data produced throughout an organization’s technology infrastructure, from host systems and applications to network and security devices such as firewalls and antivirus filters. SIEM solutions help in detecting, understanding, and responding to security incidents [194].
- **Common Vulnerabilities and Exposures (CVE):** The CVE system provides a publicly available catalog of disclosed cybersecurity vulnerabilities and exposures. Each entry in the CVE List includes an identification number, a description, and at least one public reference. CVE is used in numerous cybersecurity products and services from around the

world, including vulnerability management tools. Having a standardized identifier for each vulnerability allows security professionals to quickly and accurately assess the risks associated with specific issues.

- **Common Vulnerability Scoring System (CVSS):** The CVSS is a standardized framework for rating the severity of security vulnerabilities. Scores are calculated based on a formula that evaluates various aspects of the vulnerability, such as exploitability, the complexity of the attack needed to exploit it, and the impact on confidentiality, integrity, and availability. CVSS scores provide a way to prioritize security responses and resources according to the potential severity of the threats, with vulnerabilities rated on a scale from 0 to 10 [74].

Furthermore, creating a comparative table can help us to analyse various vulnerability assessment techniques and tools. Setting up this table involves considering a range of factors, including their methodologies, strengths, and weaknesses. Below, I have outlined a simplified Table (2.1) that compares some common vulnerability assessment techniques and tools. This table also touches upon how these methods relate to current research in cybersecurity. Please note that the field is dynamic, with tools and techniques continuously evolving. New research can introduce innovative methods or improve existing ones, impacting their effectiveness and applicability.

Table 2.1 A Comparison of common vulnerability assessment techniques and tools

Technique/Tool	Pros	Cons	Relation to Current Research
Static Analysis Tools	Can identify vulnerabilities early in the development cycle. Does not require code execution.	May produce false positives. Limited by the complexity of code analysis.	Research focuses on reducing false positives and enhancing the ability to analyse complex code structures.
Dynamic Analysis Tools	Identifies vulnerabilities in running applications. Can uncover runtime issues that static analysis misses	Can miss issues not triggered during the analysis. Requires a test environment.	Ongoing research aims to improve coverage and accuracy, including the identification of sophisticated runtime vulnerabilities.

Web Application Scanners	Automated scanning of web applications for common vulnerabilities. Can be integrated into the CI/CD pipeline.	May not identify business logic vulnerabilities. False, positives and negatives	Advances aim at better understanding of web application logic and reducing false positives/negatives.
Network Scanners	Identifies open ports and vulnerable services on network devices. Can scan large networks quickly.	Limited to network-level vulnerabilities May not provide context on the impact	Research focuses on integrating AI to predict vulnerabilities based on network behaviour and historical data.
Penetration Testing Tools	Simulates real-world attacks. Provides detailed insights into actual exploitability	Time-consuming and requires skilled personnel. Potentially disruptive to operations.	Research is directed towards automation of penetration tests and minimizing operational disruptions.
Configuration Management Tools	Ensures systems are configured securely. Can automate the compliance checking process	Configuration drift can occur over time. May not cover all security scenarios	Current research includes developing more adaptive and intelligent configuration management solutions.

2.3 What is Vulnerability Assessment?

A vulnerability assessment is an organized examination of potential security flaws in an information system. It assesses whether the system is vulnerable to any known weaknesses, categorizes the severity of those weaknesses, and advises on necessary remedial or precautionary measures [61].

There are various categories of vulnerability assessments, including:

- **Host assessment**: Host assessment refers to the process of assessing the security vulnerabilities of critical servers within an organization's IT infrastructure. These servers may be at risk of cyber-attacks if they are not properly tested or if they are not generated from a tested machine image. The host assessment identifies any weaknesses in the servers' configurations, software, or hardware, and recommends remediation or mitigation steps to address them [62].
- **Network and wireless assessment**: Network and wireless assessment refers to the evaluation of network security policies and practices to prevent unauthorised access to both private and public networks, as well as network-accessible resources. The assessment identifies potential vulnerabilities and risks associated with network and wireless configurations, such as weak encryption, unsecured Wi-Fi access points, and outdated software versions. It also provides recommendations to mitigate or remediate any identified vulnerabilities or risks [53].
- **Database assessment**: Database assessment involves the evaluation of databases or big data systems to identify vulnerabilities and misconfigurations. It also involves the identification of rogue databases or insecure development/testing environments and the classification of sensitive data across an organization's infrastructure [62][53].
- **Application scans**: This type of vulnerability assessment involves using automated scans or static/dynamic analysis of source code to identify security vulnerabilities in web applications and their source code on the front-end [62].

2.3.1 Vulnerability Assessment: Security Scanning Process

As Figure 2-1 shows, the process of security scanning can be broken down into four main steps: testing, analysis, assessment, and remediation. The testing step involves running security tests on the system or application to identify any vulnerabilities or weaknesses. The analysis step involves analysing the results of the tests to determine the severity of any identified vulnerabilities. The assessment step involves evaluating the overall risk posed by the vulnerabilities and prioritizing them for remediation. Finally, the remediation step involves addressing the identified vulnerabilities by implementing patches or other security measures to mitigate the risk. [74].

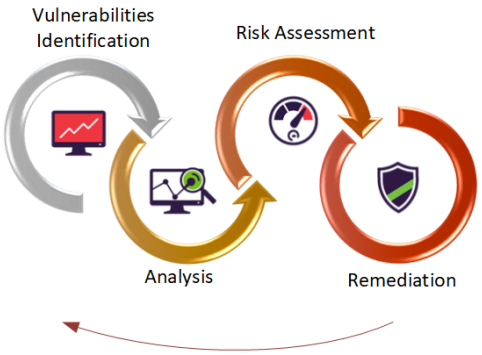


Figure 2-1. Security Scanning Process

2.4 Vulnerability Optimization

Vulnerability management optimization involves processes and strategies to identify, prioritize, and remediate security vulnerabilities in an organization's network and systems more effectively and efficiently [193]. Here are the key aspects of optimizing vulnerability management:

1. **Vulnerability Identification**: This involves using tools and techniques like automated scanners and penetration testing to find vulnerabilities in systems and software [194].
2. **Prioritization**: Not all vulnerabilities pose the same level of risk. Effective optimization requires assessing the severity and potential impact of each vulnerability, often using scores like CVSS and considering the context of the organization's specific environment.
3. **Patch Management**: This is the process of managing updates to software and systems that fix security vulnerabilities. An optimized process ensures that patches are applied as soon as they are available and verified [195].
4. **Remediation**: Beyond patching, remediation may involve configuration changes or other mitigations to reduce risk when immediate patching is not possible.
5. **Automation**: Implementing automated tools to streamline the vulnerability management process is crucial. This can include automated scanning, alerting, and even remediation tasks [193].
6. **Integration**: Integrating vulnerability management tools with other IT and security systems, like SIEM (Security Information and Event Management) and incident response platforms, can enhance the visibility and responsiveness of the security team.

7. **Continuous Improvement**: Cybersecurity landscapes and technologies evolve rapidly, making continuous assessment and improvement essential to maintaining effective vulnerability management practices [196].

2.4.1 Vulnerability Identification (Testing)

In this step, the main goal is to create an exhaustive list of vulnerabilities present in an application. Security analysts achieve this by performing security checks on applications, servers or other systems using automated tools or by manually testing and evaluating them. They also utilize various resources such as vulnerability databases, vendor vulnerability announcements, asset management systems, and threat intelligence feeds to identify potential security weaknesses [75].

2.4.2 Vulnerability Analysis

In the second step of the security scanning process, the aim is to pinpoint the origin and fundamental reason behind the vulnerabilities identified in step one. This entails recognizing the system components accountable for each vulnerability and the underlying cause of the vulnerability itself. To illustrate, the underlying cause of a vulnerability could be an outdated version of an open-source library. This, in turn, facilitates a clear path for remediation, such as updating the library to the latest version [76][77].

2.4.3 Risk Assessment

The objective of this step is to prioritize vulnerabilities based on their severity or level of risk. This step involves security analysts assigning a rank or score to each vulnerability based on various factors such as:

- Which systems are affected?
- What data is at risk.
- Which business functions are at risk?
- Ease of attack or compromise.
- Severity of an attack.
- Potential damage as a result of the vulnerability.

2.4.4 Association Between Vulnerability and Risk Assessments

The association between vulnerability and risk assessments is fundamental to effective risk management:

- Starting Point: Vulnerability assessment often serves as a starting point for a risk

assessment. Identifying vulnerabilities provides a concrete basis for Analysing risks.

- Risk Evaluation: The vulnerabilities identified provide the necessary detail to evaluate risks in terms of potential impact and likelihood. Without understanding the vulnerabilities, it would be challenging to accurately assess the risks.
- Prioritization: Both processes help in prioritizing actions. By understanding which vulnerabilities are most severe and which potential risks have the highest impact and likelihood, organizations can allocate resources more effectively to address the most critical issues first.
- Mitigation and Response: Insights from both assessments guide the development of strategies for mitigating risks. This might involve patching vulnerabilities, implementing additional security controls, or even accepting certain risks based on an informed decision.
- Continuous Process: Both are continuous processes. As new vulnerabilities are discovered and the risk landscape evolves (with new threats emerging and business objectives changing), ongoing assessments are necessary to maintain security and resilience.

2.4.5 Remediation

The goal of this step is to address the security vulnerabilities identified in the previous steps by closing the security gaps. The process typically involves collaboration between the security team, development team, and operations team. They work together to determine the most effective approach for remediation or mitigation of each vulnerability. Remediation options can include patching systems, updating software, revising configurations, or implementing new security controls. Once the vulnerabilities have been addressed, the system should be retested to ensure that the vulnerabilities have been successfully remediated [78].

Specific remediation steps might include:

- Introduction of new security procedures, measures or tools.
  - The implementation of new security procedures, measures, or tools can help prevent or mitigate vulnerabilities in an organization's IT systems. These could include measures like implementing multifactor authentication, regularly updating software, and configuring firewalls and intrusion detection systems. New security tools such as vulnerability scanners, endpoint protection software, and security information and event management (SIEM) systems can also be introduced to enhance an organization's security posture [92].

- The updating of operational or configuration changes.
  - The updating of operational or configuration changes is the process of modifying the IT infrastructure to eliminate vulnerabilities. This can include patching or updating software or hardware components, changing network configurations, or implementing new security controls to mitigate the identified risks. The goal is to address the vulnerabilities in a way that minimizes disruption to operations and ensures that the security measures are effective [95].
- Development and implementation of a vulnerability patch [99][100].
  - One of the remediation steps for addressing a vulnerability is to develop and implement a patch to fix the issue. A patch is a software update that addresses a security vulnerability or bug in an application, system, or network. Once a patch is developed, it must be thoroughly tested to ensure it doesn't cause any unintended issues or conflicts with other systems or applications. Once it has been deemed safe, the patch can be deployed to the affected systems to eliminate the vulnerability.

Vulnerability assessment cannot be a one-off activity. To ensure the effectiveness of vulnerability assessment, organizations must integrate it as a regular process and not just a one-time activity. It is essential to promote collaboration among security, operations, and development teams, which is known as DevSecOps [101]. DevSecOps involves embedding security into the quick-release cycles that are typical of modern application development and deployment, which is a culture shift in the software industry [102].

2.4.6 Vulnerability Assessment Action and Step

The process of vulnerability assessments involves identifying various system and network vulnerabilities using a range of techniques, tools, and methodologies to detect threats, risks, and vulnerabilities [13]. Table 2.2 indicates the vulnerability assessment action and steps [14].

Table 2.2. Vulnerability assessment action and steps

Vulnerability Assessment Action	Vulnerability Assessment Steps
Assets identification	Inventory the assets. Determine assets' relative value.

Threat identification	Classify threats by category. Design attack.
Vulnerability appraisal	Determine current weakness in protecting assets. Use vulnerability assessment tools.
Risk assessment	Estimate impact of vulnerability on organization. Calculate risk likelihood and impact of risk.
Risk mitigation	Decide what to do with risk.

2.4.7 What Causes Vulnerabilities?

There are many causes of vulnerabilities including [106]:

- **Complexity:** Complexity refers to the degree of sophistication of a system, and it is a factor that can contribute to the occurrence of vulnerabilities in a system. When a system is complex, it can increase the likelihood of misconfigurations or unintended access, making it easier for attackers to exploit vulnerabilities [106].
- **Familiarity:** The likelihood of finding or having information about known vulnerabilities increases with common code, software, operating systems, and hardware [106][107].
- **Connectivity:** The more connected a device is, the greater the chance of vulnerability. [107].
- **Poor password management:** Weak passwords can be broken with brute force and reusing passwords can result in one data breach becoming many [106].
- **Operating system flaws:** Operating systems, just like any other software, can have vulnerabilities. These vulnerabilities can allow viruses and malware to execute commands on a computer. Operating systems that are insecure by default and provide all users with full access can also lead to security issues [108].
- **Internet usage:** When using the internet, there is a risk of unintentionally installing spyware or adware on your computer [106].
- **Software bugs: Software bugs:** Programmers can accidentally or deliberately leave an exploitable bug in software [108].
- **Unchecked user input:** If a website or software doesn't properly verify user input, it can inadvertently run SQL commands that were not intended [107].
- **People:** The most significant weakness in any organization is the individual who is using

the system. Social engineering poses the most significant danger to most organizations. [106].

2.4.8 Vulnerability Assessment Tools

There are quite a few tools that can be used for analysing systems and identifying vulnerabilities [100].

- **Port scanners**
  - Port scanners are a type of vulnerability assessment tool that is used to identify open ports on a networked device. Ports are like doors that allow communication between devices, and each port is assigned a unique number. Port scanners can be used to scan all or a subset of the ports on a device to determine which ones are open, and therefore potentially vulnerable to attacks [141].

    There are several types of port scanners, including [143]:

    1. TCP port scanners - these scanners establish a TCP connection with each port to see if it is open or closed.
    2. UDP port scanners - these scanners send a UDP packet to each port to see if it is open or closed.
    3. SYN port scanners - these scanners send a SYN packet to each port and analyse the response to determine if the port is open, closed, or filtered.
    4. ACK port scanners - these scanners send an ACK packet to each port and analyse the response to determine if the port is open, closed, or filtered.
    5. FIN port scanners - these scanners send a FIN packet to each port and analyse the response to determine if the port is open, closed, or filtered.

  Port scanners can be used for both legitimate and malicious purposes. Legitimate purposes include assessing the security of a network, while malicious purposes include reconnaissance and planning for an attack. It is important to note that using port scanners on devices that you do not own or have permission to scan is illegal and unethical [142].

- **Banner grabbing tool**
  - A banner grabbing tool is a type of vulnerability assessment tool that is used to extract information from the banners or headers of network services running on a

target system. Banners or headers contain information about the operating system, web server, application server, and other software running on the system. This information can be used to identify vulnerabilities or misconfigurations that could be exploited by an attacker. Banner grabbing tools typically connect to the network service, such as a web server or FTP server, and retrieve the banner information. The information can then be analysed to identify known vulnerabilities associated with the software or operating system running on the system [144]. This information can be used to prioritize vulnerability remediation efforts or to plan an attack. Some common banner grabbing tools include Netcat, Telnet, and Nmap. These tools can be used to connect to network services and retrieve banner information. Nmap is a particularly popular banner grabbing tool that is used for network exploration, security scanning, and vulnerability assessment. It includes a variety of banner grabbing techniques, such as TCP/IP fingerprinting and service version detection [145].

It is important to note that banner grabbing can be used for both legitimate and malicious purposes. Legitimate uses include identifying vulnerabilities and misconfigurations on systems within an organization, while malicious uses include reconnaissance and planning for an attack. It is important to obtain permission before using banner grabbing tools on systems that do not belong to you or that you are not authorized to scan [144].

- **Protocol Analyser.**
  - A protocol analyser, also known as a packet sniffer, is a type of vulnerability assessment tool that is used to capture and analyse network traffic. Protocol analysers capture packets of data as they travel across the network and provide detailed information about the packets, including the source and destination addresses, protocols used, and the contents of the packet. Protocol analysers can be used for a variety of purposes, including network troubleshooting, performance monitoring, and security analysis. In terms of vulnerability assessment, protocol analysers can be used to identify vulnerabilities and misconfigurations in network protocols and applications [146].

    Protocol analysers can detect a variety of issues, such as [147]:

1. Unencrypted passwords and other sensitive data transmitted in clear text.
2. Malicious traffic generated by malware or other malicious software.
3. Misconfigured network devices that are generating traffic that is not compliant with industry standards.
4. Suspicious traffic patterns or traffic originating from unauthorised sources.

Some popular protocol analysers include Wireshark, TCPdump, and Microsoft Network Monitor. These tools can be used to capture and analyse network traffic and provide detailed information about the packets and protocols being used. It is important to note that protocol analysers can be used for both legitimate and malicious purposes. Legitimate uses include identifying network issues and troubleshooting problems, while malicious uses include capturing sensitive data or launching attacks. It is important to obtain permission before using protocol analysers on networks that do not belong to you or that you are not authorized to monitor [146].

- **Vulnerability scanner**
  - A vulnerability scanner is a type of vulnerability assessment tool that is used to identify vulnerabilities in networked devices, applications, and systems. Vulnerability scanners use a database of known vulnerabilities to scan devices for potential security weaknesses [148].

    Vulnerability scanners can detect a variety of vulnerabilities, including [149]:

    1. Missing security patches and updates.
    2. Weak passwords and authentication mechanisms.
    3. Misconfigured servers and network devices.
    4. Unsecured network services and protocols.
    5. Outdated software and operating systems.

    Some popular vulnerability scanners include Nessus, OpenVAS, and QualysGuard. These tools can be used to perform scans of networked devices and systems and generate reports that identify vulnerabilities and suggest remediation steps [149].

    It is important to note that vulnerability scanners are not foolproof and may generate false positives or false negatives. It is important to manually verify the results of vulnerability scans and perform additional testing and analysis as

needed. Vulnerability scanners can be used for both legitimate and malicious purposes. Legitimate uses include identifying vulnerabilities and misconfigurations in systems and applications within an organization, while malicious uses include reconnaissance and planning for an attack. It is important to obtain permission before using vulnerability scanners on systems that do not belong to you or that you are not authorized to scan [148].

- **Honeypots and Honeynets**

  Honeypots and honeynets are types of vulnerability assessment tools that are used to detect and analyse attacks against a network or system. Honeypots are decoy systems that are designed to look and behave like real systems but are isolated from the main network and contain no real data. Honeynets are multiple interconnected honeypots that can be used to detect and analyse attacks on a larger scale [109].

  Honeypots and honeynets are used for a variety of purposes, including [150]:

  1. Detecting and Analysing attacks: Honeypots and honeynets are used to lure attackers into interacting with a decoy system, allowing security professionals to analyse their methods and gather intelligence.
  2. Testing and evaluating security measures: Honeypots and honeynets can be used to test the effectiveness of security measures and to identify vulnerabilities and weaknesses.
  3. Developing and improving security policies: Honeypots and honeynets can be used to collect data and analyse attack trends, which can be used to improve security policies and procedures.

  Some popular honeypot tools include Honeyd, Dionaea, and Glastopf. Honeypots can be used to emulate a variety of systems and services, including web servers, email servers, and file servers. It is important to note that honeypots and honeynets are not foolproof and may not detect all attacks. They are also not a substitute for traditional security measures, such as firewalls and intrusion detection systems. Honeypots and honeynets should be used as part of a larger security strategy [151].

  Honeypots and honeynets can be used for both legitimate and malicious purposes. Legitimate uses include detecting and analysing attacks and improving security policies, while malicious uses include launching attacks or using the honeypots to gather sensitive

information. It is important to obtain permission before using honeypots and honeynets on networks that do not belong to you or that you are not authorized to monitor [105].

2.4.9 Modelling Vulnerabilities

There are numerous methods available for modelling vulnerabilities to perform analysis in a computer system [120].

- **Baseline Reporting:** Comparing the current state of a system to its baseline state [107].
- **Programming Vulnerabilities**: List potential threats from threat agent.

Here are some examples of potential threats that a threat agent could pose [120]:

1. Unauthorised access to sensitive data or systems.
2. Data theft or sabotage.
3. Physical theft or damage to hardware.
4. Malware or virus infection.
5. Denial of service attacks.
6. Social engineering attacks such as phishing or spear phishing.
7. Exploiting software vulnerabilities to gain access or escalate privileges.
8. Insider threats or espionage.
9. Distributed denial of service attacks.
10. Advanced persistent threats (APTs).

The area of safety critical systems investigates the process of hazard analysis. In the context of a computer system, vulnerabilities may be regarded as potential hazards [121].

The different techniques that analyse hazards include [122]:

- Checklists
- Fault tree analysis
- Event tree analysis
- Cause-consequence analysis.

There are issues with using certain techniques, such as checklists, for vulnerability analysis. Checklists are limited in that they are static and do not illustrate the connections between vulnerabilities, nor do they investigate the reasons behind why vulnerabilities are interconnected. Fault trees, on the other hand, simply depict the sequence of events over time and are insufficient for visualizing and modelling the various types of relationships between vulnerabilities [121].

Another approach to examine the relationships between vulnerabilities is the use of vulnerability trees, which can utilize historical attack data to generate attack patterns and attack trees. This technique aims to anticipate the route that a threat actor might take by analysing potential exploits. Each pathway within an attack tree represents a distinct attack on the organization. However, it is important to note that this method has its own limitations. The primary issue with attack trees is that they are unable to analyse large systems or networks with a significant degree of complexity [121][122].

One major advantage of utilizing attack trees is that they concentrate on measurable objectives that can be translated into specific tests against real-world devices, networks, and protocol implementations. This approach helps avoid the issue of overly theoretical security research that often neglects the difficulty of executing attacks and cannot measure the impact on targeted systems. Attack trees also encourage a systematic elaboration of events, such as specific attack objectives, that must happen for a successful breach to occur. This promotes the consideration of all viable approaches for an attack and enables the identification and optimal implementation of countermeasures [125]. Additionally, since each node (an individual attacker objective) can be broken down into subordinate nodes (sub-goals or a means of accomplishing the parent goal), attack trees permit security analysis to be conducted at multiple levels of abstraction, allowing researchers to focus on areas of interest while recognizing other intrusion paths. Lastly, using attack trees allows for common attacks to be referred to as reusable modules that apply to multiple network scenarios [123].

2.4.10 Vulnerability Attributes

- What are the attributes that describe a vulnerability and attacks and for every vulnerability attribute, what are the accepted values and what is the attribute type?

There are some different definitions of “cyber-attack or vulnerability”. The following are the definitions of cyberattack and vulnerability classified in six categories. Each definition can address one or more of the following five questions:

- What is the target of the attack?
- What is the objective of the attacker?
- What is the attack method?
- What is the impact on the target?

- What is the attacker's motivation or intent?

Here are the six categories with corresponding definitions of cyber-attack and vulnerability:

1. Targeted attack: A cyber-attack that is aimed at a specific individual, organization, or system. A vulnerability is a weakness that can be exploited in such an attack [121].
2. Exploit-based attack: A cyber-attack that leverages known or unknown vulnerabilities in software or hardware. A vulnerability is a weakness that can be exploited in this type of attack [123].
3. Social engineering attack: A cyber-attack that uses manipulation or deception techniques to trick individuals into divulging sensitive information or performing an action that benefits the attacker. A vulnerability is a gap in the individual's knowledge or judgment that makes them susceptible to such tactics [124].
4. Distributed attack: A cyber-attack that is launched from multiple sources simultaneously. A vulnerability is a weakness in the targeted system that allows it to be overwhelmed by the high volume of traffic [124].
5. Insider attack: A cyber-attack that is carried out by someone who has authorized access to the system or information being targeted. A vulnerability is a flaw in the system or security protocols that allows the insider to misuse their privileges [125].
6. Advanced persistent threat (APT) attack: A sophisticated, long-term cyber-attack that is aimed at a specific target, usually for the purpose of stealing sensitive information or disrupting operations. A vulnerability is a weakness in the targeted system or network that is exploited over a prolonged period of time [126].

The six definitions identified suggest that the concept of cybersecurity has at least five attributes.

- **Actors:** In each cyberattack, there are two parties involved: the target asset owner and an attacker. The definition of a cyber-attack does not specify the type or identity of the attacker. Both offensive and defencive operations can be conducted by various actors such as nation-states, companies, groups, collectives, or individuals [127].
- **Assets targeted:** Five out of six of the definitions mentioned earlier specify the assets that are targeted in cyber-attacks. These assets may include computer systems and networks, information, programs, or functions that are either in or passing through the systems or networks, computer-controlled physical infrastructure, and physical objects that are external to the computer, system, or network. These definitions can be found in sources

[121][122].

- **Motivation:** Cyber-attacks can be motivated by various reasons, such as gaining access to secure or unauthorised information, espionage, theft of data or money, national security or political interests, propaganda, or deception. These motivations are identified in source [123].
- **Effect on targeted assets:** The consequences of cyber-attacks can include various alterations or damages to assets, such as alteration, deletion, corruption, deception, degradation, disablement, disruption, or destruction, as well as denying access to assets. Different definitions of cyber-attacks also describe the effects on assets in logical, physical, and cognitive terms. For instance, denying access to assets is an example of logical effects, while the use of false information to deceive an adversary represents a cognitive effect. Physical effects, on the other hand, involve the destruction of capital assets [124][125].
- **Duration:** Out of the six definitions mentioned earlier, only one provides information on the intended duration of cyber-attacks [121].

2.4.11 Examination of High-Profile Cyber-Attacks

High-profile cyber-attacks are becoming increasingly common and can have serious consequences for individuals, organizations, and even entire countries [152]. Here are a few examples of recent high-profile cyber-attacks and their impact [153]:

- SolarWinds hack: In December 2020, it was discovered that a sophisticated cyber-attack had compromised the software of SolarWinds, a company that provides network monitoring and management software to government agencies and Fortune 500 companies. The attackers were able to insert malicious code into the software, which was then distributed to SolarWinds customers. This allowed the attackers to gain access to the networks of numerous organizations, including the US government and major tech companies. The full extent of the damage caused by the SolarWinds hack is still being assessed, but it is believed to be one of the largest and most damaging cyber-attacks in history.
- Colonial Pipeline ransomware attack: In May 2021, a ransomware attack targeted the computer systems of Colonial Pipeline, a major fuel pipeline operator in the US. The attackers were able to encrypt the company's data and demand a ransom payment in exchange for the decryption key. The attack caused widespread panic and led to fuel

shortages and price increases in several states. The Colonial Pipeline attack highlighted the vulnerability of critical infrastructure to cyber-attacks and the need for increased cybersecurity measures.

- JBS ransomware attack: In June 2021, JBS, the world's largest meat processing company, was targeted by a ransomware attack that disrupted its operations in the US, Canada, and Australia. The attack caused significant disruption to the meat supply chain and highlighted the vulnerability of the food industry to cyber-attacks.

These high-profile cyber-attacks demonstrate the increasing sophistication of cyber criminals and the need for strong cybersecurity measures. Organizations must take steps to protect their networks and systems from cyber-attacks, including implementing strong passwords and multi-factor authentication, regularly updating software and security patches, and conducting regular vulnerability assessments and penetration testing. Additionally, governments must work to improve cybersecurity laws and regulations and develop strategies to defend against cyber-attacks on critical infrastructure [152].

You can find some more examples in Appendix (Table A.1).

2.4.12 MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge)

MITRE ATT&CK is a comprehensive framework that catalogues various tactics, techniques, and procedures (TTPs) employed by threat actors during cyber-attacks. It organizes these TTPs into matrices that depict different stages of an attack, such as initial access, execution, persistence, privilege escalation, and exfiltration. By mapping out the tactics and techniques used by adversaries, MITRE ATT&CK helps organizations enhance their defences, detect threats, and respond effectively to cyber-attacks [179]. The ATT&CK framework details a variety of tactics and techniques used by adversaries, which can help security teams identify specific vulnerabilities that are being actively exploited in the wild. For example, if the framework indicates a rise in the use of spear-phishing (T1566) to deliver malware, organizations can prioritize securing their email systems against such attacks [180]. Organizations can use ATT&CK to benchmark their security posture against an industry-standard framework to identify gaps in their defences. By mapping their current security measures against the techniques listed in ATT&CK, they can see where they lack coverage and need to bolster their defences. Furthermore, ATT&CK's comprehensive listing of techniques and associated indicators of compromise (IoCs) aids in enhancing detection capabilities. Security teams can configure their security monitoring tools to flag activities related to these techniques,

allowing for quicker detection and response to potential threats. Additionally, it provides context on how attacks unfold, which helps in creating more effective incident response plans [179].

2.4.12.1 Cyber Kill Chain

The Cyber Kill Chain, originally developed by Lockheed Martin, outlines the stages of a cyber-attack from the perspective of an attacker's activities. These stages typically include reconnaissance, weaponization, delivery, exploitation, installation, command and control, and actions on objectives. The Cyber Kill Chain framework provides a structured approach for understanding and mitigating cyber threats, allowing organizations to identify and disrupt attacks at different stages of the kill chain. With a clear view of the stages of an attack, organizations can place defencive measures strategically along the kill chain. For instance, improving reconnaissance detection capabilities can alert an organization to an impending attack before it even begins, while strengthening defences against 'Command and Control' communications can help cut off an attacker's access to compromised systems. And also, it helps in planning effective incident responses by identifying which stage of the kill chain an attack is in, allowing for targeted containment and remediation strategies. If an attack is detected at the 'Exploitation' stage, immediate action can be taken to isolate affected systems and prevent the installation of malicious software. By understanding the common 'Exploitation' techniques used to compromise systems, organizations can prioritize patch management efforts to address the most likely avenues of attack first, such as patching vulnerabilities that could be exploited by recently weaponized exploits [181].

2.4.13 How Many Types of Vulnerabilities Exit?

To simplify, a vulnerability in a computer system is a weakness or defect in a system or network that can be used to cause harm or allow an intruder to manipulate the system [154]. Cyber-attacks usually consist of four primary stages.

There are four main stages present in most cyber-attacks:

- Survey: Surveying is the initial stage of a cyber-attack, which involves gathering and analysing all possible information about the target with the aim of identifying any potential vulnerabilities that can be exploited [126].
- Delivery refers to the stage of a cyber-attack where the attacker gains access to the vulnerable point in a system, allowing them to carry out the attack [126].
- Breach: Breach refers to the stage in a cyber-attack where an attacker successfully exploits the identified vulnerability or vulnerabilities to gain unauthorised access to a system [127].
- Affect: This stage refers to the attacker carrying out their intended actions within the

system once they have gained unauthorised access through exploiting the vulnerabilities. The specific activities carried out will depend on the attacker's goals, which may include data theft, system disruption, or other malicious actions [126].

A vulnerability is a weakness or a gap in an application that can result from an error during coding, compilation, or implementation of software. This error, also known as a bug, can cause harm to the stakeholders of the software indirectly. Stakeholders can refer to any entity that depends on the software, as well as the software's owner and users. In most cases, vulnerabilities can be exploited by hackers to launch an attack on the system [125].

The three elements of a vulnerability include the flaw or weakness itself, an attacker's ability to exploit it, and the potential impact on the system or network.

1. A system susceptibility or flaw
2. Attacker access to the flaw
3. Attacker capability to exploit the flaw [127].

Significant security breaches can have severe consequences, and often IT managers are held responsible for overlooking system vulnerabilities and bugs. Conducting IT security audits can reduce the risk of attacks by outsourcing the security of the system to a certified, independent IT security professional or firm. This allows for proper management of the company's IT environment without much input from IT managers. Regular penetration testing can also be applied to identify blind spots or weaknesses and provide protection to those areas to avoid catastrophic security breaches. These tests are typically carried out by white hat hackers funded by the companies [126].

The most frequent types of vulnerabilities that can pose a threat to cybersecurity are generally categorized into four types: software security and web vulnerabilities, network security vulnerabilities, configuration vulnerabilities, and cyber gaps [127]. These categories have been identified in various sources such as research studies and reports.

The most common software and web vulnerabilities are described in the next subsections.

2.4.13.1 Software Security Vulnerabilities

Software security vulnerabilities are defined as weaknesses or flaws in software that can be exploited by attackers to gain unauthorised access to systems or data, disrupt system operations, or cause damage to systems or data. These vulnerabilities can be introduced during the software development process or arise due to changes in the software's environment or interactions with

other software [155].

Some common types of software security vulnerabilities include:

2.4.13.1.1 Improper Input Validation

Input validation is a security technique that provides an additional layer of protection against attackers attempting to access unintended functionality or escalate privileges. The Department of Homeland Security recommends that input validation be used to only allow legitimate data to be entered into the system. If input validation is not properly implemented, it can leave the system vulnerable to attacks. The next section will discuss various types of attacks that can occur as a result of inadequate input validation [127].

2.4.13.1.2 Buffer Overflow

Improper input validation can lead to buffer overflow vulnerabilities, which are caused by programming errors. These errors occur when the programmer does not consider the possibility of an input from the end-user that exceeds the allocated memory space, such as entering 2000 characters for a last name. The extra data overwrites the next memory and causes the program to crash. Attackers can exploit this by executing code that allows them to establish an interactive session and send commands with the program's privileges. Network protocols that lack proper input validation are vulnerable to buffer overflow attacks. Poor coding practices can allow attackers to inject unexpected data and modify the program execution, leading to stack-based or heap-based buffer overflow vulnerabilities. These vulnerabilities can allow remote code execution on the host [128][129].

2.4.13.1.3 Lack of Bound Checking

When input validation is not properly implemented, it can lead to the insertion of invalid or unexpected data into the program, causing it to behave in an unintended manner or even crash. One example is when very large numbers are inserted into an array, which can cause the service to crash. This is due to poor coding practices that allow attackers to supply unexpected data and modify the program's execution. Therefore, it is important to implement proper input validation techniques to prevent such issues from occurring [129].

2.4.13.1.4 Command Injection

Command injection is a type of attack where the attacker injects commands or code for unauthorised execution into an application. There are two main types of command injection: SQL injection and OS command injection. The attacker typically injects a semicolon to separate one

command from another. The data being inserted is untrusted and forms part of a string that is executed as a command by the application. If the command is successfully executed, the attacker can gain the privileges or capabilities they need to compromise the system [130].

2.4.13.1.5 SQL Injection

SQL command injection is a type of attack that is commonly used on database-driven websites. The attacker is able to inject malicious scripts into a legitimate request that the website returns to the victim. The victim's web browser executes the malicious scripts because they appear to come from a trusted source, which can compromise the victim's computer through browser exploits. This is typically caused by a lack of data sanitization and can lead to a variety of malicious activities, such as the attacker sending requests to the website on the victim's behalf and potentially gaining supervisory control privileges [131].

2.4.13.1.6 Improper Limitation of a Pathname to a Restricted Directory

When input validation is not done properly, it can lead to directory traversal vulnerabilities, which occur when the software constructs a pathname using external inputs that is supposed to locate a file or directory within a sub-directory of the parent directory. Attackers can exploit this vulnerability to gain access to critical files and directories, which can result in the execution of unauthorised code and commands, modification of files and directories, and even denial of service attacks by crashing important files [132].

2.4.13.1.7 Poor Code Quality

Attackers can successfully infiltrate critical infrastructures due to poor code quality resulting from improper development and maintenance practices. These vulnerable programs do not adhere to secure development concepts and good programming practices. One of the causes of poor code quality is the use of unsafe function calls, for which the developer is responsible for validating input. This leads to buffer overflow and malformed input vulnerabilities, which pose a high risk to system security. Another issue that arises from poor code quality is the Null pointer dereference, which occurs when a pointer that is expected to be valid is NULL, leading to unexpected program crashes. Unless exception handling is implemented, null pointer dereference will usually result in process failure [133].

2.4.13.1.8 Permission, Privileges, and Access Controls

The European Commission's Cybersecurity Strategy, released in 2013, identifies that critical infrastructure is vulnerable to attacks due to inadequate permissions, privileges, and access

controls on systems. Attackers exploit these vulnerabilities to gain unauthorised access and perform illegitimate actions. Insufficient access controls and checks across all potential execution paths allow unauthorised users to access data or perform actions they shouldn't be able to. The attack takes advantage of unrestricted access to objects, common shares available on multiple systems, lack of role-based authentication, and remote users who can imitate any process without authorization. Attackers can upload files to any location on the targeted computer without restrictions using an undisclosed "back door" to gain access and remain anonymous over the network. The system fails to follow the principle of least privileges, where users should have multiple accounts for functions that require different levels of privileges, and default configurations are not changed. Due to unnecessary privileges, the attacker gains access to the network and then attempts privilege escalation to exploit vulnerable services. Full access will allow the attacker to inflict significant damage that could affect the entire operation [132][133].

2.4.13.1.9 Improper Authentication

Improper authentication mechanisms can also lead to cyberattacks on critical infrastructure, as identified by the "Information Security Breaches, 2014" report by GCHQ, UK. Inadequate verification of the user's identity can be exploited by attackers to gain unauthorised access and escalate their privileges. Weak authentication methods can be bypassed by modifying client-side credentials or removing client-side checks altogether. This can allow the attacker to perform illegal transactions by sending modified values to the server. Moreover, due to the lack of proper authentication, man-in-the-middle attacks can occur, where the attacker intercepts and modifies communication between two parties. This vulnerability in the authentication mechanism can be exploited to launch cyberattacks on critical infrastructure [133].

2.4.13.1.10 Insufficient Verification of Data Authenticity

Insufficient verification of data authenticity can lead to various types of attacks on critical infrastructure. One such attack is Cross-Site Request Forgery (CSRF), where the attacker tricks a client into making an unintentional request to the web server that is treated as an authentic request. This attack allows the attacker to change settings and hijack credentials, giving them the ability to perform any task as an authorized user. In addition, some transmission protocols do not include a mechanism for verifying the integrity of the data during transmission. Without checksum functionality, there is no way to determine if the data has been corrupted during transmission. This exclusion of checksum value removes the first application-level check of data that can be used,

leaving data vulnerable to malicious alteration during transmission. Another type of attack that can occur is the execution of source code or executable code without sufficiently verifying the origin and integrity of the code. Attackers can execute malicious code to compromise the host server, spoof an authorized server, or modify data while it is in transit [133].

2.4.13.1.11 Cryptographic Issues

Another attack that can take place on critical infrastructure is the man-in-the-middle attack, where an attacker intercepts the communication between two parties and can alter or manipulate the data being transmitted. This type of attack is particularly dangerous when encryption is weak or non-existent, as the attacker can easily read and modify the data. The attacker can impersonate the legitimate parties and communicate with them separately while relaying the messages between them, making it difficult for them to detect the attack. This can result in unauthorised access to sensitive information or even control over the critical infrastructure. To prevent this type of attack, strong encryption and authentication mechanisms must be in place to ensure the integrity and confidentiality of the data being transmitted [98].

2.4.13.2 Network Security Vulnerabilities

The network architecture needs to be securely designed to ensure that all business processes can be remotely accessed and monitored, while preventing unauthorised traffic from entering the network, the network architecture must be designed with security in mind. One way to achieve this is by implementing security zones with access control rules that restrict the flow of traffic in and out of the zone, thereby reducing the risk of intentional or unintentional attacks. In the next section, we will describe the types of attacks that can occur on critical infrastructure due to weaknesses in network security [91][92].

2.4.13.2.1 Poor Network

The lack of a defence in-depth strategy in network design is a significant cause of attacks on critical infrastructure. Networks that do not have multiple layers of security, use flat networks without perimeters or zones, lack port security, and have poor remote access policies are particularly vulnerable to successful attacks. To compound the problem, these networks are often directly connected to the corporate environment without firewalls or DMZ zones, providing direct access to the internet. Poorly designed networks make it easier for hackers to conduct successful attacks on critical infrastructure [93].

2.4.13.2.2 Security Perimeter

Defining clear boundaries for network security is essential to defend against any type of attack on critical infrastructure. The security perimeter should be logically separated from the corporate network using physical devices and additional network security controls should be put in place to prevent intrusion. If the security perimeter is not clearly defined, it can lead to unauthorised access to the system and data [90].

2.4.13.2.3 Lack of Network Segmentation

The absence or insufficient use of security zones can make it easier for attackers to gain complete control of the system, which can result in severe consequences. Without internal segmentation, servers using the Inter-Control Center Communication protocol (ICCP) and applications that handle sensitive data through dedicated serial links are not isolated within DMZ servers, making them vulnerable to successful attacks [94].

2.4.13.2.4 Firewall Issues

Poorly configured or absent firewalls can lead to unauthorised data being transmitted between networks without proper checks, making critical infrastructures vulnerable to successful attacks. Malware and viruses can easily spread between networks, and confidential data can be compromised if firewalls are not set up properly. Examples of successful attacks have been attributed to instances where connections to and from remote facilities bypass the firewall due to poor configuration [90][91].

2.4.13.3 Configuration Vulnerabilities

Successful attacks on critical infrastructure often occur due to improperly configured network devices, which can allow unauthorised access. Access control lists (ACLs) on network devices are often not properly configured, allowing unauthorised users to access the devices. Another major weakness is remote access to these network devices without using encryption or proper authentication protocols, which can allow attackers to gain root access and change the network device configuration. Additional security measures such as encryption and proper authentication protocols are necessary to prevent these attacks and restrict unauthorised access to network devices [96].

2.4.13.3.1 Permission, Privilege, and Access Controls

Successful attacks on critical infrastructure can be attributed to a lack of policies and controls governing access and permissions. This includes the absence of separation of duties, failure to

enforce lockout of systems after multiple failed login attempts, and the lack of mechanisms to terminate remote access sessions after a certain period of time. These vulnerabilities allow attackers to gain access to critical systems and cause significant damage [97].

2.4.13.3.2 Improper Authentication

Successful attacks on critical infrastructure are often due to weak or improper authentication methods, such as a lack of policies or procedures. Many organizations lack formal documentation that outlines clear authentication policies and controls and fails to specify how users and devices should be authenticated before establishing a connection. Weak authentication policies make it difficult to uniquely identify and authenticate users, and may not provide role-based, group-based, or device-based authentication. To properly manage users, organizations must verify their identities and receive authorization to provide them with appropriate authentication to access the system [98].

2.4.13.3.3 Credential Control Management

According to MITRE, (2015) “Common Attack Pattern Enumeration and Classification (CAPEC), MITRE's CAPEC highlights the importance of protecting credentials belonging to authorized users from attackers. Hackers can intercept and view credentials passed over networks in plain text, making it crucial to properly encrypt and hash passwords to prevent unauthorised access to privileged accounts. Services such as FTP, telnet, and rlogin transmit credentials in plain text, making them vulnerable to attack. Additionally, improper database service configuration can lead to administrator passwords being displayed on web pages and password hash files not being secured properly, further increasing the risk of successful attacks on critical infrastructure [94].

2.4.13.3.4 Security Configuration and Maintenance

Attacks on critical infrastructure can also occur due to poor system configuration management. Configuration management involves the process of tracking and controlling changes made to software, hardware, and system configurations in order to maintain their integrity and security. If proper configuration management practices are not followed, then the system can be left open to vulnerabilities that can be exploited by attackers. Another factor that can contribute to attacks on critical infrastructure is the lack of security awareness and training among employees. Human errors, such as using weak passwords or failing for phishing scams, can lead to successful attacks on the system. It is important for organizations to provide regular security awareness training to their employees to educate them on how to identify and avoid potential security threats. Finally,

attacks on critical infrastructure can also occur due to the lack of coordination and communication between different organizations and agencies responsible for the infrastructure. In order to effectively protect critical infrastructure, it is necessary for different stakeholders to work together and share information about potential threats and vulnerabilities. Without proper coordination and communication, critical infrastructure can remain vulnerable to attacks that can have devastating consequences [132].

A system's unpatched software that is not properly maintained or tested is a significant vulnerability. If operating system patches are not applied and computers are using outdated versions, they become vulnerable to attacks on the critical infrastructure through operating system service vulnerabilities. Therefore, timely updates and maintenance are necessary to prevent successful attacks on the systems [94].

2.4.13.3.5 Weak backup and Restore Functions

Having a comprehensive policy for creating backups, storing them in safe offsite locations, and regularly testing them is crucial for the continuity of operations in the event of an incident. However, many organizations do not have consistent policies in place for storing and testing backups, which can lead to integrity and availability issues. It is also important to protect backup information from unauthorised disclosure to prevent attacks on critical infrastructure [95].

2.4.13.3.6 Weak Port Security

Weak port security is a vulnerability that attackers exploit to gain unauthorised network access. This vulnerability allows easy access to hardware interfaces. Malicious users with physical access to an unsecured port on a network switch can easily plug into the network behind the firewall, thereby defeating incoming filtering protection. Weak port security is unable to prevent MAC address changes or the introduction of new unauthorised devices to the network without proper authorization [93].

2.4.13.3.7 Poor Monitoring of IDS

Successful attacks on critical infrastructure are also attributed to the failure to follow good cybersecurity practices for monitoring the Intrusion Detection System (IDS). Corrective actions are not taken to prevent any identified threats. There may be a lack of deployment of network-based or host-based IDS/IPS to effectively monitor the network traffic [93].

2.5 Summary

Chapter 2 provides a comprehensive overview of cybersecurity vulnerabilities, emphasizing their analysis and the implementation of a cyber-immune system. It delves into identifying, classifying, evaluating, and mitigating cybersecurity vulnerabilities as crucial elements of risk management. Methods like automated scanning, penetration testing, and code reviews are employed to uncover vulnerabilities, which are then classified by severity and type. This chapter also discusses the evaluation of potential impacts and the necessary mitigation strategies such as patching and network segmentation. Additionally, it explores various tools and frameworks that support vulnerability analysis, including vulnerability scanners and SIEM systems. Finaly Chapter 2 concludes by introducing vulnerability assessment as a structured examination to identify, categorize, and mitigate vulnerabilities, thus enhancing the security of information systems.

Chapter 3

Vulnerability Interrelationship in Cyber Security

3.1 Overview

In the realm of cybersecurity, the landscape is continuously evolving, presenting new challenges and threats at an unprecedented pace. One of the fundamental aspects of cybersecurity is understanding vulnerabilities within systems, networks, and software. However, vulnerabilities rarely exist in isolation; they often interact with each other in complex ways, creating cascading effects that amplify risks and potential damage. This chapter explores the concept of vulnerability interrelationships and their significance in cybersecurity.

3.2 The Definition of the Vulnerability Interrelationship

During the development of threat assessment methodology, the issue of examining and analysing vulnerabilities was recognized. In the context of cybersecurity, a vulnerability refers to a weakness that can be exploited by cyber attackers to gain unauthorised access to a computer system or perform unauthorised actions. These vulnerabilities can enable attackers to execute code, access a system's memory, install malicious software, and steal, damage, or alter sensitive data [105]. Furthermore, a security threat refers to the possibility of an attack that can result in the misuse of information or resources. On the other hand, a vulnerability is a weakness in a cyber system that

can be exploited by attackers. Although a security threat may or may not lead to an actual attack, it has the potential to cause significant harm. Typically, a security threat exploits one or more vulnerabilities in the system to gain access and compromise it. Therefore, understanding the relationship between security vulnerabilities and potential threats is crucial for developing an effective threat model [107].

3.2.1 What is Vulnerability Interrelationship?

In network security, interrelationship refers to the interconnectedness of various security components and their dependencies on one another. Network security is a critical aspect of information security that involves protecting the confidentiality, integrity, and availability of network resources and data [165].

Some examples of interrelationships in network security include [166]:

1. Firewalls and Intrusion Detection Systems: Firewalls are designed to prevent unauthorised access to a network, while IDS are used to detect and alert administrators to potential security breaches. These two components are often used together to provide layered protection for a network.
2. Encryption and Authentication: Encryption is used to protect data in transit or at rest by scrambling it so that it is unreadable to unauthorised users. Authentication, on the other hand, is used to verify the identity of a user or device before granting access to network resources. These two components work together to ensure that only authorized users are able to access encrypted data.
3. Access Control and Identity Management: Access control is used to restrict access to network resources based on user credentials or other factors, while identity management is used to manage user identities and permissions. These two components are closely related, as effective access control depends on accurate identity management.
4. Patch Management and Vulnerability Scanning: Patch management is the process of keeping software and systems up to date with the latest security patches and updates, while vulnerability scanning is used to identify potential vulnerabilities in a network. These two components are interrelated, as effective patch management is critical to addressing identified vulnerabilities.

Understanding the interrelationships in network security is essential for designing and implementing effective security measures that can protect against a wide range of threats. It

requires a holistic approach that considers all aspects of network security, including hardware, software, and personnel. A comprehensive security strategy should include multiple layers of protection and incorporate best practices for each of the key components [165].

In the context of security, interrelationship refers to the complex interactions and dependencies that exist between different security systems, processes, and stakeholders. This includes both physical security systems, such as access control, surveillance, and perimeter security, as well as cybersecurity systems, such as firewalls, antivirus software, and intrusion detection systems [167].

Some examples of interrelationships in security include [168]:

1. Physical and Cybersecurity: The physical security of a building or facility is closely related to its cybersecurity. For example, if an unauthorised person gains physical access to a computer network, they may be able to bypass security controls and compromise sensitive data.
2. Information Security and Risk Management: Information security and risk management are closely related, as effective risk management involves identifying potential threats and vulnerabilities to an organization's information assets and developing strategies to mitigate those risks.
3. Human Factors and Security: Human factors, such as employee behaviour and training, can have a significant impact on the effectiveness of security measures. For example, a poorly trained employee may inadvertently compromise a system, or an employee with malicious intent may intentionally bypass security controls.
4. Interagency Cooperation: Effective security often requires coordination and cooperation between multiple agencies or organizations. For example, law enforcement agencies may need to work with private security firms to investigate and prevent cybercrime.

Understanding the interrelationships in security is important for developing effective security strategies that take into account the complex and interconnected nature of security threats. It requires a holistic approach that considers all aspects of security, from physical security measures to cybersecurity controls to human factors and interagency cooperation [166].

In the context of cybersecurity, vulnerability interrelationship refers to the way in which different vulnerabilities in a system can interact with and compound one another, leading to greater risks for the security of that system. For example, a system may have multiple vulnerabilities, such as outdated software, weak passwords, and inadequate access controls. If these vulnerabilities are not

addressed, they can create opportunities for attackers to exploit the system, potentially leading to data breaches, unauthorised access, and other security incidents [168]. Moreover, vulnerabilities in one system can also impact the security of other interconnected systems, creating a ripple effect of risk. For instance, if an organization's email system is compromised, the attackers may use that access to infiltrate other systems within the organization or even spread malware to other organizations through phishing emails [165]. Understanding the interrelationship between different vulnerabilities in a system is crucial for effective cybersecurity risk management. By identifying and addressing multiple vulnerabilities simultaneously, organizations can reduce the likelihood and impact of a security incident. This can involve implementing layered security measures, conducting regular vulnerability assessments, and providing ongoing cybersecurity awareness training to employees [169].

In biology, interrelationship refers to the way in which different living organisms interact with one another and with their environment. The study of these interrelationships is known as ecology, and it is concerned with understanding how organisms interact with one another, with their physical environment, and with the larger ecosystems of which they are a part. Some examples of interrelationships in biology include predator-prey relationships, mutualistic symbioses, and competition for resources. In a predator-prey relationship, for instance, the predator and prey species are interdependent: the predator relies on the prey as a food source, while the prey must avoid being eaten in order to survive. Mutualistic symbioses, on the other hand, involve two or more species that benefit from their interactions with one another. For example, bees and flowers have a mutualistic relationship, where the bees pollinate the flowers in exchange for nectar [170]. Table 6 in appendix illustrates the relationship between security threats and vulnerabilities. It shows that a security threat can have multiple vulnerabilities, while a vulnerability can be exploited by multiple security threats. For instance, the security threat of Data Breaches involves five different vulnerabilities, including Insecure interfaces and APIs (V1), Data-related vulnerabilities (V3), Vulnerability in Virtual Machines (V4), Vulnerabilities in Virtual Machine Image (V5), and Vulnerabilities in Virtual Networks (V7). An attacker can exploit these vulnerabilities using various techniques such as brute-forcing, measuring cache usage, and load-based co-residence detection to collect data from cloud systems. Thus, the occurrence of a data leak not only depends on data-related vulnerabilities but also on virtualization vulnerabilities [107][108].

To summarize a security threat is a potential attack that may or may not happen, whereas a

vulnerability is a flaw that can be exploited by hackers. The table illustrates that a security threat may have multiple vulnerabilities, and one vulnerability may be exploited by several security threats. For example, the Data-related vulnerability (V3) is involved in three security threats: Data Breaches (DB), Weak Identity, Credential and Access Management (IAM), and Data Loss (DL). An attacker may use various techniques such as SQL injection and cross-site scripting to exploit the Data-related vulnerability and launch a Data Breach attack on a cloud system. Similarly, an attacker may leverage the data stored in clear plain text to gain access to a cloud system and launch a Weak IAM attack. Finally, an attacker may exploit different vulnerabilities to cause Data Loss, such as different located data, incomplete data deletion, and data backup vulnerabilities [107]. You can find some examples in appendix (Table A.1).

Next, our plan is to analyse various types of vulnerabilities by collecting and examining data and information. We will explore the interrelationships between vulnerabilities using vulnerability trees, which illustrate the hierarchical relationships between vulnerabilities and the steps that a threat actor might take to exploit them and move up the tree.

As we mentioned earlier, a vulnerability refers to flaws in computer programs that can be exploited by attackers to gain unauthorised access, manipulate or steal data, or disrupt system operations [154]. The security of all data and services is at risk when vulnerabilities exist in the operating system components such as the kernel, system libraries, and application tools. The number of reported vulnerabilities for networked systems has indeed grown significantly in recent years, posing a major challenge to security personnel. With the increasing prevalence of networked systems and the growing sophistication of cyber threats, there are more potential vulnerabilities and attack surfaces that need to be defended against [156].

One reason for the increase in reported vulnerabilities is the growing complexity of networked systems. As systems become more complex, they become more difficult to secure, and vulnerabilities may be introduced unintentionally during development or maintenance. Additionally, the increasing use of third-party software components and open-source libraries can also contribute to the presence of vulnerabilities, as these components may have their own security weaknesses [157]. Another reason for the increase in reported vulnerabilities is the growing sophistication of cyber-attacks. Attackers are constantly developing new and more advanced methods for exploiting vulnerabilities, and security personnel must work hard to keep up with these evolving threats. As a result, new vulnerabilities are discovered and reported on a regular

basis, and security personnel must be vigilant in their efforts to identify and address these vulnerabilities [158].

To address the challenge of growing numbers of reported vulnerabilities, security personnel must adopt a comprehensive and proactive approach to security. This includes conducting regular vulnerability assessments and penetration testing, implementing strong access controls and network segmentation, and keeping software and systems up to date with the latest security patches and updates. Additionally, organizations should prioritize security training and awareness among their staff and invest in the latest security technologies and tools to help identify and defend against emerging threats [156]. Furthermore, a methodology is needed to organize potential attack scenarios, understand their relationships, and assess their level of risk. This technique, known as “vulnerability trees”, provides a structured and flexible means of conducting security analysis of networks, protocols, and applications. While "fault trees" have been widely used as a system analysis technique, vulnerability trees are specifically designed for analysing security threats [120].

3.2.1.1 Fault Tree

Fault trees are a graphical representation of the various potential causes and events that can lead to a particular failure or incident. Fault trees are commonly used in engineering, safety, and risk management to identify potential causes of failure and to evaluate the likelihood and consequences of each possible scenario [159]. A fault tree typically consists of a series of logical gates, nodes, and events. The top event is the undesired outcome that is being analysed, such as a system failure or an accident. The events that led up to the top event are represented by intermediate events, which can be either basic or intermediate events. Basic events are the lowest level events that cannot be further decomposed, such as a component failure or a human error. Intermediate events are logical combinations of basic events or other intermediate events that contribute to the top event [160].

Fault trees can be useful in analysing potential cybersecurity incidents or vulnerabilities. By identifying the events or conditions that could lead to a particular incident, security professionals can better understand the risks and implement appropriate countermeasures. For example, a fault tree for a potential cyber-attack might start with the top event of a successful attack on a company's network. Intermediate events might include things like a phishing email being opened, a vulnerability in a software application being exploited, or a weak password being used. Basic

events might include specific details about the phishing email, or the vulnerability being exploited, such as the email containing a malicious attachment, or the vulnerability being related to a particular line of code [159].

By analysing the fault tree, security professionals can identify specific areas where countermeasures can be implemented to reduce the likelihood of a successful attack. For example, they might focus on training employees to identify and avoid phishing emails, implementing two-factor authentication to reduce the risk of weak passwords being used, or implementing software updates and security patches to address known vulnerabilities. It can also be used to evaluate the effectiveness of existing security measures. By identifying the potential paths that an attacker could take to successfully breach a network, security professionals can evaluate the existing controls in place and determine if they are effective in reducing the likelihood of a successful attack. If a particular control is found to be ineffective, the fault tree analysis can help identify alternative approaches that might be more effective. Overall, fault trees, which are commonly used in system analysis, only show chronological orderings of events over time and are not effective in visualizing and modelling the various relationships between vulnerabilities. The fault tree's levels only provide more detail on the same thing, rather than representing a new perspective or aspect of the problem [91].

3.2.1.2 Vulnerability Tree

A cyber vulnerability tree is a type of fault tree that is specifically focused on identifying and analysing potential vulnerabilities in a computer network or system. The tree starts with a top event, such as a data breach or system compromise, and works backwards to identify the underlying events or conditions that could contribute to that event. Intermediate events in a cyber vulnerability tree might include things like software vulnerabilities, weak or misconfigured passwords, unpatched systems, or social engineering attacks. Basic events might include specific details about the vulnerabilities, such as the presence of a known vulnerability in a particular software application or a password being stored in plaintext. By analysing the cyber vulnerability tree, security professionals can identify specific areas of weakness in their network or system and develop strategies for mitigating those vulnerabilities. For example, they might focus on implementing more secure password policies, regularly patching and updating software, or implementing additional security controls like firewalls or intrusion detection systems [161].

One advantage of using a cyber vulnerability tree is that it provides a visual representation of the

potential pathways that an attacker could take to exploit vulnerabilities in a network or system. This can help security professionals identify and prioritize the most critical vulnerabilities and develop more effective strategies for mitigating them. Additionally, the use of fault trees and similar tools can help organizations meet compliance requirements for security and risk management, as they provide a structured approach to identifying and addressing potential vulnerabilities [162]. As Figure 3-1 shows, vulnerability trees have modelled as hierarchy trees that the top hierarchy presents the main goal of the attacks. Vulnerability trees can be exploited by the top vulnerability, known as parent vulnerabilities (it has symbolized with a capital 'V'). There are numerous ways that such a top vulnerability can be exploited. Each of these ways has constituted a branch of the tree. Each of these ways can be considered as one of the branches of the tree and leaves can constitute child vulnerabilities (they have symbolized with the lower case 'v') [20][57]. These models usually demonstrate the relationship between one vulnerability and other vulnerabilities and/or steps that a threat agent should carry out in order to achieve the top of the tree [57].

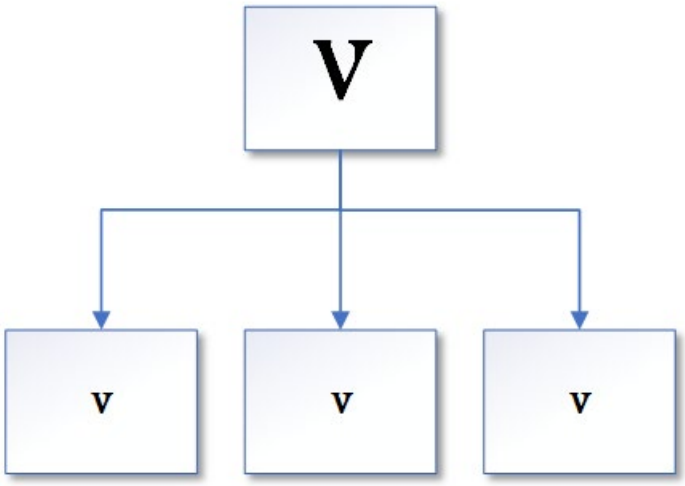


Figure 3-1. Vulnerability Tree

Vulnerability trees have several advantages over fault trees in modelling vulnerability relationships. They help to focus the analysis on specific measurable goals that can be tested in the real world, rather than just providing a chronological ordering of events. This approach avoids the pitfall of overly academic research that does not consider the practical difficulty of conducting attacks and measuring their impact [20]. Vulnerability trees also promote structured elaboration of events, which helps to identify all possible avenues of approach for an attack and optimize the

deployment of countermeasures. Moreover, attack trees allow security analysis to be conducted at multiple levels of abstraction, by decomposing each node into subgoals, which can be helpful in focusing on specific areas of interest. Finally, using attack trees allows common attacks to be referenced as reusable modules that apply to multiple network scenarios [120].

Therefore, the clearest way to demonstrate vulnerabilities inter-relationship is that analyse dataset by considering vulnerability tree, to serve this purpose we try analysing data set by Machin Learning and Natural language processing.

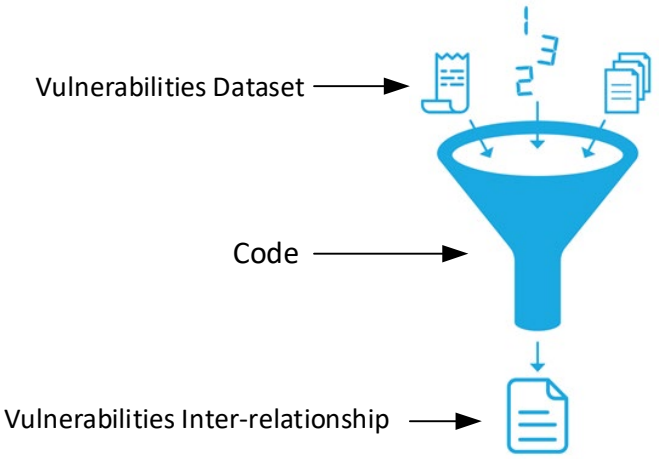


Figure 3-2. Process of Analysing the Vulnerabilities Dataset

Figure 3-2 shows the process of analysing the vulnerabilities dataset. In this study we be able to find a vulnerability tree this turn to help us to find interrelationship between vulnerabilities.

3.3 Bio-Cyber Operations

Bio-Cyber operation is a new field of research that is inspired by the Human Immune System. The human body has found solutions for problems that cybersecurity professionals have been trying to resolve for the past few decades. Cybersecurity should draw lessons from the human immune system on how to detect and deter attacks. Systems and devices are likely to leak sensitive information or data. A ‘cyber immune’ technology can be used to detect unknown cyber-attacks and provide a powerful mechanism for defence. In this paper we focus on work that describes the recent advances on Bio-Cyber operations, and we present our conceptual cyber operations model. By looking into the field of human biology we aspire to provide significant insight into the bio-cybersecurity domain. In this chapter we review some related work.

3.3.1 Why Bio-Cyber Operation is Important?

Nowadays, most devices and computers are connected to the Internet, and cybersecurity can be used against threats and attacks immediately, even over long distances can be executed. Cyberattacks can pose a very serious risk to devices such as machines, smart sensors, or even the whole infrastructure of sites that communicate through the Internet, for instance, zero-day exploits or Advanced Persistent Threats (APT). Industry 4.0, or The Internet of Things (IoT) where cyber-physical systems cooperate and communicate with each other, can expose systems or devices that often utilize firmware or legacy codes that were not originally developed with cyber threats in mind. These facilities were often constructed or designed without considering the ability to communicate, furthermore no security measures were applied. Connectivity was often added or implemented after they were built, and, exposing them to cyber-threats. The necessity to open the production sites which previously closed provides a whole new attack surface for cyber criminals, therefore the need for effective tools to against them. The amount of data created on the internet is increasing day by day and analysing and detecting it for any threats has become a serious challenge. The IoT has about millions of connections, communicating and exchanging data which are generated by using their sensors every single second, all over the world. They combine physical devices and systems with the internet and create and form a cyber environment. Any cyberattacks can pose a serious and an uncompensated threat or risk to the security of any cyber infrastructures or the reliability of them [20]. Conventional cybersecurity systems like classic firewalls, end-point protection or intrusion detection systems are not sufficient, since they are not able to detect new, or unknown vulnerabilities. To meet today's cybersecurity requirements, a cyber defines system demands the capability to identify and repel new or unknown attacks and adapt to the new threats without disrupting the systems or devices. As we do not have the ability to know or recognize all vulnerabilities, one of the strategies against cyber threats is we can apply is to learn attack patterns [3].

3.3.2 Human Immune System Factors

The human body has a significantly efficient defence mechanism, the immune system (IS), which has the capability to detect a wide range of harmful agents, named pathogens, such as microbes, viruses, and parasites. It has the ability to distinguish healthy tissue from pathogens. The skin also defends against external threats to our body, similar to a firewall. The immune system is constantly

adaptive and renewed. IS always monitors and checks the internal environment. In the absence of a working IS or If it does not work properly, even a minor infection can take hold and pose fatal. [4]. The IS includes an innate immune system and an adaptive immune system (AIS). Innate immunity is found in both invertebrates and vertebrates, while the adaptive immune system is present in invertebrates [6]. If a pathogen breaches the body, the innate immune system confers an urgent but non-specific response. However, it is not able to provide long-lasting immunity. If the innate immune system cannot fight against a pathogen, then the AIS will be activated. The AIS can produce a proper response for the pathogen specific. This specific antigen and its response are remembered if the same pathogen enters the body next time. The immune system has the ability to learn, memory as well as pattern recognition [5]. Likewise, the AIS behaves as the memory of the immune system. The AIS consists of lymphocytes, which are the major types of B-cells and T-cells. B cells can recognize pathogens when antibodies on their surface attach to a special foreign antigen. When an antigen enters the body, the IS provides and activates signal molecules that absorb specialized immune cells, named T cells or killer T cells, to the site of infection. The killer T cells are able to destroy those cells that are affected or infected by viruses, other pathogens, or other dysfunctional cells. B cells and T cells are activated, they start to replicate, and their effectiveness (or some of their offspring) becomes long-lived memory cells. A cyber immune system can imitate and be inspired by this behaviour.

3.3.3 Basics about Information Security Attributes

2020 has been an unprecedented time in our lifetime. Cybersecurity incidents are on the rise, with threat agents targeting crisis-related assets. We speculate that threat agents are taking advantage of new opportunities that have been created due to the changes in the operational framework of companies around the world. Threat agents are those individuals or groups of people that can manifest a threat [57]. The notion of threat can be defined as the function of a threat agent exploiting a vulnerability of an asset. An asset is a thing (or a service) that either now or in the future can generate or has the potential to generate (directly or indirectly) revenue for the business [57]. Today, our society is dependent on the Global Information Environment (GIE). Cyber-attacks can happen instantly. Modern threats, for example zero-days and APT pose very serious risk to the GIE. The Internet of Things, where cyber-physical systems cooperate and communicate with each other, presents devices and systems that usually use firmware or legacy code that was not improved with cyber threats in mind [27]. As we continue to embrace and use all the benefits of

cyber operations, we need to move to a more secure world that focuses on behaviours within a network to identify normal behaviours from abnormal behaviours both at the group and individual level.

The aim of trying to "secure" the information is not new. In order to have a defence chance, cyber operation, just like the human body, must be defended through focusing and understanding the information infrastructure that is at risk at any one time. To serve this purpose, we need to start to implement a cyber immune system that acquires from its environment to avoid repeating attacks or problems and combat new ones [43].

In this chapter we cut across scientific boundaries, discussing correlations between cybersecurity notions and concepts with biological notions and concepts. We draw inspiration from the human immune system, and we present a conceptual model for cyber operations.

3.3.4 The Global Information Environment

The aforementioned concepts are not new. What is new, and what is constantly changing, is the environment into which we, as individuals, and of course businesses have to operate [10]. The US DoD has defined the Information Environment as "the aggregate of individuals, organizations and systems that collect, process disseminate or act on information [51]. The global information environment has some characteristics that uniquely describe it. It is by far a non-static environment, that is making use of change as a catalyst in order to address the business needs (client needs) for constantly generating revenue streams. It can consist of a plethora of devices, using diverse mobile architectures. Furthermore, it is hyper-charged in that every device is performing several roles, offering services to other devices, all related to strangeness [20].

Due to technological advances, communication is being exchanged at unprecedented rates. According to recent estimates, the amount of data created globally each day is closer to 59 zettabytes, which is equivalent to 59 trillion gigabytes. This is a significant increase from the previously cited figure of 2.5 quintillion bytes [141]. The modern information environment attempts (and not with great success) to constantly change in order to enable the aforementioned communication to take place [5].

Any member (or operator) of the environment can produce and consume information, often at the same time. The average decision maker today has access to platforms allowing for capabilities that until the recent past were limited to specific sections of the defence community [30]. In addition, everyone wants to get involved, but more of the point, there is the expectation that everyone will

be allowed to be involved in every decision originating from the environment, and for the environment.

Finally, almost everyone listens. It is hard to define boundaries (internal or external) to the information environments. Ownership and responsibilities can be two grey areas.

The Information Environment is not the only concept that has become global in the current computing evolution era. Cyber Operators have also become global [51]. Instead of looking at local systems and local environments, cyber operators are now truly operating across geographical boundaries, cutting across jurisdictions, having to manage and protect complex interrelationships between tangible and intangible assets [63]. We argue that this also applies to cyber as well as kinetic operations.

3.3.5 The Challenge of the GIE

The same concepts that were discussed as attributes of the GIE in a previous section, can also be seen as challenges that global cyber operators are trying to overcome. The main objective of any business is the generation of revenue. From the perspective of a cyber operator, an increase in revenue will be the by-product of achieving information superiority for their organization. In agreement to [46], information superiority is the operational advantage derived from the ability to collect, process and disseminate an uninterruptible flow of information, while exploiting or denying an adversary to do the same. Information superiority is a state achieved as the result of successful information operations [12].

Information operations are continuous acts of force in the GIE [51], to compel our adversaries to do our will. We classify information operations in the following types:

Intelligence Operations (Inc counterintelligence) Cyber intelligence is a cyber discipline that exploits a wide range of information collection and analysis methods to provide decisions and direction to cyber operation units and cyber commanders [3].

3.3.6 Cyber Operation

A cyber operation is a coordinated and targeted activity conducted in digital space for the purpose of achieving a specific objective. These objectives can range from stealing data, disrupting systems, or even causing physical damage. Cyber operations can be carried out by various entities, including nation-states, criminal organizations, and hacktivist groups. They can also take many forms, such as hacking, phishing, social engineering, or distributed denial of service attacks [12].

It often involves the use of advanced technologies and techniques, such as malware, encryption, and data exfiltration tools. These tools can allow attackers to gain access to target systems, exfiltrate sensitive information, or cause disruption or destruction. To carry out a cyber operation, attackers typically go through a multi-step process that includes reconnaissance, exploitation, and exfiltration [13]. During the reconnaissance phase, attackers gather information about the target's systems, vulnerabilities, and defences. During the exploitation phase, attackers use this information to gain access to the target's systems or data. Finally, during the exfiltration phase, attackers steal or exfiltrate the data they are interested in [14]. Cyber operations can have serious consequences, both for the targeted organizations and for society as a whole. They can lead to financial losses, data breaches, and even physical harm or loss of life. As a result, it is important for organizations and individuals to take steps to protect themselves against cyber-attacks, such as implementing strong cybersecurity measures, staying vigilant for suspicious activity, and regularly backing up important data [15].

3.3.7 Psychological Operations

Psychological operations that transport information, for instance broadcasting satellite radio messages, with the object of manipulating the views of organizations, foreign governments, or individuals. Psychological operations, also known as psyops, are defined as the planned use of propaganda, deception, and other techniques to influence the beliefs, emotions, and behaviour of a target audience. These operations can take many forms, including the use of satellite radio messages [8].

Satellite radio messages can be used as a means of broadcasting propaganda or other forms of persuasive messaging to a wide audience, with the aim of influencing their attitudes or behaviour. For example, a government or military organization might use satellite radio messages to disseminate disinformation or propaganda to foreign populations in order to undermine the credibility of a particular government or ideology [9].

It's important to note that psychological operations can be used for both defencive and offensive purposes. Defencive psyops may be used to protect a government or organization from threats by disseminating accurate information or correcting false rumors. Offensive psyops, on the other hand, may be used to influence the opinions or behaviour of an adversary in a conflict situation.

Overall, psychological operations can be a powerful tool for shaping public opinion and influencing behaviour. However, they can also be controversial and raise ethical concerns,

particularly when used to deceive or manipulate people for political or military purposes [8].

3.3.8 Deception

Cyber deception refers to the use of techniques or technologies to mislead or deceive attackers attempting to breach a network or system. Cyber deception can be a proactive defence mechanism that aims to divert or delay attackers, gather intelligence on their tactics, and ultimately prevent or minimize the impact of a cyber attack [8]. Deception usually uses computer networks and information technology to intentionally deceive adversary decision-makers as friendly military abilities, deliberate and operation, so the adversary takes an action in ways that aid to friendly forces' mission [24].

3.3.9 Computer Network Operations

Computer Network Operations (CNO) refer to the use of computer networks and related technologies to conduct various activities that are aimed at achieving a specific objective. These activities can range from simple tasks such as network management and monitoring to more complex activities like cyber-attacks and network exploitation [100].

CNO can be used for a wide range of purposes, including:

1. **Network Defence**: This involves using CNO to monitor and protect networks from unauthorised access, cyber-attacks, and other security threats [101].
2. **Network Exploitation**: This involves using CNO to gather information from target networks, such as data on users, applications, and hardware [102].
3. **Cyber Warfare**: This involves using CNO to conduct offensive operations against enemy networks, such as conducting cyber-attacks to disrupt or disable critical infrastructure [103].
4. **Cyber Intelligence**: This involves using CNO to collect and analyse information on the activities of potential adversaries, including their capabilities, intentions, and vulnerabilities [104].

CNO is used by various entities, including government agencies, military organizations, and private companies. The use of CNO raises various legal and ethical issues, including privacy concerns, international law, and the use of force in cyberspace [100].

3.3.10 Computer Network Attack

A Computer Network Attack (CNA) is a type of offensive operation that aims to disrupt, damage, or destroy computer systems or networks. This can be achieved through a variety of techniques,

including malware, denial of service attacks, hacking, and social engineering. CNA is often carried out by hackers, cybercriminals, and state-sponsored actors [105]. The goals of CNA can vary widely, depending on the motivations of the attacker. For example, some attackers may seek to steal sensitive data or intellectual property, while others may aim to disrupt critical infrastructure or cause harm to an organization or nation [106]. CNA can have serious consequences, including financial losses, reputational damage, and even physical harm in the case of attacks on critical infrastructure. As a result, organizations and governments around the world invest significant resources in defending against CNA and responding to attacks when they occur [107].

3.3.11 Computer Network Defence

Computer Network Defence (CND) refers to the actions and processes taken to protect computer systems and networks from unauthorised access, exploitation, and attack. CND is an essential part of cybersecurity and involves a range of activities, including network monitoring, threat analysis, incident response, and vulnerability management [108]. CND strategies typically focus on preventing, detecting, and responding to cyber threats. Prevention measures can include firewalls, intrusion prevention systems, and access controls, while detection techniques may involve network traffic analysis, behaviour monitoring, and threat intelligence. If a cyber-attack is detected, CND teams will typically respond by isolating affected systems, identifying the source of the attack, and taking steps to remediate any damage. This can involve patching vulnerabilities, removing malware, and restoring backups [109]. CND is a critical component of modern cybersecurity, as cyber threats continue to grow in frequency, sophistication, and impact. By implementing robust CND strategies and processes, organizations can minimize the risk of cyber-attacks and protect their sensitive data, systems, and networks [108].

3.3.12 Computer Network Exploitation

Computer Network Exploitation (CNE) refers to the practice of using hacking techniques and other cyber-attacks to gain unauthorised access to computer systems and networks for intelligence gathering or other malicious purposes. CNE is typically carried out by government agencies or other state-sponsored groups, although it can also be conducted by criminal organizations or individuals [110]. CNE can involve a wide range of techniques, including phishing, malware attacks, password cracking, and social engineering. The goal of CNE is often to gain access to sensitive information, such as classified government data, corporate trade secrets, or personal

information about individuals [111]. CNE is often used as part of a larger strategy of cyber espionage or cyber warfare, in which nations or other groups attempt to gain advantage over their rivals through covert cyber operations. However, CNE can also be used for more mundane criminal purposes, such as stealing financial information or intellectual property [112]. CNE is a highly controversial practice, as it often involves violating the privacy and security of individuals and organizations without their knowledge or consent. However, it is also seen as a necessary tool for national security and law enforcement and is therefore a subject of ongoing debate and controversy [111].

3.3.13 Situational Awareness Operations

Situational awareness is the understanding of environmental events and elements with regard to space or time, the perception of their meaning, and the plan of their future status [22].

Situational Awareness Operations (SAO) refer to the collection, analysis, and dissemination of information to provide a comprehensive understanding of a particular situation or environment. SAO is used in various domains such as military, law enforcement, emergency response, and cybersecurity to enhance decision-making and improve operational effectiveness [113].

SAO involves the following key elements:

1. Collection: Gathering data from various sources, including sensors, intelligence reports, social media, and other open-source information.
2. Analysis: Processing and analysing the collected data to identify patterns, trends, and potential threats.
3. Visualization: Presenting the analysed data in a way that allows decision-makers to quickly and easily understand the situation and make informed decisions.
4. Dissemination: Sharing the analysed information with relevant stakeholders in a timely and secure manner [114].

SAO is a critical component of many operations, as it helps to identify potential threats, assess risks, and develop effective response strategies. For example, in military operations, SAO is used to monitor enemy movements and activities, identify vulnerabilities, and provide commanders with real-time situational awareness. In law enforcement and emergency response, SAO is used to monitor public safety incidents, track suspects, and coordinate response efforts [22].

In cybersecurity, SAO is used to monitor networks and systems for potential threats, such as malware infections or unauthorised access attempts, and to respond quickly to any incidents that

may occur. By maintaining a high level of situational awareness, organizations can better protect themselves against cyber-attacks and other security threats [114].

3.3.14 Operational Security

Operational security (OPSEC) which is known as procedural security, is a risk management operation, that encourages managers to take in the process from the vision of an opponent in order to protect sensitive data and information from the leak into the wrong hand [35].

Operational Security is the process of identifying, analysing, and controlling critical information that could be exploited by adversaries. OPSEC is used to protect sensitive information, prevent unauthorised access, and maintain operational effectiveness across various domains such as military, law enforcement, business, and personal security [115].

OPSEC involves the following key elements:

1. Identification of critical information: Determining what information needs to be protected and what potential adversaries are interested in that information.
2. Analysis of threats: Identifying and assessing potential threats to critical information, such as espionage, social engineering, or cyber-attacks.
3. Development of countermeasures: Implementing measures to prevent or mitigate the identified threats, such as encryption, access controls, or security awareness training.
4. Evaluation of effectiveness: Continuously evaluating the effectiveness of the countermeasures and adjusting them as necessary to ensure the protection of critical information [116].

OPSEC is used in various applications, such as military operations, where it is used to protect sensitive information such as troop movements, tactics, and other operational plans from the enemy. In law enforcement, OPSEC is used to protect sensitive information such as witness identities or investigative techniques from criminals. In the business world, OPSEC is used to protect proprietary information, such as trade secrets, from competitors. OPSEC is essential to maintaining operational effectiveness and protecting critical information. It is an ongoing process that requires continuous evaluation and improvement to ensure that threats are identified and mitigated effectively [117].

3.3.15 Information Security

Information security is a set of activities that try to keep data and information secure from unauthorised alterations or access [118]. Information Security (InfoSec) refers to the practice of

protecting sensitive and confidential information from unauthorised access, use, disclosure, disruption, modification, or destruction. Information Security is concerned with maintaining the confidentiality, integrity, and availability of information [119].

Information Security involves the following key elements:

1. Confidentiality: Ensuring that sensitive information is only accessed by authorized individuals and not disclosed to unauthorised parties.
2. Integrity: Maintaining the accuracy and completeness of information and preventing unauthorised modification or deletion.
3. Availability: Ensuring that information is available to authorized individuals when needed and not disrupted by unauthorised parties.
4. Non-Repudiation: Ensuring that the origin and integrity of information can be verified, and that individuals cannot deny their involvement in a transaction.
5. Authentication: Verifying the identity of individuals accessing sensitive information or systems.
6. Authorization: Granting individuals access to sensitive information or systems based on their level of clearance or need-to-know [120].

Information Security is a critical aspect of any organization's overall security strategy, as it protects sensitive information from theft, unauthorised access, or misuse. Information Security is particularly important in the digital age, where information is often stored electronically and can be easily accessed and shared across multiple devices and networks [118]. Common Information Security measures include encryption, access controls, firewalls, intrusion detection systems, and security awareness training. Information Security is an ongoing process that requires regular assessments, updates, and improvements to ensure that information remains secure in an ever-changing threat landscape [119].

3.3.16 Cyber-Physical Security

Cyber-Physical security is the activity of personnel, data, networks, software, and hardware from physical action from physical threats such as theft, vandalism, sabotage, terrorism, or natural disasters [121]. It encompasses a wide range of measures and systems designed to protect the physical environment and prevent unauthorised access, damage, or disruption to an organization's operations. Physical Security is an important aspect of an organization's overall security strategy and is essential to ensure the safety of personnel and protection of assets and event that could

reason serious loss or damage to an agency, enterprise, or institution [30].

It is not our intension to discuss the above operation types in isolation. Instead, we will discuss their correlations and how we have used the human immune system for modelling interrelationships between operations, aiming to better defend an information environment.

3.3.17 Human Immune System Correlated to Cyber Operations

Cyber Operations (Cyber-Ops) is an interdisciplinary field encompassing the whole scope of cyberspace and related activities that are both technical and non-technical in nature, for instance ethical, legal, human-centered, etc. Cyber Operations is a supplementary subject to Cybersecurity. Cyber Operations places special emphasis on techniques and technologies applicable to all system and operational levels [54]. If cyber operations are compared to the human body, then cyber-attacks can compare to viruses. The human body has a significant effective immune mechanism called the immune system which can protect and detect a wide range of harmful agents, such as microbes, viruses, and parasites which are known as pathogens [48]. The human immune system includes special organs, cells, and chemicals that can fight pathogens or any infection. The main part of the immune system is made up of blood cells, antibodies, the complement system, the lymphatic system, the spleen, the thymus, and the bone marrow, they are an internal part of the immune system. However, the external part of our immune system is the skin, which fends off external threats similar to a firewall. It is continuously adaptive and renewed. Our immune system monitors the internal environment regularly and constantly [50]. In general, the immune system contains an adaptive immune system as well as an innate immune system. Vertebrates and invertebrates have innate immunity whereas the adaptive immune system is found only in invertebrates [50]. If a pathogen breaks physical barriers of the body, the innate immune system presents an immediate response, but it is a non-specific response, and it is not able to confer long term immunity. The adaptive immune system can present a pathogen specific, tailored response. If the same pathogen enters the body, the response is remembered, and the immune system can present a quick and specific response to the antigen. The immune system has the ability to learn, to remember, and to identify patterns. Likewise, the adaptive immune system acts as the memory of the immune system [48]. The adaptive immune system includes the lymphocytes which consist of significant types of B-cells and T-cells. B cells recognize pathogens when antibodies on their outside connect to a certain foreign antigen. When an antigen enters the body, the immune system sends a signal to direct specific immune cells, known as killer T cells, to the infection's site. The

killer T cells annihilate cells that are affected by viruses or any other pathogens and dysfunctional cells. An adaptive immune system recognizes and neutralizes antigens (think of ransomware or Trojans as antigens) by erasing or quarantine them. This function is similar to the biological systems' function which kills cells affected by known or unknown viruses [49]. In case B cells and T cells are activated, they reproduce, and some of their children become long-lived memory cells. A cyber immune system acts very similar to this behaviour.

3.3.18 Bio Inspired Cyber Security

All living creatures take advantage of the wide range of natural forms of protection to help them to reduce as well as to avoid possible risks and adapt to their environment for survival. These biological predispositions and instincts, which are different among living creatures, can be synthetically implemented and replicated to the cyber immune systems in order to increase the system’s resilience when facing an attack [37]. This method of cybersecurity (bio-inspired cybersecurity) can be categorized by the weaknesses in security systems such as transparency, bioinformatic analysis, security standards, and cryptography [59].

3.3.19 Transparency

The Ant-Based Cyber Defence (ABCD) [34] is a bio-inspired method that monitors and defends computer networks by adapting the algorithms and metrics by patterning the social insects, especially ants. ABCD is such a society in which humans and autonomous adaptive software agents cooperate [34]. In the changing attackers' strategies, ABCD can supply fast, stable adaptation and a dynamic environment. The ABCD software uses a biological concept known as swarm intelligence to implement. Swarm intelligence not only affected making of decisions but also allows ants to communicate and gathering in a risk's instance in a colony. ABCD utilizes a hierarchy of digital instruments, which are looking for harmful activity by comparing different machines in the system [40].

3.3.20 Bioinformatic Analysis

Bioinformatic analysis transform cycles of behaviours or instructions to the data set for each same metrics that may be estimated by applying protein-sequencing tools [53]. Similarity can end up in the interrelationship groupings, description, and fast recognition of beforehand hidden suspect behaviours or attack designs. Similarity can end up in the interrelationship groupings, description, and fast recognition of beforehand hidden suspect behaviours or attack designs. Bioinformatic

analysis is applied to an attack on a behavioural motif that has been seen before and to help determine how the malware can be neutralized [26].

3.3.21 Security Standards and Metrics

In general, hierarchy trees imitate a real tree structure. These structures are widely used, as they are associated with computer science, they can be used to explain system processes where relationship paths exist [37].

3.3.21.1 Attack Tree Models

In recent years, attack trees have been significantly developed to identify processes by which attackers or malicious users try to break or exploit computer software and/or network. An attack tree is a conceptual diagram indicating how a target, or an asset, might be attacked. Attack trees are extensively applied to threat scenarios in a concise and intuitive manner, which is suitable to express security information to non-experts [41]. As Figure 3-3 illustrates, which can be considered as multi-level diagrams and inspire the relationship between different parts of the real tree for example root, leaves, and children. The top of the tree is considered a top attack [55]. From the bottom up, child nodes are the status that should be satisfied to construct the direct parent node true; when root is content the attacks are complete. These models are usually used to experiment with control centers and communication devices. In order to recognize and countermeasure against modelled viral and Internet attacks, they improve cybersecurity systems by producing flexibility in facing future attacks [42].

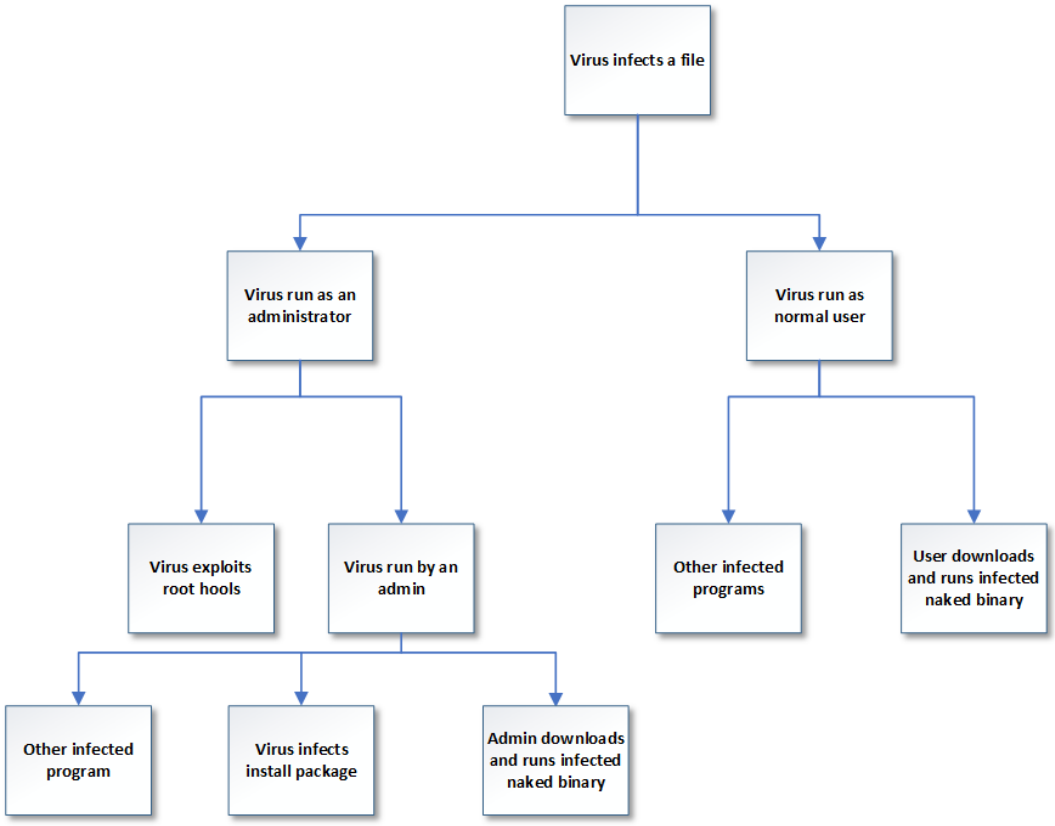


Figure 3-3. Attack Tree

3.3.22 Cyber Immune System

Cyber immunity is a bio-inspired method based on the human adaptive immune system that is able to learn and recognize attacks with unknown signatures [11]. Cybersecurity can be defined as a combination of technologies and processes designed and produced to protect and defended networks, computers, data, and programs from possible attack or risk, unauthorised access, variations, or any destruction. The DNA of a virus changes continuously, so the immune system should adapt to detect the signature of the virus. Likewise, in cybersecurity, we face an ever-evolving adversary. Due to new attacks being unknown, it is very difficult to determine the signature of the attack from previous ones [29]. Bio-inspired cyber immunity is able to learn and recognize the attacks with unknown signatures such as the human adaptive immune system (Figure 3-4) [59]. A cyber immune system tries to learn what is the algorithm of the network traffic during a specific time, instead of learning attack signatures [37]. Once learned, it determines the possibility that a certain abnormal algorithm is malicious. It regularly updates its information and results based on the new observations. It is able to interrupt an attacking agent by surveillance and

recognition what information the cause is and where it comes from. One of the main responsibilities in data mining is to find and select the monitoring points [43].

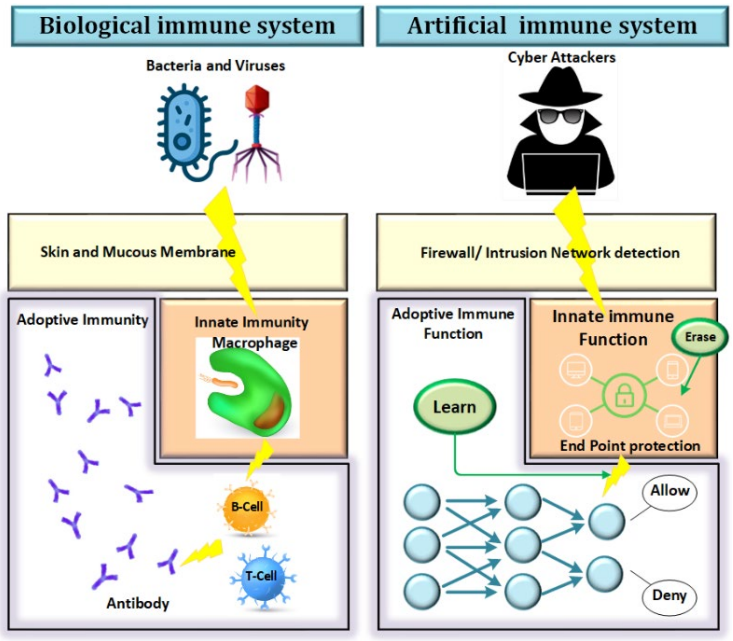


Figure 3-4. Brief comparison of Biological and Artificial Immune System.

It should be noted that the immune system usually considers a metaphor. There is no connection between artificial immune systems and biological in the "mechanisms" of inherent and adaptive Immunity [56].

3.4 Summary

Chapter 3 delves into the complex interrelationships between vulnerabilities in cybersecurity, emphasizing how these vulnerabilities rarely exist in isolation but interact in complex ways to amplify risks. This chapter explains the importance of understanding the interrelationship of various security components, such as the interaction between firewalls and intrusion detection systems, and encryption with authentication. It introduces the concept of vulnerability trees as a visualization tool to analyse how different vulnerabilities may link together and exacerbate security incidents. Furthermore, the chapter draws an analogy to biological systems, particularly the human immune system, to illustrate how cybersecurity systems can dynamically adapt and respond to threats. This bio-cyber operations discussion underscores the need for cybersecurity systems to learn and adapt continuously to address new and evolving threats effectively.

Chapter 4

Linking Biology to Cybersecurity

4.1 Overview

This chapter presents the link biology to cybersecurity. The research used a mixed method, this is a combination of both quantitative and qualitative techniques.

The modern information environment is characterized by various factors, which have been identified and analysed. This includes the detection and identification of potential threats, along with an understanding of how threat agents can exploit systems connected to the internet. Such vulnerabilities can be assessed by testing the system in a controlled environment. As previously noted in a registration report, the aim is to develop a new approach to cybersecurity that draws upon the principles of immunology, where the human body defends itself against unknown threats in a new environment. This approach aims to address existing gaps in the cybersecurity domain by leveraging insights from nature and immunology.

The current speed of security tools is inadequate, which allows intruders a wider window of opportunity to attack systems. Since it takes time for these tools to identify new threats, the impact of attacks can be catastrophic. For example, current IDS/IPS tools cannot identify new attacks without a signature in their database. Moreover, current network security tools are composed of various barriers such as firewalls and signature-based IDS, which are static and require manual intervention from network admins. This design is vulnerable to new threats and becomes obsolete with any changes to network design or assets. As a result, there is an urgent need for new approaches to cybersecurity. Research has shown that the impact of an attack and the resulting

vulnerability on the main network can affect any private or public sector linked to the organization's processes. To address this issue, we plan to identify the weaknesses and strengths of existing security protocols by testing them in an environment designed by us. Based on the outcomes of the testing phase, we will develop a semi-automated defence solution to prevent, deter, and eliminate threats from accessing the system. In addition, we aim to create a semi-automated decision tree based on the immune system design, which identifies the links between nodes and different sections of the network. We will determine how compromising one section can lead to the next possible part of the system getting affected, and we will be able to identify and deter dangerous activity before it reaches vulnerable parts of the system. During our work on the IUK project, we encountered TensorFlow, an open-source library for dataflow programming across a range of tasks. It has a significant advantage over other products due to its symbolic math library, which is useful in machine learning applications such as neural networks.

This study, we plan to work with the TensorFlow environment, which benefits from AI and machine learning technology. With TensorFlow, we will generate a database consisting of interconnection vulnerabilities and train it is using a set of data. The aim will be to identify different platforms, determine the existence of vulnerabilities, their links, and the potential exploitations and outputs of each vulnerability. Furthermore, we have tried to show the relation between biology and cyber security (Table 4.1. indicates some examples of them). These similarities between biology and cyber operations have introduced the field of bio-inspired cybersecurity, and it has been recognized by cyber professionals. Professional recognition of overlap between biology and cyber operation has also emerging in the field of cybersecurity, which can identify blurring of the domain borders between biology and cyber and acknowledges new risks arising from the growingly cyber-physical nature of biotechnology.

Table 4.1. The link between Biology and Cyber-Security [140]

The link Function	IT Infrastructure Action or Term	Cell Biology Action or Term
Barrier defence	Exterior Router Packet Filter/Stateful Inspection Firewalls Intrusion Detection Systems	Plasma Membrane/ Plant cell wall Oligasaccharins “oxidative brust”.
Barrier Transmission and Communication	Tunneling protocols Secure Sockets Layer Virtual Private Networks Advanced Encryption Mechanisms Network Ports	Variety of Member Channels Gap Junctions Facilitated Diffusion and Transporters (i.e., Glucose) Extrasellar matrix signaling.
Internal Organization	Internal Firewalls Network DMZ (Buffer Zones)	Membrane- bound organelles, mitochondria Nuclear pore complex, double membrane envelope.
Internal Routing and Sorting	Email, standard. Fax, telephone IPv6 Routers, routing Table	Endocytosis, Exocytosis Golgi Apparatus Cell Nucleus
Virus Defence and Response	Infection Carrier System Cleansing (anti-virus s/w) System Isolation System Corruption	Infection Carrier Intracellular digestion, endocytosis Cell Division Cell Death

4.2 Bio Inspired Cybersecurity

This section presents Bio Inspired Cybersecurity more closely. All phases of biology and Cybersecurity and the processes of each phase will be described in extent. The main goal of this research is to show all aspects of the relationship between the human immune system and cybersecurity and a deep understanding of bio-Cybersecurity.

Each phase classified into four categories and is presented in the following structure:

- Phase Overview: A high-level explanation of the processes that organized the phase.
- Process Overview: analysis and explanation of the process and the activities that organized the process.
- Process activities: it shows what happens at each step of the process. Included: the description of each activity and a description of the inputs and outputs of each process.

4.2.1 Phase 1: Biology and Immune System

4.2.1.1 Phase 1 Overview

Phase 1 is named "Biology and Immune System. It has five processes. In this phase examine all aspects of Human immune system. This phase examines by acquiring information for Human body [182]. These processes provide a structured and comprehensive framework without being overly complex. Here's how they justify the number:

Process 1 Establishes the foundational knowledge of the immune system's structure and functions.

Process 2 Focuses on the specific mechanisms employed in defence.

Process 3 Introduces the concept of artificial support and its applications.

Process 4 Critically examines the immune system's potential weaknesses.

Process 5 Aims to dispel misunderstandings that can obscure accurate knowledge of the immune system [182].

By identifying and classified this information and put them in assessment, we can present unique research in Bio-Cybersecurity.

Process 1 examines **Human Immune System Analysis**. Process 2 identifies the **Line of Defence**. Process 3 examines **Artificial Immunity** and shows when we must use this system. Process 4 analyses **Vulnerability and the Immune System Response**, and process 5 examines the **Common Mistakes and Misconceptions** in the human immune system.

4.2.1.2 Processes: Human Immune System Analysis

The first process is Human Immune system Analysis, defines the activity of defence system from the human body.

The main goal of this process is to understand the different aspects of the human immune system. The process classified the below category [183]:

- Human Immune system Processes Analysis

- Line of Defence Analysis,
- Vulnerability and the Immune System Response Analysis,
- Common mistakes and misconceptions Analysis.

The following sections examine these activities and their inputs and outputs.

4.2.1.3 Processes Activities:

4.2.1.3.1 Human Immune System Processes Analysis

We first examine the human immune system process by identifying the different aspects of this system. The human immune system can lead us to bearing variable information about the different steps of human defence system and response of it in different situation.

The activity uses the following inputs:

- Current knowledge of human immune system
- Current knowledge of different terms of human immune system

The activity produces the following outputs:

- The human immune system is classified into some different terms which shows a signification information about the human body defence system and the relationship of it with Cybersecurity. This trend is the main concern of researchers in today's world [182].

4.2.1.3.2 Immune System Process Analysis

As Figure 4-1 shows the immune system is composed of various terms that can be identified and analysed. These terms may vary depending on the situation, such as the term "Pathogen" which refers to a disease-causing organism, including bacteria [183]. The human body uses different terms and exhibits varying responses depending on the specific conditions present.

The immune system terms include:

- **Innate Immune System:** The innate immune system represents one of the primary approaches to immunity. In terms of evolutionary history, this system is considered relatively ancient, and it is the predominant immune response mechanism seen in plants, fungi, insects, and early multicellular organisms [70].
- **Adaptive Innate Immune System**: The adaptive immune response, which is also known as acquired or specific immunity, serves as the second line of defence against non-self-pathogens. Unlike the innate immune system, the adaptive immune response is exclusive to vertebrates. This type of immunity is tailored to the specific pathogen encountered, and

it is designed to combat non-self-pathogens. However, there are instances where the adaptive immune response can misidentify self as non-self, leading to attacks on healthy tissues and the development of autoimmune diseases [73].

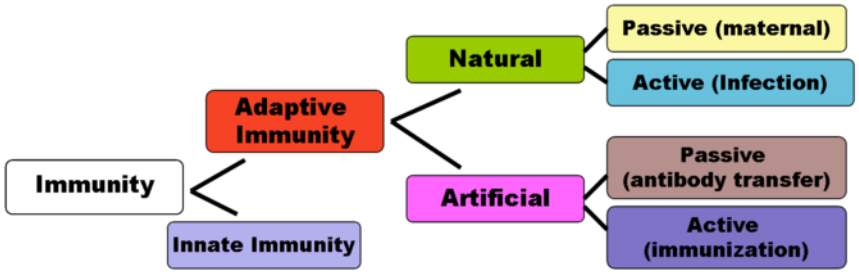


Figure 4-1. Immune System Diagram [182].

The primary activities are:

- The terms "vulnerability" and "attacks" refer to actions related to the reception, storage, and distribution of potential attacks.
- The human body responds to and defends against attacks, processing information about these attacks and transmitting it to the brain.
- The human immune system reacts to attacks by mounting a defence against the invading pathogens, which may involve the activation of various immune cells and the production of specific antibodies to neutralize or eliminate the threat.

The support activities are:

1. The systems involved in planning, controlling, and other related infrastructure are essential to an organization's strategic capabilities across all primary activities.
2. The term "artificial immune system" typically refers to the process of acquiring various resource inputs required for primary activities.

The activity uses the following process:

- Current knowledge of vulnerability and attacks
- Current knowledge of immune system
- Current knowledge of different terms of immune system
- Current knowledge of human body
- Organizational information: The organizational chart, the policies and procedures can be used for acquiring human immune system information.

4.2.2 Phase 2: Vulnerability and Risks in Cyber-Security

4.2.2.1 Phase Overview:

Phase 2, titled "Vulnerability and Risks in Cyber-Security," consists of four distinct processes. During this phase, all aspects of vulnerabilities and risks in systems and devices are examined, and information is gathered to identify weaknesses. This information is then classified and assessed to present a unique research perspective on cybersecurity vulnerabilities.

Process 1 examines **Vulnerability and Risks Analysis**. Process 2 identifies **Vulnerability Assessment**. Process 3 presents a **Definition of System Baseline**, and process 4 illustrates the systems or devices **Response and Reaction.**

The main goal of these process is to understand the different aspects of the Vulnerabilities and risks in Cyber-Security.

4.2.2.2 Process 2.1 Vulnerability and Risks Analysis

Vulnerability and risk in cyber-security refer to weaknesses that can be exploited by a cyber attacker to gain unauthorised access to a computer system. These vulnerabilities may allow attackers to access system memory, execute malicious code, install malware, and potentially damage, modify, or steal sensitive data.

4.2.2.3 Process 2.1 Activities

4.2.2.3.1 Vulnerability and Risks Process Analysis and Objectives

The initial step in examining vulnerability and risks involves identifying the various aspects of weakness that may threaten systems and devices. This process involves a comprehensive evaluation of potential vulnerabilities to determine the specific areas that require additional attention and mitigation measures.

analysing vulnerabilities in cyber-security can lead to the use of diverse information related to the various stages of system responses in different situations. This information can be leveraged to improve the overall security posture of the system and enhance the ability to detect, prevent, and respond to potential cyber-attacks.

The activity uses the following inputs:

- Current knowledge of Weaknesses in Cybersecurity
- Current knowledge of different terms of Cybersecurity

The activity produces the following outputs:

- Cyber-security can be classified into various terms, each of which provides significant information about the system defence. This trend is a primary focus for researchers in today's world, as the increasing prevalence of cyber threats underscores the importance of effective cyber-security measures.

4.2.2.4 Process 2.2 Vulnerability Assessment

A vulnerability assessment is a systematic process that involves defining, identifying, classifying, and prioritizing vulnerabilities in computer systems, applications, and network infrastructures. The objective of this assessment is to provide the organization with the necessary knowledge, awareness, and risk background to understand the threats to its environment and respond appropriately. This process includes an evaluation of potential vulnerabilities, an analysis of the risks associated with each vulnerability, and a prioritization of these vulnerabilities based on their potential impact on the organization. The result is a comprehensive report that can be used to guide the organization's efforts to improve its overall security posture.

4.2.2.5 Process 2.2 Activities

4.2.2.5.1 Vulnerability Assessment Analysis and The Initial Assessment in Any Devices

By conducting a vulnerability assessment and initial assessment of devices, it is possible to access a wide range of information about the risks that can threaten the systems. This information includes potential vulnerabilities in the systems, as well as any weaknesses that could be exploited by attackers. Additionally, the assessment can provide insights into the current security posture of the organization and identify areas where improvements are needed to enhance overall cyber-security. By leveraging this information, organizations can take proactive measures to mitigate potential risks and strengthen their security defences.

The activity uses the following inputs:

- Current knowledge of the Vulnerability assessment
- Current knowledge of different terms the Vulnerability assessment

The activity produces the following outputs:

- A vulnerability assessment categorizes information into various terms that provide significant information about vulnerabilities and risks. These terms may include identifying and classifying potential vulnerabilities in computer systems, applications, and network infrastructures, analysing the potential impact of these vulnerabilities on the organization, and prioritizing the vulnerabilities based on their level of risk. Additionally,

the assessment may provide recommendations for mitigation strategies that can help the organization reduce the potential impact of these vulnerabilities and strengthen its overall security posture. By leveraging this information, organizations can take proactive measures to improve their cyber-security defences and reduce the risk of a successful cyber-attack.

4.2.2.6 Process 2.3 System Baseline Definition

A baseline configuration refers to a predefined set of settings that are established on a system before it is approved for production use. Baselines are typically developed through the use of administrative checklists or other standardized processes to ensure that systems are configured correctly for their intended purpose. By establishing a baseline configuration, organizations can help ensure that their systems are secure, reliable, and perform as intended. This can include setting security policies, configuring system settings, installing required software and updates, and implementing other best practices to meet the needs of the organization. Once a baseline configuration has been established, it can be used as a reference point to monitor and manage the system, ensuring that any changes made are consistent with the organization's security policies and objectives.

4.2.2.7 Process 2.3 Activities

4.2.2.7.1 System Baseline Definition Analysis and Steps

A system baseline analysis can help identify which parts of a computer system have not been set up correctly for their intended purpose. By comparing the system's current settings to a baseline configuration, administrators can identify any deviations or inconsistencies and take corrective action to ensure that the system is configured properly. This can help improve the system's security, reliability, and performance, as well as ensure that it meets the organization's specific needs and objectives. Additionally, a baseline analysis can help identify potential vulnerabilities or weaknesses in the system, allowing administrators to take proactive measures to mitigate the risk of cyber-attacks or other security incidents.

The activity uses the following inputs:

- Current knowledge of the group setting of the system.
- Current knowledge of the system intended purpose.

The activity produces the following outputs:

The system baseline settings can be classified into different categories that provide significant information about the design of the system and its environment. These categories can include:

- Security settings - these are settings related to the security of the system, such as passwords, firewalls, access controls, and encryption. They help protect the system from unauthorised access, data breaches, and other security threats.
- Operating system settings - these are settings related to the operating system on the computer, such as user accounts, network settings, system services, and installed applications. They help ensure that the system runs smoothly and can be managed effectively.
- Hardware settings - these are settings related to the hardware components of the computer, such as the processor, memory, storage devices, and peripheral devices. They help ensure that the hardware is configured optimally and can perform as intended.
- Application settings - these are settings related to the specific applications or software installed on the computer, such as email clients, web browsers, and productivity software. They help ensure that the applications are configured correctly and can operate effectively.

By analysing these different baseline settings, administrators can gain a better understanding of the system's design, identify potential vulnerabilities or weaknesses, and take corrective action to improve the system's performance and security.

4.2.2.8 Process 2.4: System and Device Responses

When a system receives some information about vulnerabilities, its classifieds them in some categories, and they show some different reactions to them. This classification helps in prioritizing and addressing the identified vulnerabilities. While the system itself may not react in a traditional sense, the information is used to guide decision-making about how to mitigate or address the vulnerabilities. Some common vulnerability classifications include critical, high, medium, and low severity, as well as classifications based on the type of vulnerability. It's important to note that when a system receives information about vulnerabilities, it may not necessarily "react" in the traditional sense. Instead, the information is typically used by cybersecurity professionals to prioritize and address the identified vulnerabilities. The classification of vulnerabilities can help in determining the severity and potential impact of the vulnerabilities, which can guide decision-making about how to mitigate or address them. Some common vulnerability classifications include critical, high, medium, and low severity, as well as classifications based on the type of vulnerability, such as SQL injection or cross-site scripting.

4.2.2.9 Process 2.4 Activities

4.2.2.9.1 System Responses and Solutions

Based on the vulnerabilities and risks identified in a system, it is important to conduct a thorough analysis to determine the appropriate response. The nature and severity of the vulnerabilities may require different levels of action and response, ranging from minor patches and updates to more significant changes in system architecture and security measures. In some cases, the system may need to be taken offline or isolated from other networks to prevent further damage or exploitation. An in-depth analysis of the vulnerabilities and risks is crucial for determining the appropriate response and mitigating potential threats to the system and its users.

The activity uses the following inputs:

- Current knowledge of system vulnerabilities.
- Current knowledge of the system response and defence.

The activity produces the following outputs:

- When a system faces a risk, it first analyses it and then shows some related reactions..

4.2.3 Phase 3: Vulnerability Interrelationship in Cybersecurity

4.2.3.1 Phase Overview

Phase 3 of Cyber-Security, known as "Vulnerability Interrelationship", involves the examination of the interconnection between vulnerabilities in systems and devices. This phase aims to conduct an in-depth analysis of all aspects of vulnerabilities and their relationships. The information is identified, classified and put into a system that allows for the understanding of how vulnerabilities are interconnected. This understanding can help in finding ways to handle vulnerabilities in Cyber-Security effectively.

Process 1 in Phase 3 of Cybersecurity Vulnerability and Risk Analysis, focuses on examining the interrelationship between vulnerabilities in a system or device. This involves analysing how different vulnerabilities are interconnected and how one vulnerability can lead to the exploitation of another vulnerability.

Process 2 identifies the challenges that arise in analysing the interrelationship between vulnerabilities. This involves understanding the complexity of the system or device, the variety of vulnerabilities that may exist, and the difficulty in prioritizing which vulnerabilities to address first.

Process 3 illustrates how the system or device responds and reacts to the interrelationship between vulnerabilities. This involves understanding the system's or device's defence mechanisms and how they can be improved to prevent the exploitation of vulnerabilities. It also involves developing a plan for addressing vulnerabilities in a prioritized manner.

The processes in Phase 3 aim to examine and understand how vulnerabilities in a system or device are interrelated, identify challenges that may arise in the analysis, and illustrate how the system or device responds and reacts to vulnerabilities. By understanding the interrelationship of vulnerabilities, it is possible to better identify and address potential security threats in the system or device.

4.2.3.2 Process 3.1 Vulnerability Interrelationship Analysis

Yes, that is correct. Understanding the interrelationship of vulnerabilities can provide a better understanding of the system or asset, and how different vulnerabilities are connected to each other. This can help in prioritizing which vulnerabilities need to be addressed first and developing a more comprehensive plan to improve the overall security of the system or asset.

4.2.3.3 Process 3.1 Activities

4.2.3.3.1 Vulnerability Interrelationship Analysis and objectives

Identifying the different aspects of system vulnerabilities is an essential step in examining the vulnerability interrelationship. This involves analysing and understanding the various vulnerabilities that exist in the system and how they relate to each other. It is necessary to determine the impact that each vulnerability could have on the system as a whole and how they can interact with each other to create a more significant security risk.

This process can involve conducting a thorough vulnerability assessment, reviewing system logs, and examining network traffic to identify potential vulnerabilities. Once identified, the vulnerabilities are categorized and prioritized based on their severity, impact on the system, and the likelihood of exploitation.

By understanding the interrelationship between vulnerabilities, organizations can develop a more comprehensive and effective security strategy to mitigate the risks associated with cyber threats. This can involve implementing additional security controls, updating software and hardware, and educating employees on best practices to reduce the risk of human error.

The activity uses the following inputs:

- Current knowledge of the vulnerabilities in Cyber-Security

- Current knowledge of the interrelationship of the vulnerabilities

The activity produces the following outputs:

- The Vulnerability interrelationship can be classified into various terms that provide significant information about the interdependence of vulnerabilities. These terms include:
  - Vulnerability chain: A series of vulnerabilities that can be exploited in a particular order to gain unauthorised access to a system or network.
  - Attack vector: A path or means by which a hacker or attacker can gain unauthorised access to a system or network.
  - Attack surface: The set of all possible entry points that an attacker can use to gain access to a system or network.
  - Risk propagation: The spread of risk from one system or component to another due to interdependencies or common vulnerabilities.
  - Threat correlation: The analysis of multiple threats and their interrelationships to identify common patterns or causes.
  - Impact analysis: An assessment of the potential consequences of a successful attack or exploitation of a vulnerability.
  - Understanding these terms and their interrelationship can help in better assessing and mitigating vulnerabilities in a system or network.

4.2.3.4 Process 3.2: Vulnerability Interrelationship Challenges

Identifying vulnerabilities interrelationship can pose some challenges. One of the major challenges is that vulnerabilities often have complex interdependencies and may not be easy to isolate. Additionally, vulnerabilities may have different levels of criticality and impact on the system, making it difficult to prioritize which ones to address first. Another challenge is that the relationships between vulnerabilities may change over time as new vulnerabilities are discovered or existing vulnerabilities are mitigated. It is also possible that fixing one vulnerability may inadvertently create a new vulnerability or affect the overall security posture of the system in unexpected ways. Therefore, it is important to carefully analyse the vulnerabilities' interrelationship to ensure that all risks are properly addressed.

4.2.3.5 Process 3.2 Activities

4.2.3.5.1 Vulnerability Interrelationship Analysis and Objective

Initially, we analysed the challenges associated with the interrelationship of vulnerabilities by

identifying various aspects of system vulnerabilities.

By analysing these challenges, we can gather various information regarding the difficulties and obstacles that we may encounter when trying to understand the vulnerability interrelationship of a system.

The activity uses the following inputs:

- Current knowledge of the vulnerability interrelationship challenges.
- Current knowledge of how to handle vulnerabilities interrelationship.

The activity produces the following outputs:

- The vulnerability interrelationship classified information into some different terms to address possible challenges.

4.2.3.6 Process 3.3 Device Responses

When a system is provided with information about the interrelationship between vulnerabilities, it becomes easier to identify and recognize those vulnerabilities.

4.2.3.7 Process 3.3 Activities

4.2.3.7.1 The system and devices responses and reaction

Based on the Vulnerabilities interrelationship, systems should make a comparison with its information and then shows reaction, which necessitates to an in-depth analysis.

The activity uses the following inputs:

- Current knowledge of the system vulnerabilities interrelationship.
- Current knowledge of the system response and defence.

The activity produces the following outputs:

- When a system encounters a potential risk or threat, it typically performs an analysis of the risk and then proceeds to implement relevant responses or reaction.

4.2.4 Phase 4: Machine Learning and Natural Language Processing in Cybersecurity

4.2.4.1 Phase Overview

Phase 4, which is called "Machine Learning and Natural Language Processing in Cyber-Security", consists of three processes. The aim of this phase is to explore the potential applications of ML and NLP algorithms in Cyber-Security. By conducting an in-depth analysis of these algorithms,

we can gain a better understanding of their capabilities and limitations. The ultimate goal is to develop and implement ML and NLP algorithms in Cyber-Security to enhance the effectiveness of security measures.

Process 1 involves examining the concepts of Machine learning and natural language processing in the context of Cyber-Security. This includes understanding how these techniques can be used to improve the security of computer systems and networks.

Process 2 focuses on the implementation of Machine learning and natural language processing in Cyber-Security. This involves exploring the different tools and technologies available for implementing these techniques, as well as understanding the challenges and limitations of using them in a Cyber-Security context.

Process 3 involves collecting and analysing datasets for Machine learning and natural language processing in Cyber-Security. This includes identifying relevant data sources, cleaning and pre-processing the data, and selecting appropriate algorithms for analysis. The goal of this process is to build a comprehensive dataset that can be used for training and testing Machine learning and natural language processing models in the context of Cyber-Security.

The primary objective of these processes is to implement ML and NLP algorithms in Cyber-Security to tackle the vulnerabilities.

4.2.4.2 Process 4.1: ML and NLP in Cybersecurity Analysis

Machine learning has been used for various applications in cybersecurity, including intrusion detection, malware classification, android malware detection, spam and phishing detection, and binary analysis. Due to the unstructured nature of the information corpus, Machine learning algorithms are crucial in increasing the use of NLP in cybersecurity. Machine learning has demonstrated significant advancements overrule-based and classic learning-based solutions in detecting malware and network intrusions [165].

4.2.4.3 Process 4.1 Activities

4.2.4.3.1 The process of ML and NLP in Cyber-Security Analysis and Objectives

To examine Machin Learning and NLP in Cyber-Security, the algorithms of ML and NLP need to be identified first. This analysis can help in understanding how the ML and NLP algorithms can be used to analyse information about system vulnerabilities and effectively handles them in Cyber-Security [165].

The activity uses the following inputs:

- Current knowledge of Machin Learning in Cybersecurity.
- Current knowledge of the algorithms of Machin learning and Natural Language Processing.

The activity produces the following outputs:

- Machine learning and natural language processing can categorize the information about vulnerabilities in cybersecurity into various terms.

4.2.5 Process 4.2: ML and NLP in Cybersecurity Implementation

Machin learning is a subset of Artificial Intelligence (AI) and Cognitive Computing. It emerged at the same time as cognitive science and linguistics, all influenced by the computer's invention. Cognitive computing has the ability to learn and adapt, and it understands context, such as text meaning or domain knowledge. This makes cognitive computing suitable for cyber immunity solutions. However, Machine learning algorithms may not perform well with small data volumes because they need a significant amount of data to understand it accurately [165].

4.2.5.1 Process 4.2 Activities

4.2.5.1.1 The Implementation of ML and NLP on Any System and Devices

ML and NLP algorithms can be applied in the field of cybersecurity to perform various tasks such as identifying and categorizing malware, detecting network intrusions, detecting phishing/spam attacks, and inspecting website defacements.

The activity uses the following inputs:

- Current knowledge of the implementation of Machin learning and Natural Language Processing in Cyber-Security

The activity produces the following outputs:

- Machine learning and natural language processing techniques can be used to analyse information about vulnerabilities in cyber security and represent them in different terms. This can involve extracting relevant features and patterns from raw input data, classifying them based on their characteristics, and identifying potential threats or vulnerabilities. By applying these techniques, complex and unstructured data can be transformed into structured and meaningful representations, allowing for more effective analysis and decision making in the context of cyber security.

4.2.5.2 Process 4.3: ML and NLP Data Collection

Traditionally, NLP for machine learning involved both pre-processing of textual data and feature

engineering. However, a new Machine learning model called Word2Vec-Keras Text Classifier has been developed for text classification without the need for feature engineering. This model utilizes the Word2Vec model from Genism, which is a Python library for tasks such as topic modelling, document indexing, and similarity retrieval for large text corpora, and combines it with KNN through an embedding layer used as input.

4.2.5.3 Process 4.3 Activities

4.2.5.3.1 ML and NLP Dataset Collection Operation

We first examine ML and NLP dataset collection, by Using them can be led to:

- Keep a single copy of data in our storage, referenced by datasets.
- Seamlessly access data during model training without worrying about connection strings or data paths.
- Share data and collaborate with other users.

The activity uses the following inputs:

- Current knowledge of the ML and NLP dataset collection
- Current knowledge of the Operation of ML and NLP dataset collection

The activity produces the following outputs:

- The Machine Learning and Natural Language Processing dataset collects the data and keeps a copy of our storage.

4.2.5.3.2 Phase 4 Diagram

Figure 4-2 illustrates the process of the Machin Learning and Natural Language Processing in Cyber-Security. The diagram highlights the inputs and outputs of the process.

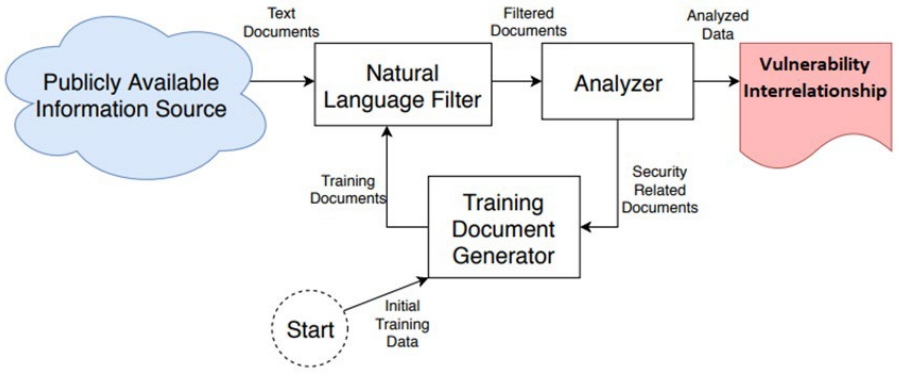


Figure 4-2. Process of the Machin Learning and Natural Language Processing in Cyber-Security

4.3 Summary

Linking biology and cybersecurity is an emerging field that has gained significant attention in recent years. With the increasing use of technology and data in biological research and healthcare, there is a growing need to ensure the security and integrity of biological data and systems. However, this is a complex task that requires interdisciplinary expertise and knowledge, and there is a lack of well-defined requirements and specifications in this field.

In the next chapter, we aim to provide more specific requirements and specifications that can help to provide better insight into the link between biology and cybersecurity. These requirements are based on a comprehensive analysis of the current state of the field and the challenges and opportunities that it presents. By providing these specific requirements and specifications, we aim to contribute to the development of a more comprehensive and effective approach to linking biology and cybersecurity. It is worth noting that this analysis with these details is the first of its kind, and we hope it will serve as a useful resource for researchers, practitioners, and policymakers in this emerging field.

Chapter 5

Bio Inspired Cybersecurity Framework

5.1 Overview

In this chapter, we will take a closer look at the framework used in this study. As previously stated, we have divided bio Inspired Cybersecurity into different phases and processes. Here, we provide a more detailed explanation of our methodology and the technical work involved. The primary objective of this section is to gain a deeper understanding of the interrelationships between vulnerabilities and create a vulnerability tree. Figure 5-1 displays the data-flow diagram of the various phases of our methodology.

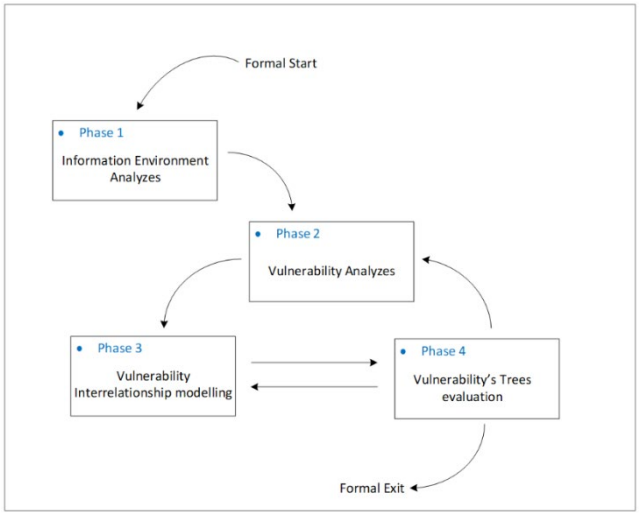


Figure 5-1. Diagram of Phases Overview.

Each phase of the methodology is presented in Table 5.1.

Table 5.1 Phase of the methodology

Phase/Process/Activity	Literature Sources	Original Contributions
Phase Overview	Cybersecurity frameworks and models, biology textbooks	Identification and definition of phase processes
Process Overview	Peer-reviewed journals, textbooks on relevant disciplines	Analysis and definition of the process within the phase
Process Activities	Research articles, case studies, educational materials	Definition of activities, inputs, and outputs within the process
Process Diagram	Academic literature on data-flow diagrams, process modeling techniques	Development of data-flow diagram illustrating the process

5.2 Framework Evaluation

Framework Evaluation introduces the importance of evaluating the proposed framework to ensure its effectiveness, efficiency, and applicability in real-world cybersecurity scenarios.

- **Findings**: The framework was found to comprehensively address key aspects of bio-inspired cybersecurity, including threat identification, analysis, and mitigation strategies. It incorporates insights from the human immune system to predict and respond to cyber threats effectively.
- **Gaps and Enhancements**: While the framework covers a broad spectrum of cybersecurity concerns, areas such as real-time threat response and automatic adaptation to new threats were identified as needing further development. Future work could focus on enhancing these aspects to improve the framework's completeness.

Highlight the evaluation criteria: completeness, explainability, and potential for exploitation.

5.2.1 Explainability Evaluation

The framework was designed with a focus on user-friendliness and explainability, employing clear, non-technical language and visual aids where possible. Researcher findings indicated that the framework's concepts and processes are accessible to individuals with varying levels of cybersecurity expertise [184].

5.2.2 Potential for Exploitation

- Practical Applications: The framework's bio-inspired approach offers new avenues for developing cybersecurity solutions that are adaptive, resilient, and capable of identifying and mitigating previously unknown threats. It is particularly applicable in sectors where cybersecurity threats can have critical consequences, such as finance, healthcare, and critical infrastructure.
- Areas for Exploitation: Beyond direct cybersecurity improvements, the framework can be exploited for educational purposes, helping to train the next generation of cybersecurity professionals in innovative defence mechanisms. It also holds potential for influencing policymaking by providing a robust model for understanding and combating cyber threats [185].

5.3 Phase 1 Information Environment Analysis

5.3.1 Phase 1 Overview

Phase 1 of the methodology is called "Information Environment Analysis," which comprises four processes. The purpose of this phase is to analyse the information environment in which the enterprise operates. Figure 5-2 shows the process of information environment analysis. The first process, Process 1, examines the business process by analysing the goals, processes, and environment of the enterprise. The second process, Process 2, collects and analyses data related to the enterprise's assets. Process 3 identifies the boundaries of the system and the internal and external stakeholders of the enterprise. Throughout this phase, the identified stakeholders are included to gain their perceptions of the business, threat agents, and the system itself. The goal of this phase is to define a preliminary set of security requirements, which may be required in subsequent phases.

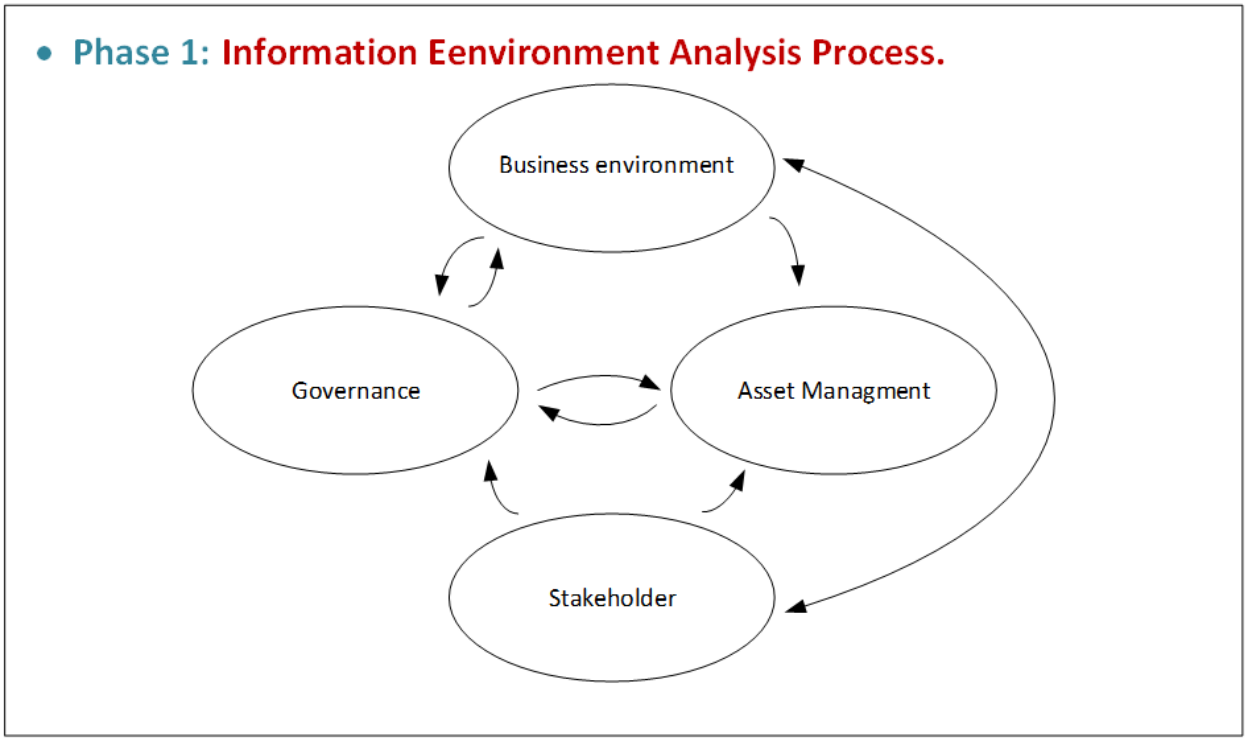


Figure 5-2. Process of Information Environment Analysis.

5.3.2 Process 1: Business Environment

The Function of understanding the business context, identifying critical resources and associated cybersecurity risks allows an organization to effectively prioritize and focus its efforts, based on its risk management strategy and business needs. The outcomes of this Function can be categorized into areas such as Asset Management, Business Environment, Governance, Risk Assessment, and Risk Management Strategy. These outcomes help organizations define and implement effective cybersecurity measures to protect their critical assets and meet their business goals [1][7].

In the context of business strategy, it is common to conduct a SWOT analysis (Strengths, Weaknesses, Opportunities, and Threats) to understand the current situation and develop a plan. However, cyber security has some unique characteristics that require a different approach than a SWOT analysis [8]. Firstly, cyber security is not a standalone function and is always connected to the whole organization. Secondly, it is reactive in nature and responds to incidents rather than proactively preventing them. Finally, cyber security is asymmetrical, meaning that it is easier for an attacker to exploit vulnerabilities than it is for the defender to prevent them [2].

Cybersecurity is a set of practices and technologies that are designed to protect computers, networks, and other digital systems from unauthorised access, use, or destruction. The ultimate

goal of cybersecurity is to protect the information assets of an organization, which can include sensitive data, intellectual property, financial information, and other valuable resources. By safeguarding these assets, an organization can maintain its competitive advantage and ensure the continued success of its business operations [9].

In cybersecurity strategy, business analysis includes areas that should all be considered:

- Business
- Strategy
- External Factors
- IT Technology
- Roadmap
- IT Capabilities the Organization Must Acquire to Accomplish the Roadmap
- Current Security State
- Need State Based on the Technology Roadmap
- Benefits
- Resources
- Constraints
- Assumptions
- Risk to the Plan (Not the Organization)
- Security Governance [10].

5.3.3 Business Processes Analysis

The identification of critical business processes can help in identifying the assets that support these processes, which in turn can reveal vulnerabilities that may exist in these assets. Therefore, it is essential to identify and prioritize critical business processes to ensure that they are adequately protected from potential cybersecurity threats. Figure 5-3 shows the diagram of business analysis. For analysing business process, we consider following input:

- Current Knowledge of Stakeholder
- Information Security Policy
- Organization data
- Information Gathering

And the process produces the following outputs:

- Technical Environment Report
- Business Environment Report
- Physical Environment Report
- Business Process list

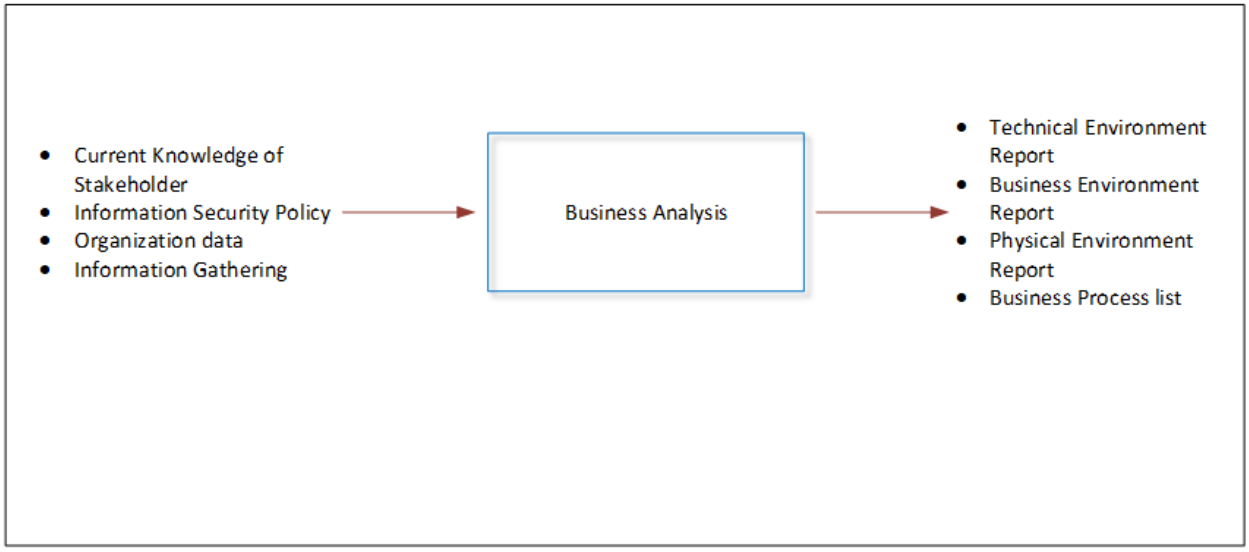


Figure 5-3. Diagram of the Analysing Business

5.3.4 Environmental Analysis

Environmental analysis is a strategic tool. PESTLE analysis is a widely used tool for environmental analysis in strategic planning. It stands for Political, Economic, Social, Technological, Legal, and Environmental factors [3]. By analysing these factors, organizations can better understand the external environment and its potential impact on their operations and strategic goals. PESTLE analysis helps in identifying potential threats and opportunities, and in formulating strategies to respond to them [6].

It is worth noting that while Porter's 5 Forces model (Figure 5-4) is primarily used in marketing and business strategy, it can also be adapted and applied in the context of cybersecurity [8]. By examining the five forces (Threat of New Entrants, Bargaining Power of Suppliers, Bargaining Power of Buyers, Threat of Substitutes, and Competitive Rivalry) within the cybersecurity industry, organizations can gain a better understanding of their competitive environment and make informed decisions about their cybersecurity strategy. For example, understanding the threat of new entrants can help organizations anticipate new cybersecurity threats and adjust their defences accordingly [6].

The 5 forces that make up Porter's model are:

- Supplier power
- Buyer power
- Threat of substitution
- Threat of new entry
- Competitive rivalry

Marketers who can understand and formulate strategy around these five forces should be able, in theory, to improve the profitability of their brands and/or products [5].

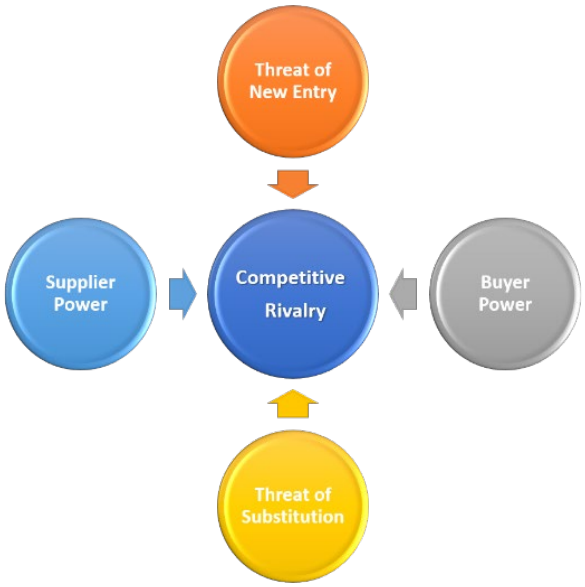


Figure 5-4. Michael Porter's 5 Forces model

5.3.5 Governance

Enterprise governance is crucial for effective risk management and ensuring that an organization achieves its objectives while using resources responsibly. In general, governance involves a set of responsibilities and practices exercised by those responsible for the enterprise. This can include the board and executive management in a corporation, or the agency head for a federal agency [11].

Governance is aimed at providing strategic direction, ensuring that objectives are achieved, managing risks appropriately, and verifying that the enterprise's resources are used responsibly. Different domains within an organization, such as IT, finance, legal and regulatory compliance,

and information security, may require specialized expertise in order to manage risks effectively. Therefore, enterprise governance is often organized by domain [4].

Cyber security governance is a subset of enterprise governance that focuses specifically on managing risks and protecting assets in cyberspace, given the presence of adversaries who may seek to exploit vulnerabilities in information systems [5]. Information systems security governance, which is concerned with protecting the confidentiality, integrity, and availability of information in general, can be seen as a component of cyber security governance. However, because so much of today's information is stored, transmitted, and processed electronically, it is often more practical for cyber security governance to encompass information security governance more broadly [6][12].

5.3.6 Stakeholder Identification

Identifying stakeholders and their responsibilities is a crucial step in the cyber security governance process. It helps to clarify who is responsible for what within the organization and ensures that everyone is aware of their roles and responsibilities in ensuring the security of the organization's information assets.

The first step in this process is to identify the different stakeholders within the organization. These may include executives, managers, employees, customers, suppliers, regulators, and other stakeholders who have a vested interest in the organization's information security. Once the stakeholders have been identified, their roles and responsibilities can be defined.

For example, executives and managers may be responsible for setting the overall cyber security strategy for the organization and ensuring that it is implemented effectively. Employees may be responsible for following security policies and procedures and reporting any security incidents or vulnerabilities they encounter. Customers may expect the organization to safeguard their personal and financial information, while suppliers may need to be aware of and adhere to the organization's security policies and procedures.

Once the stakeholders and their responsibilities have been identified, the organization can develop policies and procedures that ensure that everyone is aware of their roles and responsibilities and that the organization's information assets are protected. These policies and procedures should be communicated clearly to all stakeholders and should be reviewed and updated regularly to ensure that they remain effective in the face of evolving cyber security threats [20]. This section aims to identify the individuals or groups with an interest in a company, known as stakeholders, and their

corresponding responsibilities within the organization. Stakeholders can be internal or external to the organization, depending on their direct or indirect relationship to the company. The three main types of stakeholders for a computer system are the management, users, and developers. These stakeholders' roles and responsibilities may vary depending on the type of business the organization conducts.

The process consists of the following activities:

- Identify stakeholders,
- Identify stakeholder responsibilities.

5.3.6.1 What Is a Stakeholder?

A stakeholder refers to any party that holds an interest in a company and can either have an impact on or be impacted by the business. The primary stakeholders in a typical corporation include investors, employees, customers, and suppliers [7]. Stakeholders can be internal or external to an organization. Internal stakeholders are those with a direct relationship to the company, such as employees, owners, or investors. External stakeholders are individuals or groups that do not directly work with the company but are still affected by its actions and outcomes. Examples of external stakeholders include suppliers, creditors, and public organizations [8].

In the case of computer systems, there is a specific set of stakeholders that can be identified to define its function and structure. These stakeholders are classified into different categories depending on the type of business the enterprise is conducting. However, there are generally three types of stakeholders in a computer system: management, users, and developers [13].

The Identify Stakeholders process has the following Inputs:

- **Project Charter:** A project charter is a top-level document that grants permission for a project to proceed and designates the project manager with the authority to carry it out [21].
- **Procurement Documents:** Procurement documents are a set of documents that are used to identify the stakeholders involved in the procurement process of goods or services. These documents include a request for proposal (RFP), request for quotation (RFQ), invitation to bid (ITB), and other legal and financial documents related to the procurement process. The procurement documents help to identify and evaluate potential suppliers, negotiate contracts, and ensure that the procurement process is transparent and fair to all stakeholders involved [23].
- **Enterprise Environmental Factors:** Enterprise Environmental Factors refer to all

external and internal environmental factors and conditions that can influence the project's success or failure. These factors include organizational culture, governance structure, political and social climate, economic conditions, technology, infrastructure, industry standards, and regulations, among others. Understanding these factors can help project managers plan and execute their projects in a way that maximizes the chances of success [25].

- **Organizational Process Assets:** Organizational Process Assets refer to the company's internal resources that are used to execute projects. They include standard policies and procedures, templates, lessons learned, historical data, and knowledge bases. These assets help to ensure consistency and efficiency across all projects, as well as provide a basis for continuous improvement. They are an important input for project planning and execution and help to guide decision-making throughout the project lifecycle. Examples of organizational process assets include project management methodologies, guidelines for project management, communication plans, and risk management strategies [13].

The Identify Stakeholders process uses the following Tools & Techniques:

- **Stakeholder Analysis:** Stakeholder analysis is the process of collecting and evaluating information to identify the parties whose interests should be taken into account for a project.
- **Expert Judgment:** Expert judgment refers to the input and insights provided by individuals who have a high level of expertise and experience in a particular field. This input can be technical, managerial, or both, and is used to inform decision-making and problem-solving processes within a project. Expert judgment can come from a variety of sources, including internal or external stakeholders, consultants, subject matter experts, or industry experts. It is an important tool for project managers to use in order to ensure that decisions and actions are based on informed and knowledgeable input [19].

The Identify Stakeholders process has the following Outputs:

- **Stakeholder Register:** The Stakeholder Register is a paper that lists all details related to project stakeholders, including their requirements and categorization [14].
- **Stakeholder Management Strategy:** The Stakeholder Management Strategy outlines a plan to enhance stakeholder support and minimize any negative effects, as identified in a stakeholder analysis matrix [15].

5.3.6.2 Identify Stakeholders Diagram

This is a visual representation of the stakeholder identification process, which illustrates the inputs and outputs of the process through a diagram (Figure 5-5).

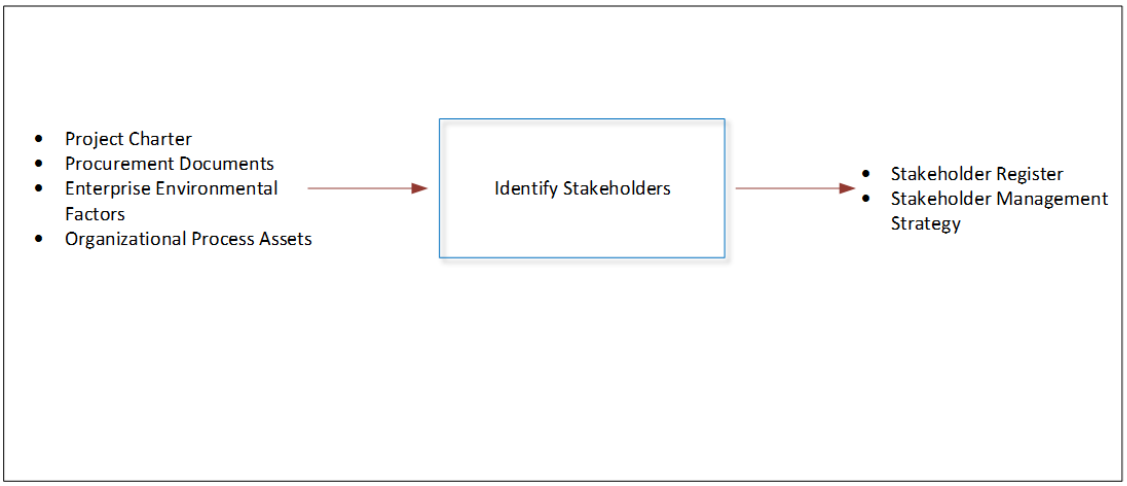


Figure 5-5. Diagram of the Stakeholder Identification

5.4 Phase 2 Vulnerability Analysis

5.4.1 Phase 2 Overview

Phase 2 is referred to as "Vulnerability Analysis" and comprises three distinct processes. The purpose of this phase is to recognize and assess vulnerabilities and risks associated with various systems and devices. The approach involves gathering information on weaknesses in systems, categorizing and analysing the collected data, and incorporating it into the assessment.

The primary objective of this process is to classify vulnerabilities in a manner that facilitates the identification of any interrelationships between them. Process 2 involves the identification of vulnerabilities through a vulnerability assessment, while Process 3 entails the analysis of these vulnerabilities. The key goal of this process is to classify and identify the various types of vulnerabilities. It is worth noting that, given the nature of this methodology, the assessor can exercise flexibility in determining the sequence of process execution.

5.4.1.1 Process 2.1: Vulnerability Identification

The identification of vulnerabilities and the assurance of secure functionality through security testing are commonly used methods for evaluating and enhancing the security of software. Given the openness of contemporary software-based systems, the use of appropriate security testing techniques is becoming

increasingly crucial and necessary to conduct effective and efficient security testing [16]. Thus, it is essential to provide an overview of current security testing techniques that can be of high value to researchers seeking to evaluate and improve the techniques, as well as to practitioners looking to apply and disseminate them [17]. Figure 5-6 illustrates the data-flow diagram of Vulnerability Identification and Analysis.

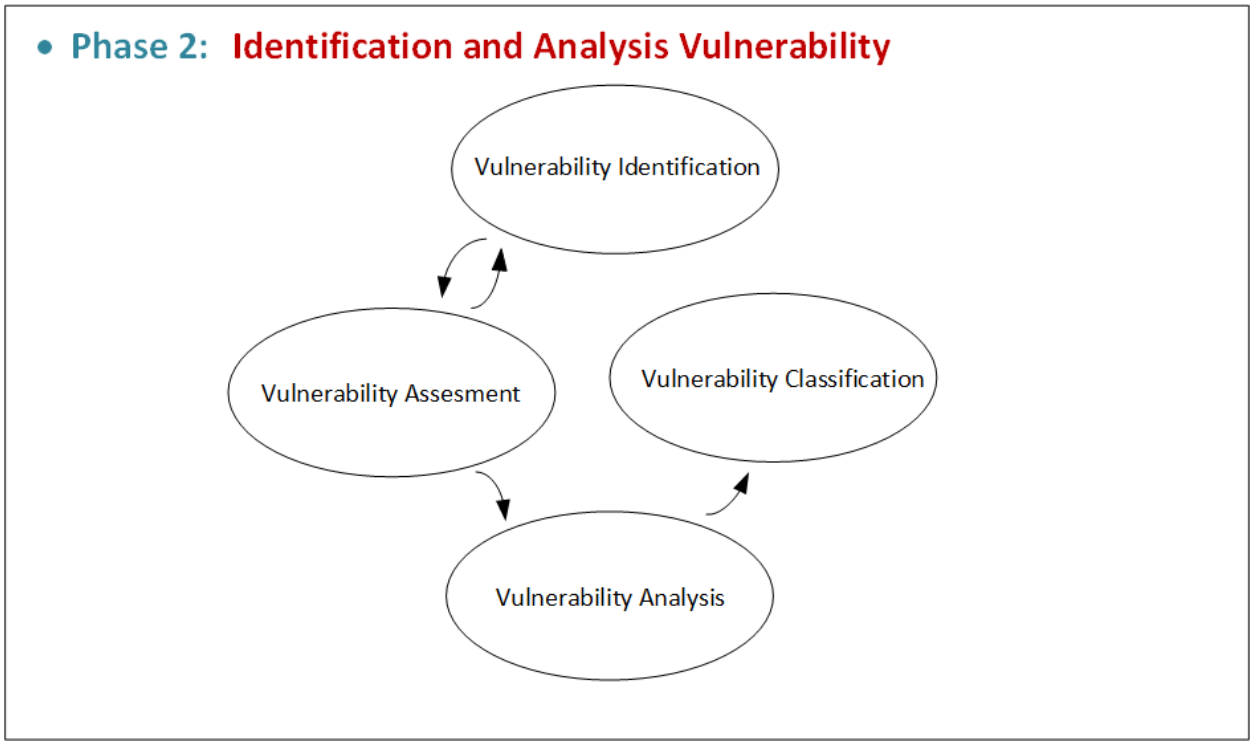


Figure 5-6. Process of Vulnerability Identification

5.4.1.2 Activity 2.1.1: Vulnerability Identification

The Oxford Dictionary provides a brief definition of the term "vulnerability" as being susceptible to damage. However, there have been various definitions of vulnerability offered by different sources:

- A vulnerability can be defined as a point or weakness in a system that makes it susceptible to an attack.
- A vulnerability is a flaw or weakness in a security system that could be exploited to cause harm or result in loss.
- A vulnerability is a weakness within a system that could potentially allow security to be violated [17].

The activity uses the following inputs:

- Current knowledge of Vulnerability: The current understanding is that weaknesses within a system can be planned and exploited in order to cause harm or result in loss. Current knowledge of System weakness: The current understanding is that weaknesses within a system can be planned and exploited in order to cause harm or result in loss.

The activity produces the following outputs:

- Vulnerability List: The Vulnerability List will be arranged in order of priority, with the most significant vulnerabilities listed first, in order to facilitate the identification of any interrelationships between them.

5.4.2 Process 2.2: Vulnerability Assessment

5.4.2.1 Activity 2.2.1: Vulnerability Assessment Identification

Vulnerability assessment is considered a supplementary methodology [18]. Essentially, vulnerability assessments involve identifying and reporting vulnerabilities, but it is important to note that the selection of vulnerability assessment methods will vary depending on the purpose and scope of the assessment, which will also impact the degree of analysis required [19].

5.4.2.2 Activity 2.2.2: Vulnerability Assessment Type Selection

Vulnerability Assessment can classify in two categories:

- Network-based assessment.
- Host-based assessment.

Network-based vulnerability assessment tools enable network administrators to detect and eliminate security vulnerabilities present in their organization's network infrastructure. In contrast, host-based scanning tools assist network administrators in securing the internal systems of their organization by adding an extra layer of protection. These tools limit access to the hosts, thus preventing unauthorised access to sensitive data within the organization. To summarize, network-based analysis is used to test and monitor the overall network, whereas host-based analysis focuses on specific host [5].

The activity uses the following inputs:

- Current knowledge of Vulnerability assessment

The activity produces the following outputs:

- Qualified Vulnerability: The Qualified Vulnerability Analysis will assess a broad spectrum of network-related concerns and subsequently identify the specific weaknesses in the

network that require remediation [5].

5.4.3 Process 2.3: Vulnerability Analysis

5.4.3.1 Activity 2.3.1: Vulnerability Type Identification

In any system there are six types of vulnerabilities can exist:

- Physical,
- Natural,
- Hardware/Software,
- Media,
- Communication,
- Human [19].

5.4.3.2 Activity 2.3.2: Vulnerability Type Selection

In order to conduct a comprehensive vulnerability assessment, all types of vulnerabilities should be taken into consideration and thoroughly analysed [18].

The activity uses the following inputs:

- Vulnerability Type List: listing the different types of vulnerabilities that were identified in the system,
- Technical Environment Report: definition of the technical environment, which is the system operating,
- Physical Environment Report: definition of the physical environment, which is the system operating.

The activity produces the following output:

- Vulnerability Classification: listing the types of vulnerabilities that were selected for further examination.

5.4.4 Phase 2 Diagram

The process diagram for Phase 2 is provided in Figure 5-7, which outlines the various steps involved in identifying and analysing vulnerabilities, as well as the inputs and outputs associated with this phase.

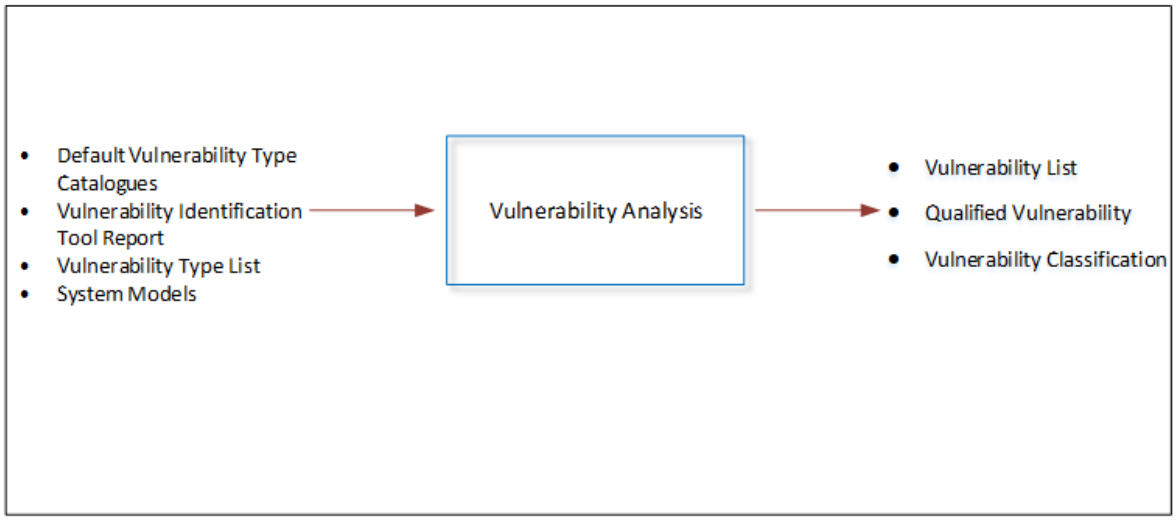


Figure 5-7. Diagram of Analysing Vulnerability

5.5 Phase 3 Vulnerability Interrelationship Modelling

5.5.1 Phase 3 overview:

Phase 3, titled "Vulnerability Interrelationship Modelling," consists of six processes that aim to identify, analyse, and model the interrelationship of vulnerabilities in systems and devices. The process begins with a pre-analysis of vulnerability interrelationship in process 1, where vulnerabilities are classified to find their interrelationship. Process 2 identifies the vulnerability structure, while process 3 involves node analysis. Process 4 examines value analysis, process 5 presents vulnerability optimization, and process 6 involves modelling the vulnerability interrelationship. Figure 5-8 indicates this process.

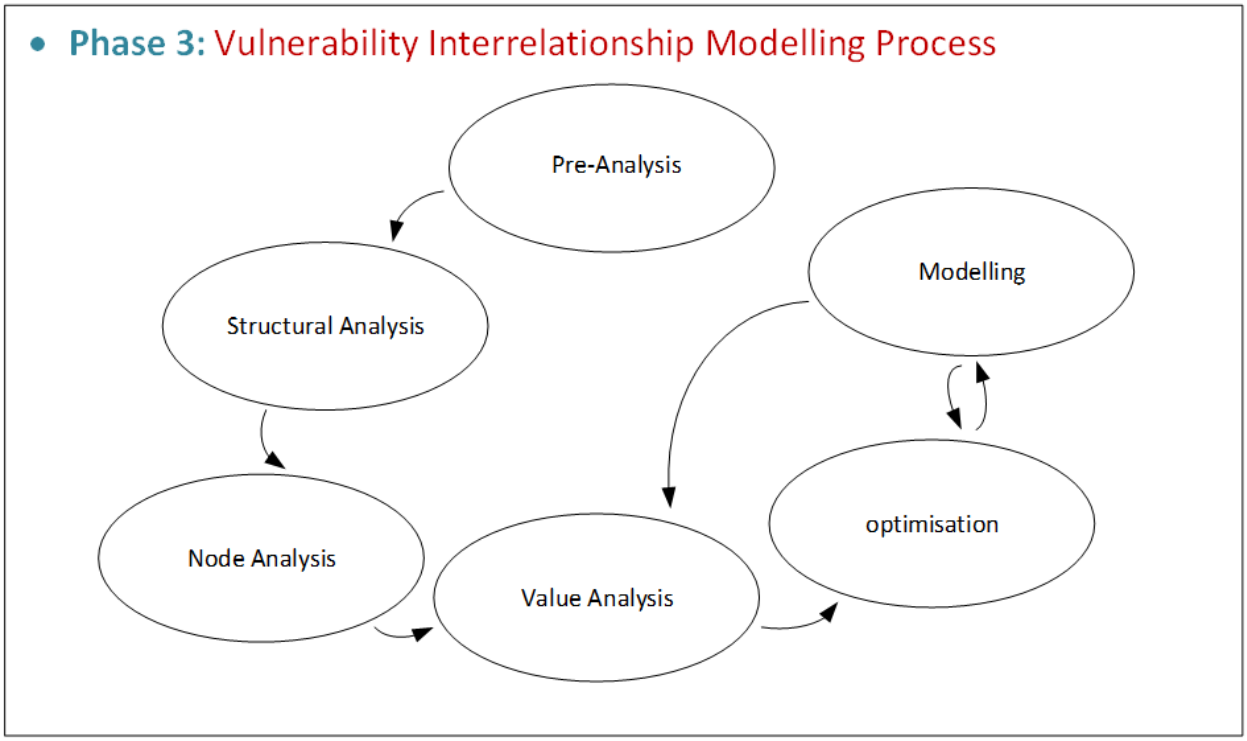


Figure 5-8. Process of Vulnerabilities Modelling and Their Interrelationship

5.5.2 Process 3.1: Pre-Analysis

5.5.2.1 Activity 3.1.1: Pre-Analysis Identification

In this activity, the information collected during the vulnerability identification process is utilized to identify a specific vulnerability, which is referred to as the top or parent vulnerability and represented by a capital 'V'. Assets related to this vulnerability are identified and selected from the asset table created during the asset identification step. These assets are associated with multiple vulnerabilities, which are selected from the vulnerability table created during the vulnerability identification step.

The activity uses the following inputs:

- Vulnerability Preference List: listing the vulnerabilities that were filtered for further analysis.
- Asset List: listing all identified assets of the enterprise.
- Vulnerability List: listing all identified vulnerabilities of the enterprise.

The activity produces the following outputs:

- Pre-analysis Asset List: listing all assets associated with the top vulnerability, selected from

the Asset List.

- Pre-analysis Vulnerability List: listing all vulnerabilities associated with the assets contained in.

5.5.3 Process 3.2: Structural Analysis

5.5.3.1 Activity 3.2.1: Structural Analysis Identification

During this step, trees are created by determining the branches and nodes. The process starts from the top vulnerability and moves downwards until it reaches the level where no more steps are needed. Usually, it is rare to have only a level 0 vulnerability tree. The level 0 tree does not have any nodes as steps. The number of levels in the tree depends on how detailed the analysis needs to be. The assessment's scope may limit the types of vulnerabilities that need to be considered, so the analyst may choose not to analyse a particular child vulnerability.

The activity uses the following inputs:

- Top Vulnerability: the top vulnerability of the tree,
- Control List: describing which stakeholder has control over which asset, and to the extent over which vulnerability.

The activity produces the following output:

- Vulnerability Tree: the constructed tree

5.5.4 Process 3.3: Node Analysis

5.5.4.1 Activity 3.3.1: Node Analysis Identification

In this activity, the attributes of the vulnerabilities contained in the Node List are being filled out. These attributes include ID, name, type, category, complexity value, educational complexity, time to exploit, family position, head, asset ID, tree ID, and description. The time to exploit (T.T.E.), vulnerability complexity (V.C.), and educational complexity (E.C.) values are being evaluated and assigned based on the type of threat agent that is selected. Additional information about the threat agents can be found under the threat agent identification activity [85].

The activity uses the following inputs:

- Threat Agent Preference List: listing the threat agents that are considered to have an interest in the enterprise.

The activity produces the following output:

- Listing the Nodes of the Tree: listing the nodes of the vulnerability tree.

5.5.5 Process 3.4: Value Analysis

5.5.5.1 Activity 3.4.1: Value Analysis Identification

In this activity, the main focus is on identifying the critical paths that lead to the top vulnerability, which are determined based on the selected type of threat agent. By identifying the critical paths, it becomes possible to determine the complexity value of the top vulnerability. Once the complexity value is established, it is then easier to select the threat agents that have the capability to exploit that vulnerability. Additionally, the top vulnerability will also have an associated educational value [79].

The activity uses the following inputs:

- Vulnerability Tree: the constructed tree.
- Node List: listing the nodes of the tree.

The activity produces the following output:

- Constructed Tree and Listing the Nodes of the Tree.

5.5.6 Process 3.5: Optimization

5.5.6.1 Activity 3.5.1: Vulnerability Optimization

In this activity, the focus is on identifying and minimizing the costs of countermeasures for vulnerabilities that occur more than once in a vulnerability tree or across different trees. This can be achieved either through manual observation or with the help of an automated tool that examines all the trees to identify duplicate vulnerabilities. The more instances a vulnerability has, the more cost-effective the countermeasures will be. Additionally, vulnerabilities that are lower in the family hierarchy will be easier to address with more efficient countermeasures. While this activity is not currently part of the vulnerability assessment process, it is an idea that could be developed in the future.

5.5.7 Phase 3 Diagram

Figure 5-9. shows the process diagram of Phase 3. The diagram highlights the inputs and outputs of the Phase.

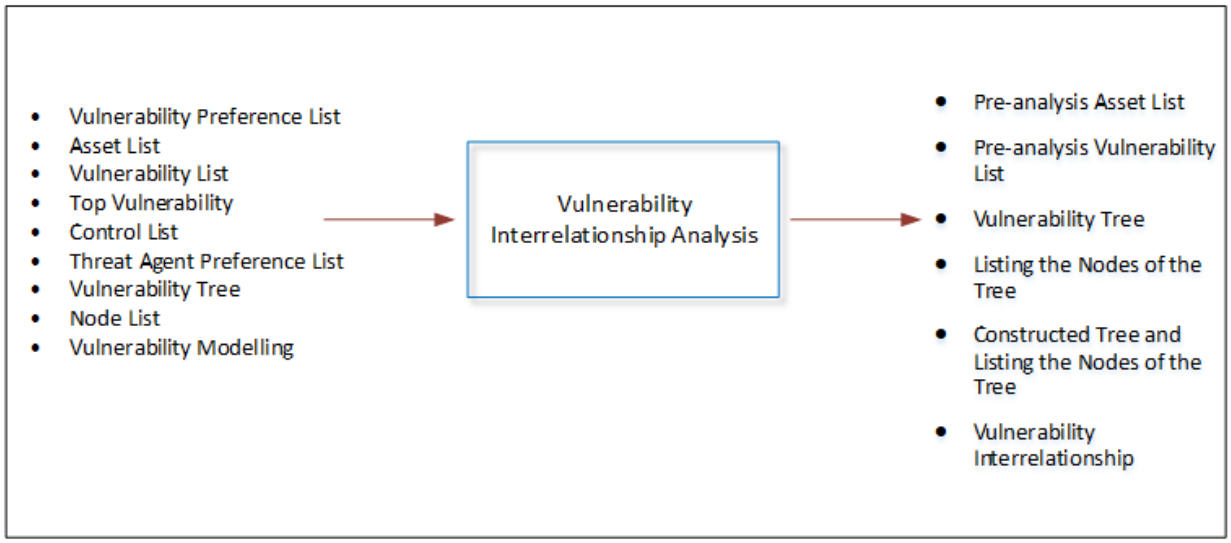


Figure 5-9. Diagram of Vulnerability Interrelationship Analysis.

5.6 Phase 4 Vulnerability Tree Evaluation

5.6.1 Phase 4 overview:

The "Vulnerability Tree Evaluation" is the name of the fourth phase, which involves three processes. The objective of this phase is to assess the vulnerabilities identified in the previous phases, by gathering information on their interrelationships and organizing it for analysis. The three processes are: Stakeholder Evaluation, which involves classifying vulnerabilities to identify interrelationships; Impact Analysis, which focuses on evaluating the potential impact of vulnerabilities; and Vulnerability Statement, which summarizes the findings of the previous processes [11].

- Process 1: Vulnerability Tree Analysis - In this process, the vulnerability tree is analysed to determine the likelihood and impact of each vulnerability. This includes assessing the potential threats, the assets affected, and the potential consequences of an exploit [21].
- Process 2: Risk Assessment - This process involves evaluating the risks associated with each vulnerability and determining the level of risk to the organization. This includes assessing the likelihood and impact of each vulnerability, as well as the overall risk posed to the organization [22].
- Process 3: Vulnerability Statement - In this process, a vulnerability statement is created for each vulnerability identified in the vulnerability tree. This statement includes a description

of the vulnerability, its potential impact, and any recommended countermeasures or mitigation strategies. The vulnerability statement serves as a reference for addressing and mitigating the identified vulnerabilities [23].

5.6.2 Process 4.1: Stakeholder Evaluation

5.6.2.1 Activity 4.1.1: Output Identification

In this process, the stakeholders' information is utilized to determine who will evaluate each process output. Evaluating the outputs is crucial as it is the only way to ensure that all stakeholders' opinions are incorporated and considered throughout the assessment's lifespan [47].

5.6.2.2 Activity 4.1.2: Output Allocation

In this activity, the role of the stakeholders in the enterprise is analysed, and they are assigned the identified outputs from the Output List for evaluation purposes.

The activity uses the following inputs:

- Stakeholder List
- Output List
- Current knowledge of stakeholders

5.6.3 Process 4.2: Impact Analysis

In this activity, the environmental reports obtained in Phase 1 are utilized to determine the various business areas that could potentially be impacted by a threat.

The activity produces the following input:

- Current knowledge of stakeholders
- Technical environment report
- Business environment report
- Physical environment report

The activity produces the following output:

- Impact Field List

5.6.4 Activity 4.2.1: Impact Field Identification

In this process, the business information and the threat information analysed in Phase 3 are integrated to determine the impact that each identified threat will have on the various business levels of the organization.

5.6.5 Activity 4.2.2: Tangible Impact Analysis

In this activity, the impact of a threat on an enterprise is calculated using both threat and asset information. If the preceding processes were successful, applying the results to the equation presented below should provide estimations of the threat's impact on the enterprise, we should get estimations of the threat impact in the enterprise under discussion. The impact equation uses the threat likelihood (L), asset value (A), and threat impact (I) as variables [55]. The impact equation is:

$$f(I) = f(L) * A \tag{5.1}$$

The output of this calculation provides a monetary estimate of the cost impact on the business. However, it is important to note that the equation is only suitable for assessing the impact on tangible assets and is not applicable to intangible assets. It is crucial to address threats that are related to intangible assets separately and use a different method to evaluate their impact [43].

The activity uses the following inputs:

- Threat List: The Threat List is a document that contains all the identified threats towards the enterprise, based on the scoping criteria. This list is developed in Phase 1 - Scoping and is used throughout the risk assessment process to evaluate the potential impact of each threat [78]. The Threat List typically includes a description of each threat, its source or origin, its potential impact on the organization, and any known vulnerabilities or weaknesses that the organization may have that could be exploited by the threat. The list is continually updated and refined as new threats are identified or as the organization's threat environment changes [84].
- Impact Field List: The impact filed list is a catalogue of all the different areas or fields of the enterprise that could potentially be impacted by the threats identified in the Threat List [63]. It provides a comprehensive overview of all the potential impacts that the enterprise may face, allowing for a more thorough and effective risk assessment. By analysing the potential impact to each field, the organization can develop targeted strategies to mitigate the risks associated with each threat. The Impact Field List is a crucial component of the risk assessment process as it helps to identify and prioritize the most critical areas of the enterprise that require protection from potential threats [53].
- Asset List: In this activity, the Asset List is utilized to identify the value of the assets that

are associated with the examined threat. The Asset List contains a comprehensive list of all the identified assets of the enterprise and their corresponding attributes. By utilizing the asset value attribute, this activity can calculate the potential financial impact of the threat to the enterprise [61].

- Threat Agent Preference List: The Threat Agent Preference List is a list that includes all the selected threat agents from the Threat Agent List for further investigation. The attribute that will be used in this activity is the likelihood of the threat agent [33].

5.6.6 Activity 4.2.3: Intangible Impact Analysis

This activity aims to assess the impact of threats on the intangible assets of the enterprise, which were not considered in the previous activity. Instead of trying to calculate a monetary value, this activity focuses on the potential severity of the impact of a threat. The assessment considers how many assets could be affected, interrupted, or disrupted by the manifestation of a specific threat. The more assets that could be impacted, the greater the potential impact of the threat on the enterprise.

The activity uses the following inputs:

- Threat List: listing all the identified threats to the enterprise, based on the scoping criteria.
- Impact Field List: listing all the identified fields of the enterprise that might be affected by the manifestations of the threats included in the Threat List.

5.6.7 Process 4.3: Vulnerability Statement

Vulnerability assessment is the process of detecting weaknesses and potential security risks in various components of the IT environment, such as computer networks, hardware, applications, and systems. The aim of vulnerability assessments is to give security teams and other stakeholders the necessary information to analyse and rank risks so that appropriate measures can be taken to address them [23].

Vulnerability assessments play a crucial role in managing vulnerabilities and risks in IT systems. They are essential for safeguarding systems and data from security breaches and unauthorised access. Vulnerability assessments commonly utilize tools such as vulnerability scanners to detect potential vulnerabilities or risk exposures in an organization's IT infrastructure [28].

5.6.8 Phase 4 Diagram

Figure 5-10. illustrates the process diagram of Vulnerability Tree Evaluation. The diagram

highlights the inputs and outputs of the Phase.

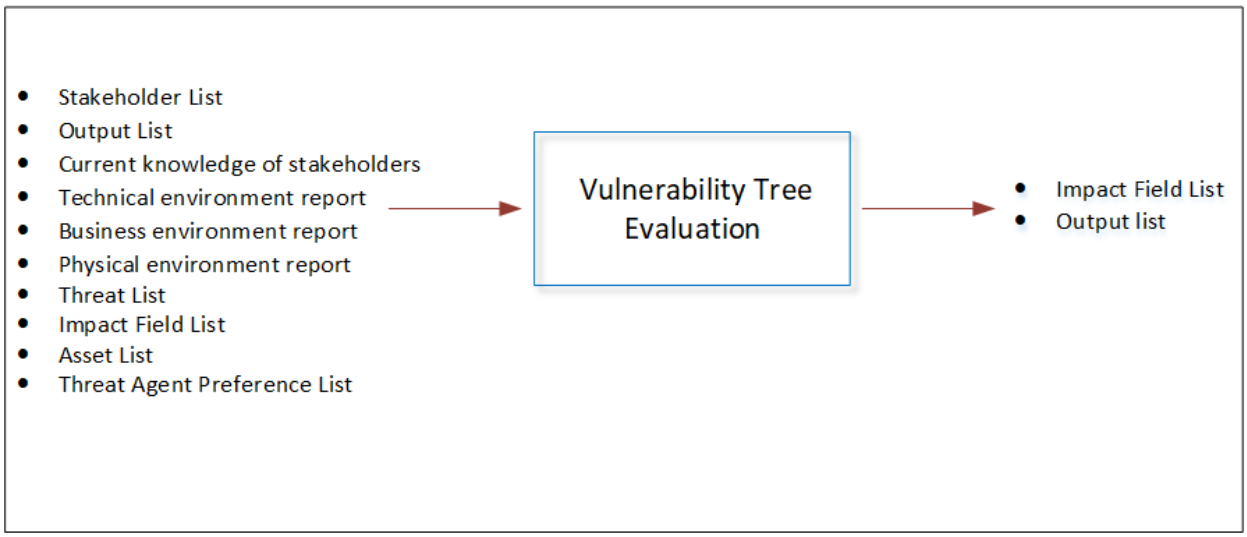


Figure 5-10. Diagram of Vulnerability Tree Evaluation

As Figure 5-10. indicates, Phase 4 focuses on evaluating the interrelationships of vulnerabilities through a Vulnerability Tree, which have been explicitly prepared for in earlier phases. This phase is a combination of previous phases; for instance, data and insights from Phase 1 (identification and analysis), and phase 2 (assessment) are integrated into the development and evaluation of the vulnerability tree and finally by phasing 3 (mitigation strategies) we can propose a proper defence strategy.

5.7 Summary

Chapter 5 delves into the challenges of developing a bio-inspired cybersecurity framework, offering a structured approach to safeguarding digital ecosystems by mimicking the resilience and adaptability of biological systems. The framework is delineated into four critical phases: information environment analysis, vulnerability analysis, vulnerability interrelationship modelling, and vulnerability tree evaluation, each underscoring a unique aspect of cybersecurity from a bio-inspired perspective. In general, this chapter represents a significant step forward in understanding and applying cybersecurity strategies. By taking advantage of insights from biology, the framework introduces a multifaceted approach that enhances resilience, adaptability, and effectiveness in combating cyber threats. Its emphasis on understanding and modelling the interrelationships among vulnerabilities provides a deeper insight into the cybersecurity landscape, allowing for more informed and strategic decision-making. As cyber threats continue to grow in complexity and scale, adopting bio-inspired principles in cybersecurity frameworks offers new pathways for developing more robust, dynamic, and resilient defences.

Chapter 6

Methodology and Implementation

6.1 Overview

This chapter presents the experimental results of our methodology. The primary objective of this project was to develop a cybersecurity model that is clear in its design and capable of predicting potential attacks while identifying interrelationships between vulnerabilities. The experimental results highlight the effectiveness of our approach in achieving this goal. By using the proposed methodology, we were able to create a reliable and efficient cybersecurity framework that can accurately predict potential threats and uncover hidden vulnerabilities. Overall, the results of this project demonstrate the importance of using a structured approach to cybersecurity and highlight the potential of our methodology in enhancing existing security practices.

6.2 Introduction

The danger posed by cyber-attacks remains a top concern for global security and economic stability, with a rise in frequency and severity in recent years. An instance of this was seen in May 2021, when Colonial Pipeline Co., a fuel transportation firm, experienced a ransomware attack that caused a complete shutdown of its pipeline network, leading to fuel shortages and a state of emergency in Florida. According to Interesting Engineering [97], the likelihood of an F-35 fighter jet being downed by a cyber-attack is higher than that of a missile. Therefore, it is clear that cyber-attacks are a significant threat to public safety.

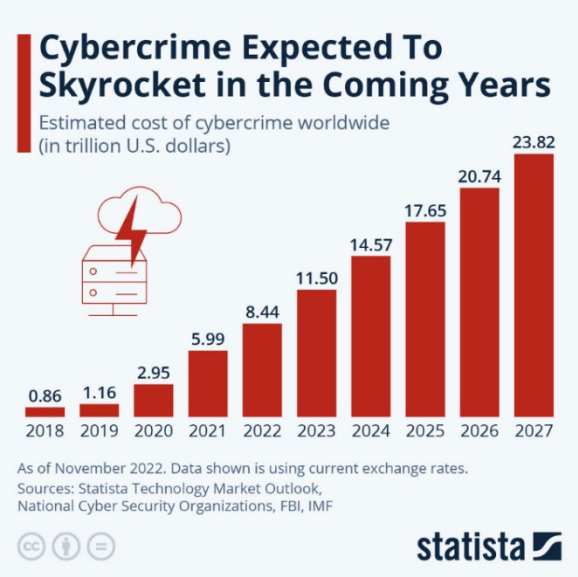


Figure 6-1. Global Cybercrime Expected Costs [99]

As Figure 6-1. illustrates and according to estimates, the worldwide expense of cybercrime reached around 8.4 trillion U.S. dollars in 2022. It is projected that the cost of internet-based criminal incidents will exceed 11 trillion U.S. dollars in 2023. By the year 2026, the annual costs of cybercrime worldwide could surpass 20 trillion U.S. dollars, showing an increase of nearly 150 percent compared to the year 2022. Furthermore, cyber-attacks can have severe consequences. Most cybersecurity tools struggle to distinguish between legitimate and false alarms, as hackers continuously modify and adapt their techniques to evade detection. The traditional approach of using static signatures to detect cyber threats becomes ineffective almost as soon as they are released. This can result in complacency and delayed action when cybersecurity tools detect a threat. However, modern technology has taken a new approach by recognizing the constantly evolving nature of cyber-attacks [99]. One significant challenge with threat detection is that even slight variations in cyber-attacks can deceive the system. Cybercriminals can quickly modify code with minimal adjustments to evade signature detection, and this process can be automated. As a result, attacking computers has become simpler, faster, and cheaper than defending them. To address these challenges, various strategies have been proposed, such as incorporating bio-cybersecurity or taking inspiration from the human immune system to enhance cybersecurity [98]. This study is introduced a technique based on Machin Learning and NLP solution for analysing CVE and predicting possible attacks and find vulnerability interrelationships.

6.2.1 The Information of Cybersecurity Vulnerabilities Sources

The realm of cybersecurity is not a new concept and has relied on various sources and approaches to study security vulnerabilities. One of the most widely used sources is the National Vulnerability Database (NVD), a U.S. government repository that contains information about security vulnerabilities dating back to 1988. Each vulnerability in the NVD is identified by its unique CVE ID. Another example is the Zero-Day Initiative (ZDI), which allows security researchers to privately report Zero-Day vulnerabilities to the affected vendors. ZDI then collaborates with the vendor to release a joint advisory that discloses the vulnerabilities publicly. Other helpful sources are also available:

- CVE ® Program Mission
- https://nvd.nist.gov/
- https://www.cve.org/
- Mozilla's security advisory https://www.mozilla.org/en-US/security/advisories/
- Redhat's product security center
- https://access.redhat.com/security
- The Apache Security Team https://www.apache.org/security/

6.2.2 Information Classification with NLP Support

Recent research has enabled the use of NLP for analysing vulnerability information sources, utilizing database systems [90] and Semantic Web technologies [96]. Some examples include Tablan et al.'s QuestIO, a Natural Language Interface (NLI) for analysing structured information in a domain-independent style [99], and Lopez et al.'s PowerAcqua methodology, which is a system for analysing information stored in heterogeneous, semantic resources [100]. Despite the benefits of these existing works, our proposed work focuses on vulnerability interrelationships, inspired by the human body, which has not been addressed before. Our approach aims to identify relationships between different vulnerabilities, enabling users to address related risks or vulnerabilities simultaneously, and create an environment where sensitive data cannot be easily leaked by attackers.

6.3 Experimental Implementation

Our experimental include different steps:

6.3.1 Download and Read Data

For the first step we have downloaded data from https://nvd.nist.gov/ (Figure 6-2.). CVE records can be obtained through direct querying of the NVD database API in JSON format.

```
[23] file_path = '/content/gdrive/MyDrive/data/Data/cve.json'
     base_url = 'https://services.nvd.nist.gov/rest/json/cves/1.0'
```

Figure 6-2. Download data from NVD.

We attempted to read the CVE.Json file, as shown in Figure 6-3.

```
Retrieving page: 191
Scrapped: 191000
Retrieving page: 192
Scrapped: 192000
Retrieving page: 193
Scrapped: 193000
Retrieving page: 194
Scrapped: 194000
Retrieving page: 195
Scrapped: 195000
Retrieving page: 196
Scrapped: 196000
Retrieving page: 197
Scrapped: 197000
Retrieving page: 198
Scrapped: 198000
Retrieving page: 199
Scrapped: 198260
Completed scrapping..
Total scrapped: 198260
Total unique CVE: 198260
```

Figure 6-3. Read Dataset.

6.3.2 Dataset Description

As Figure 6-3. illustrated the dataset of CVE comprises more than 198260 entries, each providing particular details about the vulnerability, including CVE ID, date of addition and modification, textual description, related CWEs, and other relevant information. The combined file size of all CVEs in JSON format is anticipated to exceed 693MB and continue to expand.

6.3.2.1 Preparing data for further processing or analysis

The datasets utilized in this study were produced by human experts, which makes them less prone to noise and ensures the presence of high-quality textual data. In contrast, extracting value from a data source like Twitter would demand considerably more processing and effort. However, employing human-generated data is not without drawbacks. Unlike machine-generated data, human-created datasets are susceptible to becoming a bottleneck because they require additional effort to accumulate. Consequently, data augmentation was necessary to generate a comprehensive text dataset focused on cybersecurity vulnerabilities [91].

To create a cybersecurity text corpus, various curated data sources were utilized, including CVEs, which provide text descriptions but are often short and limited to one sentence. While other data sources, such as CAPEC and CWE, offer high-quality data, they do not contribute as much information as CVEs. Fortunately, CVEs have been in use for a long time and are integrated into the workflows of major corporations and open-source projects. These entities regularly provide information about new or updated CVEs relevant to their day-to-day activities, which is embedded in CVE records as web links and can be scraped for additional data. Each CVE contains multiple links, many of which lead to common websites, which can be scraped to expand the cybersecurity corpus [92].

To analyse the data, we need to classify them. To implement our model, have the following assumptions:

6.3.2.2 Data Explore

The objective of a data explorer is to analyse, interpret, and extract insights from data. A data explorer seeks to understand patterns, relationships, and trends within a dataset, and to identify opportunities for further analysis or investigation. The ultimate goal of data exploration is to generate actionable insights that can inform decision-making and drive outcomes. Data exploration involves a combination of technical skills, including statistical analysis, data visualization, and programming, as well as an understanding of the domain in which the data is being analysed. In this research the data explorer's objective is to serve as a centralized platform for cybersecurity engineers to review CVE records, along with their descriptions and levels of severity. Although it offers all the necessary functionalities.

6.3.2.3 Visualization Data

To get better understanding of CVE dataset, we downloaded and read cve.csv file (Figure 6-4) and

generate a graphical format of our dataset based on CVE-Id description (Figure 6-5).

```
df = pd.read_csv('/content/gdrive/MyDrive/data/MyNewData/cve_test.csv')
df = df[['cve_id','attack_vector','attack_complexity', 'description']]
df.head(10)
```

	cve_id	attack_vector	attack_complexity	description
0	CVE-2020-16220	ADJACENT_NETWORK	LOW	Patient Information Center iX (PICiX) Versions...
1	CVE-2019-5080	NETWORK	LOW	An exploitable denial-of-service vulnerability...
2	CVE-2020-14856	NETWORK	LOW	Vulnerability in the Oracle Trade Management p...
3	CVE-2019-19748	NETWORK	LOW	The Work Time Calendar app before 4.7.1 for Ji...
4	CVE-2020-12131	NETWORK	LOW	The AirDisk Pro app 5.5.3 for iOS allows XSS v...
5	CVE-2021-24197	NETWORK	LOW	The wpDataTables – Tables & Table Charts premi...
6	CVE-2020-16989	PHYSICAL	LOW	Azure Sphere Elevation of Privilege Vulnerabil...
7	CVE-2022-32775	NETWORK	LOW	An integer overflow vulnerability exists in th...
8	CVE-2019-8139	NETWORK	LOW	A stored cross-site scripting (XSS) vulnerabil...
9	CVE-2022-20446	LOCAL	LOW	In AlwaysOnHotwordDetector of AlwaysOnHotwordD...

Figure 6-4. Read CVE.csv File.

```
#'CVE-Description' visualization with WordCloud
def disp_wordcloud(df):
    stop_words = nltk.corpus.stopwords.words('english')
    wc = WordCloud(max_words=250, stopwords=stop_words).generate(' '.join(df['description'].tolist()))
    plt.figure(figsize=(10,10))
    plt.imshow(wc)
    plt.axis('off')
    plt.show()
    print()
print('\nVisualizing titles from CVE-Description:\n')
disp_wordcloud(cve_df)
```

```
Visualizing titles from CVE-Description:
```

Figure 6-5. Visualization of the CVE Description.

6.3.3 Train-Test Split Data

As Figure 6-6. shows we can split our dataset into two parts named Train-Test set. The train-test split method is utilized for evaluating the effectiveness of machine learning algorithms in predicting outcomes on data that were not used during the model training phase. This technique is quick and straightforward, providing insights into the performance of different algorithms for your specific predictive modelling task [122]. In general, we train the ML models, because during the learning phase of machine learning, the model is trained by utilizing these observations. In essence, the model's parameters are adjusted by inputting these observations to the model. Once the training phase is finished, the performance of the machine learning model is evaluated by applying it to the test set, which consists of new observations. The objective is to assess how the model handles novel data. It is important to ensure that the train and test sets have a similar distribution [123].

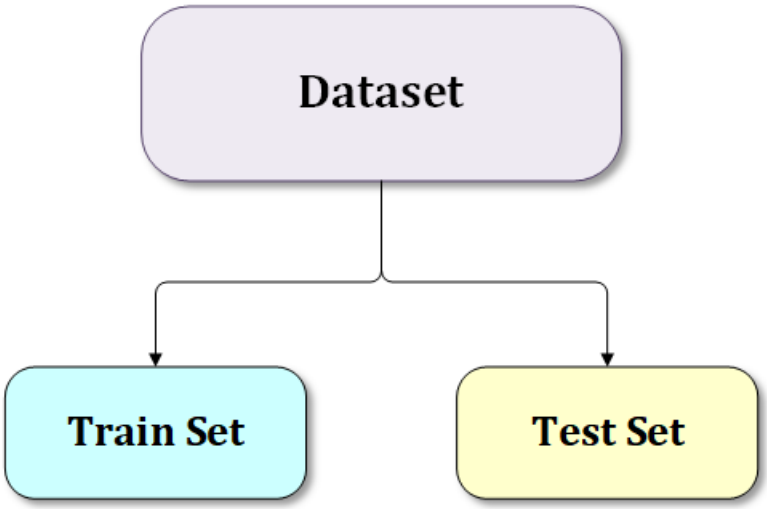
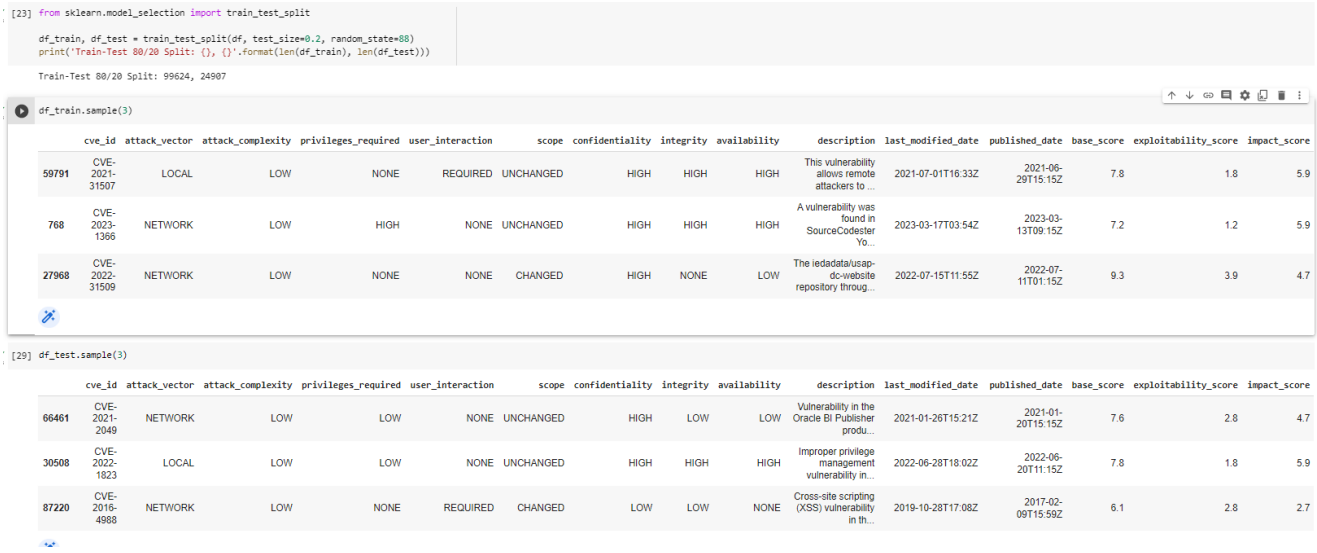


Figure 6-6. Train-Test Split Data

6.3.3.1.1.1 How to Split the Dataset?

According to recent research, the optimal train-test ratio is 80:20, which indicates that 80% of the data should be used for training and the remaining 20% for testing [122]. Therefore, for this study we consider this ratio for our data (Figure 6-7.).



```
[23] from sklearn.model_selection import train_test_split

     df_train, df_test = train_test_split(df, test_size=0.2, random_state=88)
     print('Train-Test 80/20 Split: {}, {}'.format(len(df_train), len(df_test)))

     Train-Test 80/20 Split: 99624, 24907
```

df_train.sample(3)

	cve_id	attack_vector	attack_complexity	privileges_required	user_interaction	scope	confidentiality	integrity	availability	description	last_modified_date	published_date	base_score	exploitability_score	impact_score
59791	CVE-2021-31507	LOCAL	LOW	NONE	REQUIRED	UNCHANGED	HIGH	HIGH	HIGH	This vulnerability allows remote attackers to ...	2021-07-01T16:33Z	2021-06-29T15:15Z	7.8	1.8	5.9
768	CVE-2023-1366	NETWORK	LOW	HIGH	NONE	UNCHANGED	HIGH	HIGH	HIGH	A vulnerability was found in SourceCodester Yo...	2023-03-17T03:54Z	2023-03-13T09:15Z	7.2	1.2	5.9
27968	CVE-2022-31509	NETWORK	LOW	NONE	NONE	CHANGED	HIGH	NONE	LOW	The iedadata/usap-dc-website repository throug...	2022-07-15T11:55Z	2022-07-11T01:15Z	9.3	3.9	4.7

```
[29] df_test.sample(3)
```

	cve_id	attack_vector	attack_complexity	privileges_required	user_interaction	scope	confidentiality	integrity	availability	description	last_modified_date	published_date	base_score	exploitability_score	impact_score
66461	CVE-2021-2049	NETWORK	LOW	LOW	NONE	UNCHANGED	HIGH	LOW	LOW	Vulnerability in the Oracle BI Publisher produ...	2021-01-26T15:21Z	2021-01-20T15:15Z	7.6	2.8	4.7
30508	CVE-2022-1823	LOCAL	LOW	LOW	NONE	UNCHANGED	HIGH	HIGH	HIGH	Improper privilege management vulnerability in...	2022-06-28T18:02Z	2022-06-20T11:15Z	7.8	1.8	5.9
87220	CVE-2016-4988	NETWORK	LOW	NONE	REQUIRED	CHANGED	LOW	LOW	NONE	Cross-site scripting (XSS) vulnerability in th...	2019-10-28T17:08Z	2017-02-09T15:59Z	6.1	2.8	2.7

Figure 6-7. Train-Test Ratio Dataset (80:20)

6.4 Prediction

According to research and studies, the Common Vulnerability Scoring System (CVSS) score is one common method used to anticipate vulnerabilities. This score is assigned by humans to indicate the degree of severity or risk that a CVE poses. Cybersecurity engineers use CVSS scores to prioritize the order in which CVE records should be evaluated and addressed, making it a valuable tool in their arsenal [124]. However, there are various approaches and frameworks that can be used for predicting cybersecurity risk besides the CVSS [125].

Some of these include:

1. **NIST Cybersecurity Framework:** The NIST framework provides guidelines, standards, and best practices for managing and reducing cybersecurity risk. It includes five core functions: identify, protect, detect, respond, and recover [126].
2. **MITRE ATT & CK Framework:** The MITRE ATT & CK framework provides a comprehensive knowledge base of tactics, techniques, and procedures used by threat actors. It can be used to identify potential vulnerabilities and develop effective mitigation strategies [127].
3. **FAIR (Factor Analysis of Information Risk):** FAIR is a quantitative risk assessment framework that uses mathematical models to calculate the probability and impact of various cyber threats. It can help organizations prioritize their cybersecurity investments and resources [128].
4. **CIS Controls:** The CIS Controls are a set of best practices and guidelines for securing IT

systems and networks. They provide a prioritized list of security actions that organizations can take to improve their cybersecurity posture [129].

Overall, there are several approaches and frameworks available for predicting cybersecurity risk. The choice of framework will depend on the specific needs and context of the organization, as well as the nature of the threats they face [130].

Previously, there have been efforts to anticipate potential cyber-attacks through textual analysis of vulnerabilities. Bozorgi et al. [90] employed Bag-of-Words and SVM classification to categorize vulnerabilities into various exploitability groups established by the researchers. While the study did not explicitly categorize the CVSS vector metrics, it demonstrated that textual analysis could be useful in assessing the severity of vulnerabilities. In another study by Khazaei et al. [91] the prediction was further enhanced by utilizing Bag-of-Words, SVM, Random Forest, and fuzzy system to classify the numerical severity score, generating an integer value between 0 and 10 without extracting the CVSS metric vector. Despite the challenges, research has indicated the possibility of predicting the CVSS through text mining. Elbaz et al. [93] utilized Bag-of-Word to process text, but they improved the transparency of the predicted score by classifying the metric vectors rather than the numerical score. Similarly, Yin et al. [94] leveraged transfer learning by employing a pre-trained ExBERT model.

In this study we will measure CVSS to predict the possible vulnerabilities based on different model. The objective of this study is to use the NLP utilized in CVE descriptions to anticipate whether a vulnerability is exploitable or not and find its interrelationship.

6.4.1 Create Models

This study has considered two algorithms Gradient Sensitivity (GS) and (Gradient Sensitivity Input (GI) to extract words that exerted the greatest impact on the sequence classifier of the model, thereby enhancing its explainability. These techniques facilitated a better understanding of classification process and the words that significantly influenced the model's decision-making. The extracted words could also aid users in comprehending the output and assessing its credibility [95]. For a comparative analysis focusing on the decision to utilize GS and GSI techniques within an ML framework developed for cybersecurity, several key considerations come into play. These considerations are deeply rooted in the unique demands of cybersecurity data analysis, the nature of the vulnerabilities to be detected or predicted, and the overall objectives of the bio-inspired cybersecurity framework. These models are determined by their capability to deal effectively with high-dimensional and imbalanced data, their contribution to enhancing model explainability, their adaptability to evolving threat landscapes, and their comparative advantages in sensitivity analysis

and feature importance assessment [131]. This strategic choice aligns with the overarching goal of developing ML models that are not only predictive but also interpretable and adaptable, catering to the critical needs of cybersecurity applications.

- Gradient Sensitivity: The GS method is based on the gradient of inputs relative to each embedding and is inspired by the backpropagation algorithm in vision [6]. The algorithm determines the contribution of each input word to the final decision by calculating the first derivative of the input embedding. To explain the concept more formally, let's consider a classification model that uses a scoring function S to classify an input with embedding as X and assign it to a golden class c. The goal of the GS model is to identify which i-th dimension of the embedding c(x) contributed to the scoring S that led to the decision of the class label [131].

$$R_i^{GS}(x) = \frac{\partial f_c(X)}{\partial X_i} \tag{6.1}$$

Where the derivative pertains to the i-th dimension of the X embedding.

Figure 6-8. shows the Python implementation of the code for more clarification.

```
def backward_gradient(sensitivity_grads):
classifier_output = func_activations['model.classifier']
embedding_output = func_activations['model.bert.embeddings']
sensitivity_grads = torch.autograd.grad(classifier_output, embedding_output,
                                        grad_outputs=sensitivity_grads,
                                        retain_graph=True)[0]
return sensitivity_grads
```

Figure 6-8. Python Implementation of the Code with GS Model.

- Gradient Sensitivity Input: The GI technique is an extension of the GS approach that involves multiplying the GS output by the i-th dimension of the input embedding. By doing so, this method calculates the extent to which the output would differ if the input were modified [97].

$$R_i^{GI}(X) = X_i \cdot R_i^{GS}(X_i) \tag{6.2}$$

Figure 6-9. illustrates the Python implementation of the code.

```
def backward_gradient_input(sensitivity_grads):
    embedding_output = func_activations['model.bert.embeddings']
return backward_gradient(sensitivity_grads) * embedding_output
```

Figure 6-9. Python Implementation of the Code with GI Model.

The relationship between vulnerabilities is defined by the following attribute [57]:"

- Availability (access)
- Attack Vector
- Attack Complexity
- Confidentiality
- Privilege Required
- User Interaction
- Scope
- Integrity

In our methodology, we identified the relationship between vulnerabilities using a specific attribute. This attribute provided us with valuable insights into how vulnerabilities are connected and allowed us to develop a more comprehensive understanding of the cybersecurity landscape. By leveraging this attribute, we were able to identify potential attack vectors and mitigate security risks more effectively. Overall, the use of this attribute proved to be a crucial factor in the success of our methodology, and its inclusion in our cybersecurity framework has significantly improved our ability to identify, prevent, and respond to cyber threats. Therefore, we have exploited eight metrics from CVE based on the vulnerability relationship and misuser the prediction score (Figure 6-10).

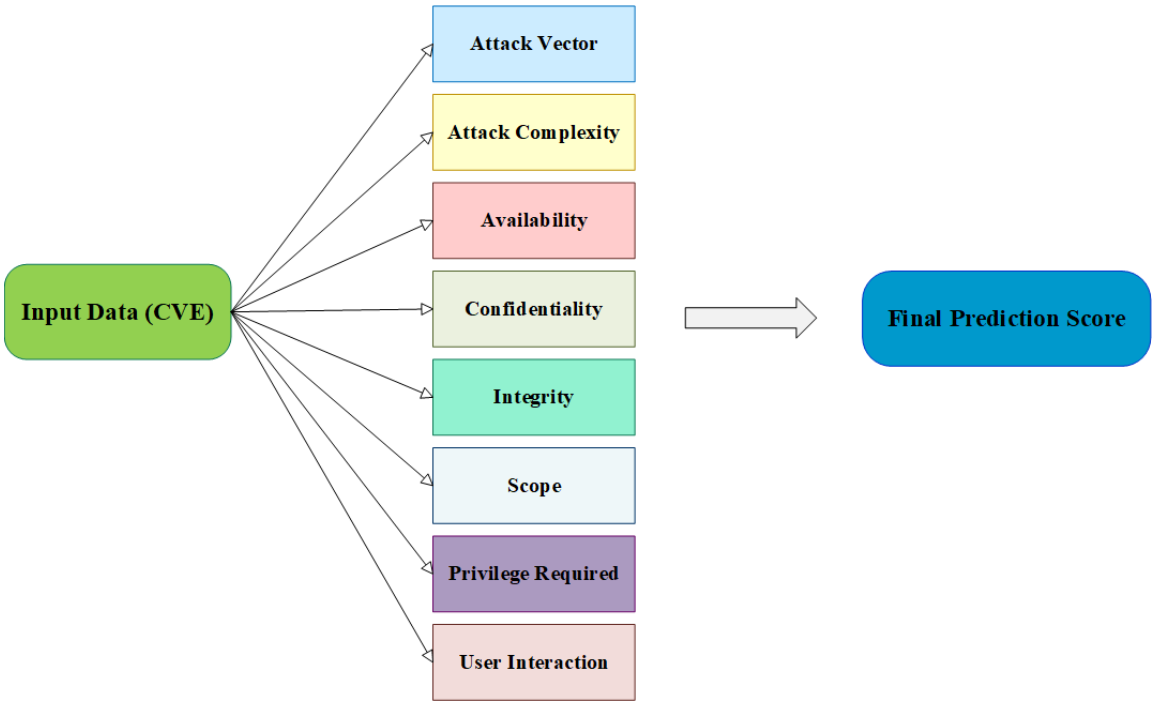


Figure 6-10. Predict the Possible Vulnerabilities Based on Different Models.

Figure 6-11. shows the outcome of our creating models. These models will use for the rest of our proposed work.

Where:

- **AV** is Attack Vector
- **Ac** is Attack Complexity
- **PR** is Privilege Required
- **UI** User Interaction
- **S** is Scope
- **C** is Confidentiality
- **I** is Integrity
- **A** is Availability

All of these models exploited from the CVE.Json.

```
ac_output_dir = '/content/gdrive/MyDrive/Model/AC'
a_output_dir = '/content/gdrive/MyDrive/Model/AI'
av_output_dir = '/content/gdrive/MyDrive/Model/AV'
c_output_dir = '/content/gdrive/MyDrive/Model/CI'
i_output_dir = '/content/gdrive/MyDrive/Model/II'
pr_output_dir = '/content/gdrive/MyDrive/Model/PR'
s_output_dir = '/content/gdrive/MyDrive/Model/SC'
ui_output_dir = '/content/gdrive/MyDrive/Model/UI'
print ('All models create and save to the directory.')
```

```
All models create and save to the directory.
```

Figure 6-11. Outcomes of Creating Model

As we mentioned earlier the implementation of GS and GI techniques within our framework has markedly improved the model's explainability. This is a crucial advancement, as it not only aids in the precise identification of vulnerabilities but also provides users with an understandable rationale behind the model's predictions. Such transparency is invaluable in the cybersecurity field, where understanding the 'why' behind a vulnerability prediction can significantly impact mitigation strategies and policy decisions. Our comparative analysis and methodology have demonstrated that models equipped with GS and GI are not only predictive but also interpretable and adaptable, catering to the critical needs of cybersecurity applications. Furthermore, these techniques provide users with an understandable rationale behind the model's predictions. This achievement is a foundational shift in how cybersecurity models communicate with their human counterparts. In an environment where decision-making processes are becoming increasingly automated, maintaining a channel through which human users can comprehend and validate these processes is essential. This transparency ensures that decisions made by the model can be trusted and, more importantly, acted upon with confidence.

6.4.2 Accuracy of Vulnerability Prediction

As previously mentioned, in order to measure the prediction score, it was necessary to create sub-models from the CVE dataset. We then evaluated the accuracy of each sub-model based on the CVSS. This approach allowed us to better understand the effectiveness of each sub-model in predicting the severity of vulnerabilities. By using the CVSS scoring system, we were able to quantify the accuracy of our predictions and make informed decisions about how to improve our models. Overall, this approach helped us to develop a more effective and reliable cybersecurity

framework. The CVSS score is calculated based on eight distinct metrics that are determined independently and then combined [125]. To implementing our work, we used Bidirectional Encoder Representations from Transformers (BERT) model which a popular model of NLP. BERT is a pre-trained language model developed by Google AI Language. It is based on the Transformer architecture and is designed to understand the context of words in a sentence by considering the surrounding words on both sides (i.e., bidirectional) [132]. BERT is trained on a large amount of text data, such as Wikipedia articles and books, using a masked language modelling task and a next sentence prediction task. In the masked language modelling task, BERT is trained to predict the missing word in a sentence. In the next sentence prediction task, BERT is trained to determine whether two sentences are related or not. The main advantage of BERT is that it can be fine-tuned for a wide range of natural language processing tasks, such as question-answering, sentiment analysis, and named entity recognition, by adding a few additional layers on top of the pre-trained model. This allows developers to create more accurate and effective NLP applications with less training data and fewer resources [132].

Furthermore, for analysing the CVSS score we consider three main components score of CVSS score:

- Base score: Which is referring to the severity rating of a vulnerability, based on the CVSS. It represents the intrinsic characteristics of a vulnerability. It ranges from 0 to 10 and is calculated using a formula that takes into account factors such as the attack vector, attack complexity, and authentication requirements. The base score is then used as a starting point to calculate the overall CVSS score, which also includes temporal and environmental factors that can affect the severity of the vulnerability in a specific context [133].
- Impact score: The impact score refers to the portion of the CVSS that assesses the potential impact of a vulnerability if it is successfully exploited. The impact score takes into account three factors: confidentiality, integrity, and availability. Each factor is scored on a scale from 0 to 10, with 10 being the most severe impact. The three scores are then combined to produce an overall impact score [133].
- Exploitability score: It refers to the portion of the CVSS that assesses the likelihood that a vulnerability will be exploited in the wild. The exploitability score is one of three main components of the CVSS score, along with the base score and the impact score. It is intended to provide an estimate of the ease of exploitation and the level of skill and

resources required to exploit the vulnerability [134]. The exploitability score takes into account three factors: attack vector, complexity, and privilege required. Each factor is scored on a scale from 0 to 1, with 1 being the most severe. The three scores are then combined to produce an overall exploitability score [133]. Attack vector refers to the way in which an attacker can exploit the vulnerability, such as over a network or through physical access. Complexity refers to the level of expertise and resources required to exploit the vulnerability. Privilege required refers to the level of access an attacker needs to successfully exploit the vulnerability [134].

Figure 6-12 provides a result representation of the performance of eight trained models in terms of their prediction accuracy and mean confidence. The results show that all of the models performed exceptionally well, with consistently high accuracy and mean confidence across all the tested data. Accuracy refers to the degree to which the predicted output of a model matches the actual output, and mean confidence refers to the degree of certainty with which the model makes its predictions. High accuracy and mean confidence are essential qualities of a good machine learning model, as they indicate that the proposed model is making accurate and reliable predictions.

The fact that all eight models demonstrated outstanding prediction accuracy and mean confidence is a significant achievement, as it indicates that the models are robust and effective in their respective domains. This is particularly noteworthy given the complexity and variability of the data being analysed, which can often present significant challenges for machine learning models. Overall, the results shown in Figure 6-12. and Figure 6-13. suggest that the trained models are capable of making accurate and reliable predictions across a wide range of data and scenarios. This bodes well for the future of machine learning in various fields, as it demonstrates the potential for these models to drive innovation and progress in a range of applications, from healthcare and finance to engineering and beyond.

```
Predictions:

  AV: Network          Confidence: 0.9864
  AC: Low              Confidence: 0.6990
  PR: None             Confidence: 0.8334
  UI: None             Confidence: 0.9243
  S : Unchanged        Confidence: 0.8818
  C : High             Confidence: 0.9277
  I : High             Confidence: 0.8758
  A : High             Confidence: 0.9326

  Base score: 9.8
  Impact score: 5.9
  Exploitability score: 3.9

Truths:

  AV: Network
  AC: Low
  PR: Low
  UI: None
  S : Unchanged
  C : High
  I : High
  A : High

  Base score: 8.8
  Impact score: 5.9
  Exploitability score: 2.8
```

Figure 6-12. Prediction Results

In addition to the risks associated with the lack or gap of authentication checks on SAP NetWeaver Java, there is also the potential for unauthorised users to fully compromise a targeted system. This is a serious threat, as it could lead to significant damage to the affected organization's data, systems, and reputation.

SAP NetWeaver Java is a key component of the SAP NetWeaver platform, which is widely used in many industries to develop and deploy enterprise-level applications. The platform provides a range of features and services, including a web application server, development tools, and APIs, all of which are designed to support the creation of scalable, secure, and interoperable applications [135].

Figure 6-13 shows that by using a specific methodology, it is possible to predict this type of threat. This is a significant achievement, as it enables organizations to proactively identify potential security risks associated with SAP NetWeaver Java and take steps to address them before they can be exploited by malicious actors. By leveraging the power of NLP and data analysis, security professionals can stay ahead of the curve when it comes to identifying and mitigating potential threats to their systems and data. This approach can help to ensure that organizations remain secure and resilient against emerging threats, even as the threat landscape continues to evolve and become increasingly complex.

```
Predictions:

AV: Network          Confidence: 0.9954
AC: Low              Confidence: 0.9283
PR: None             Confidence: 0.9885
UI: None             Confidence: 0.9540
S : Unchanged        Confidence: 0.9027
C : High             Confidence: 0.9095
I : High             Confidence: 0.8618
A : High             Confidence: 0.9607

Base score: 9.8
Impact score: 5.9
Exploitability score: 3.9
```

Figure 6-13. Prediction of the Unauthorised Access to a System or Web Service.

Figure 6-14 depicts the accuracy of the proposed model, which is designed to predict threats associated. The results show that the suggested model is highly accurate, with a consistently high accuracy score across all the tested data.

```
Proposed Model Accuracy:
   AV : 0.9640
   AC : 0.9948
   PR : 0.9452
   UI : 0.9877
   S  : 0.9936
   C  : 0.9682
   I  : 0.9793
   A  : 0.9797
```

Figure 6-14. Accuracy of Proposed Model.

Chapter 7

Finding The Vulnerability Interrelationship

7.1 Overview

Vulnerability interrelationship (Cyber vulnerability semantic similarity) specifically refers to the degree of similarity between the meanings or contexts of different cybersecurity vulnerabilities. This involves Analysing and comparing the characteristics and attributes of various vulnerabilities to identify commonalities and patterns. By understanding the vulnerabilities interrelationship between different cybersecurity vulnerabilities, security professionals can better anticipate potential threats and develop more effective mitigation strategies [165].

To find the similarity we used K-Nearest Neighbor (KNN) algorithm which operates by comparing the similarity between a new case or data point and existing cases, and then assigning the new case to the category that is most similar to the existing categories. It stores all of the available data and utilizes similarity measures to classify a new data point [136].

7.1.1 Processing Index of Dataset

Processing an index in machine learning involves using it to access or manipulate data or models in some way. For example, you might use a feature index to select a subset of features for training a model, or you might use a data index to split a dataset into training and testing sets Machine learning involves a wide range of techniques that enable computers to learn from data and make decisions based on that knowledge. Processing an index is an essential component of machine

learning, as it enables us to access or manipulate data or models in various ways.

One common application of index processing is in selecting a subset of features for training a model. Features are the inputs to a machine learning model, and selecting the right set of features is critical to the model's performance. By using a feature index, we can select the most relevant features for a given task and use them to train the model. Another way to use an index in machine learning is to split a dataset into training and testing sets. This technique is essential for evaluating the performance of a model accurately. By using a data index, we can randomly split the data into two sets: one for training the model and one for testing its performance. This helps to ensure that the model has not merely memorized the training data but has learned to generalize to new, unseen data [137].

Figure 7-1 illustrates the process of index processing in machine learning. As data flows through the system, various indices are used to access or manipulate the data in different ways. These indices can be featuring indices, data indices, or other types of indices depending on the specific application.

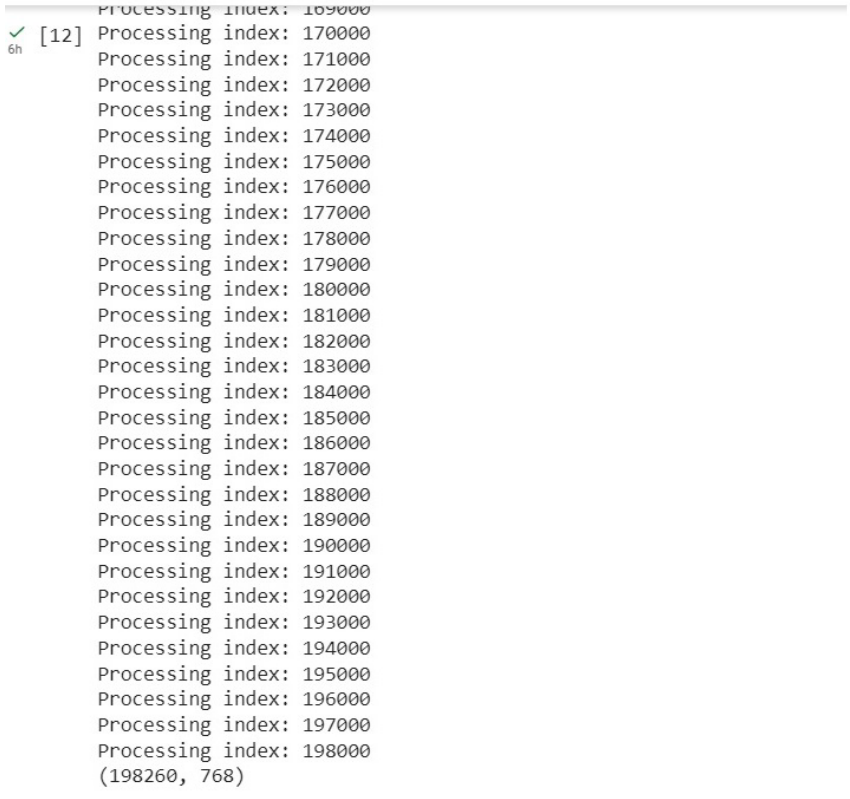


Figure 7-1. Process of Index Processing in Our Proposed Model

After the indexing process is completed, the system proceeds to generate a file named "cve_vectorized.csv," which is then saved to a designated location on your Google Drive, as shown in Figure 7-2. This file contains the vectorized representation of the CVE data that was indexed. This vectorized representation enables efficient and effective querying of the CVE data based on various search criteria. Additionally, it allows for the application of machine learning algorithms to identify patterns and trends in the data, aiding in the identification of potential security threats and vulnerabilities.

The storage of the file on Google Drive provides a secure and easily accessible location for the data, facilitating collaboration among team members and ensuring that the data is always available when needed. Overall, this process helps to streamline the management and analysis of CVE data, leading to improved cybersecurity practices and stronger overall security for your organization.

```
[ ]  Processing index: 193000
     Processing index: 194000
     Processing index: 195000
     Processing index: 196000
     Processing index: 197000
     Processing index: 198000
     (198260, 768)
```

```
df_vectorized = pd.DataFrame(data=vecs_stacked)
df_vectorized.insert(loc=0, column='cve_id', value=df['cve_id'])

save_path = '/content/gdrive/MyDrive/Mydata200/cve_vectorized.csv'
df_vectorized.to_csv(save_path)
```

Figure 7-2. Generate CVE_Vectorized.csv.

To identify the interrelation of vulnerabilities, a common approach is to analyse the descriptions of CVEs and identify similarities among them. In this context, the input for the analysis is a set of CVE descriptions that are deemed to be related to a particular security vulnerability. Table 7.1. presents the details of the input used for the analysis, which typically includes information about the CVE ID, the associated products or systems, and a brief description of the vulnerability. By processing this information, it is possible to identify similarities and patterns among the CVE descriptions that may indicate a common source or underlying vulnerability. Once the CVE descriptions are analysed, the next step is to use this information to improve the security of the systems or products that are affected by the vulnerabilities. This can involve implementing patches or other security measures to address the identified vulnerabilities and prevent potential attacks.

Overall, Analysing the interrelation of vulnerabilities is an important task in the field of cybersecurity, as it helps organizations to proactively identify and address potential security risks before they can be exploited by malicious actors. By using techniques such as CVE analysis, security professionals can stay ahead of the curve and ensure that their systems and products remain secure and resilient against emerging threats.

Table 7.1. The detail of CVE-ID input.

CVE_ID	Vulnerability Type(s)	Description
CVE-2016-9560	Overflow	Stack-based buffer overflow in the jpc_tsfb_getbands 2 function in jpc_tsfb.c in JasPer before 1.900.30 allows.
CVE-2016-8866	Overflow	Remote attackers to have unspecified impact via a crafted image.
CVE-2018-12591	Execute Code	Ubiquiti Networks EdgeSwitch version 1.7.3 and prior suffer from an improperly neutralized element in an OS command.
CVE-2018-12590	Execute Code	Due to lack of protection on the admin CLI, leading to code execution and privilege escalation greater than administrators themselves are allowed. An attacker with access to an admin account could escape the restricted CLI and execute arbitrary shell instructions.
CVE-2017-5420	XSS	'A "javascript:" url loaded by a malicious page can obfuscate its location by blanking the URL displayed in the addressbar, allowing for an attacker to spoof an existing page without the malicious page\'s address being displayed correctly. This vulnerability affects Firefox < 52.'
CVE-2007-0994	Bypass, Execute Code, XSS	A regression error in Mozilla Firefox 2.x before 2.0.0.2 and 1.x before 1.5.0.10, and SeaMonkey 1.1 before 1.1.1 and 1.0 before 1.0.8, allows remote attackers to execute arbitrary JavaScript as the user via an HTML mail message with a javascript: URI in an (1) img, (2) link, or (3) style tag, which bypasses the access checks and executes code with chrome privileges.
CVE-2020-15778	DoS	Cloud Vulnerability caused by spoofing

The graphical representation of the process and its outcomes can be seen in Figure 7-3, while a more detailed breakdown of the results can be found in Table 7.2. Figure 7-3. provides a clear visual representation of the process flow, highlighting the various steps involved in achieving the desired outcomes. Table 4, on the other hand, presents a comprehensive breakdown of the results obtained from the process, including metrics such as accuracy, precision, recall, and F1 score. These metrics are essential in evaluating the effectiveness of the process and determining whether

it met the desired objectives. Furthermore, the outcomes presented in both Figure 7-3. and Table 7.2. provide valuable insights into the strengths and weaknesses of the process, enabling further refinement and improvement in future iterations. This level of evaluation and analysis is crucial in ensuring that the process continues to meet the evolving needs of the organization and delivers tangible benefits to all stakeholders. In general, the presentation of the process and its outcomes in both visual and tabular formats enable easy understanding, interpretation, and communication of the results to stakeholders, facilitating informed decision-making and ensuring the continued success of the process.

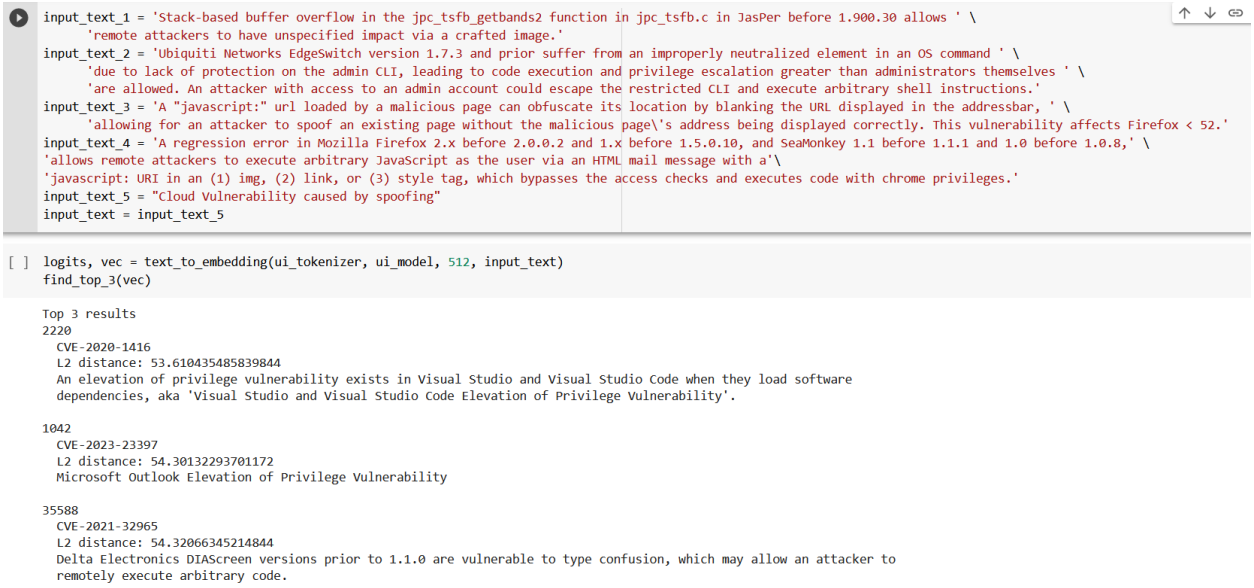


Figure 7-3. Vulnerability Interrelationship

Table 7.2. Experimental Results of Vulnerability Interrelationship

CVE_ID	Vulnerability Type(s)	Description
CVE-2020-1416	Execute Code	An elevation of privilege vulnerability exists in Visual Studio and Visual Studio Code when they load software dependencies, aka 'Visual Studio and Visual Studio Code Elevation of Privilege Vulnerability'.
CVE-2023-23397	SQL Injection	Microsoft Outlook Elevation of Privilege Vulnerability
CVE-2021-32965	Execute Code	Delta Electronics DIAScreen versions prior to 1.1.0 are vulnerable to type confusion, which may allow an attacker to remotely execute arbitrary code.

These results provide an analysis of the similarity between different vulnerabilities based on the L2 distance, also known as the Euclidean distance. L2 distance is a commonly used measure of

the distance between two points in a multidimensional space and is often used in machine learning and other fields to calculate the similarity or dissimilarity between data points.

In our analysis, we have used the L2 distance to calculate the similarity between different vulnerabilities, based on their characteristics and features. For example, we have calculated the L2 distance between the characteristics of a data breach vulnerability and a system vulnerability, to determine how similar or dissimilar these vulnerabilities are.

The L2 distance between two points (x1, y1) and (x2, y2) in a two-dimensional space can be calculated using the formula [138]:

$$d = \sqrt{((x2 - x1)^2 + (y2 - y1)^2)} \quad (7.1)$$

In machine learning, L2 distance is often used as a measure of similarity between two data points in clustering or classification tasks. For instance, in the K-Nearest Neighbors (KNN) algorithm, the L2 distance is used to find the k nearest neighbors of a test data point among the training data points. The closer the data points are in the L2 distance, the more similar they are considered to be [138].

Our analysis has shown that there is a significant degree of similarity (around 54%) between different vulnerabilities, based on their characteristics and features. To present the results, we have considered the top three vulnerabilities with the highest degree of similarity. These results can be used to identify patterns and trends in vulnerability data, and to develop more effective strategies for mitigating vulnerabilities and reducing the risk of cyberattacks.

In summary, the L2 distance is a commonly used measure of similarity between data points and can be used to analyse and categorize vulnerabilities in the context of biology and cybersecurity. Based on our analysis, we have found that there is a considerable level of similarity between various vulnerabilities that exist in computer systems. This discovery can be instrumental in devising more efficient and successful approaches to mitigating these vulnerabilities and, consequently, minimizing the risk of cyberattacks.

In essence, our analysis suggests that instead of treating each vulnerability in isolation, we can group them according to their interrelationship as well as similarities and address them collectively. By doing so, we can identify patterns and commonalities between different vulnerabilities, which can enable us to develop more robust and comprehensive strategies for preventing cyberattacks. By utilizing these findings, we can adopt a more proactive and targeted

approach towards securing computer systems against cyber threats. This can involve implementing specific security measures to address vulnerabilities that share similar characteristics, such as prioritizing patching or updating software, implementing access controls, or deploying intrusion detection systems. Overall, our results highlight the importance of vulnerability interrelationship approach to cybersecurity and utilizing all available information to create effective and tailored strategies for mitigating vulnerabilities and preventing cyberattacks.

7.2 Vulnerability Tree

Now that we have conducted a comprehensive analysis of the link between biology and cybersecurity and identified potential threats and vulnerabilities, we can use this information to create a vulnerability tree. To create this vulnerability tree, we have utilized the Natural Language Toolkit (NLTK). NLTK is a powerful framework for creating Python applications that process natural language data. It offers a range of user-friendly interfaces to more than 50 language resources and databases, such as WordNet, which can be used to analyse and categorize text data. The toolkit also provides a range of text processing tools, including classification, tokenization, stemming, tagging, parsing, and semantic reasoning, that can be used to analyse natural language data [139]. Furthermore, the NLTK is a versatile framework that can be used to analyse natural language data and create vulnerability trees in the context of biology and cybersecurity. With its user-friendly interfaces, range of text processing tools, and wrappers for powerful NLP libraries, NLTK offers a comprehensive solution for analysing and categorizing text data. Moreover, its thriving online community forum provides a valuable resource for discussion, learning, and support [139]. The decision to choose NLTK over other libraries can be influenced by various factors, including the specific requirements of the project, the familiarity of the research team with the tool, and the types of analyse needed. Here's a breakdown of why NLTK chosen and how it compares to other options [186]:

- **Comprehensive Library**: NLTK provides a wide array of tools and resources for almost every NLP task, making it a one-stop-shop for many researchers, especially those in the academic field.
- **Ease of Learning and Use**: Its straightforward and intuitive API is particularly appealing for beginners and for those who prioritize rapid development and testing of NLP concepts.
- **Strong Community and Documentation**: NLTK benefits from a large community of

users and extensive documentation, which can be invaluable for troubleshooting and learning, especially for students and new researchers.

- **Educational Resource**: NLTK is widely used in educational settings, making it a familiar choice for researchers with an academic background in NLP.
- **Versatility**: It supports a variety of tasks essential in cybersecurity applications, such as tokenization, part-of-speech tagging, named entity recognition, and sentiment analysis.

Figure 7-4 shows different steps of generating a vulnerability tree.

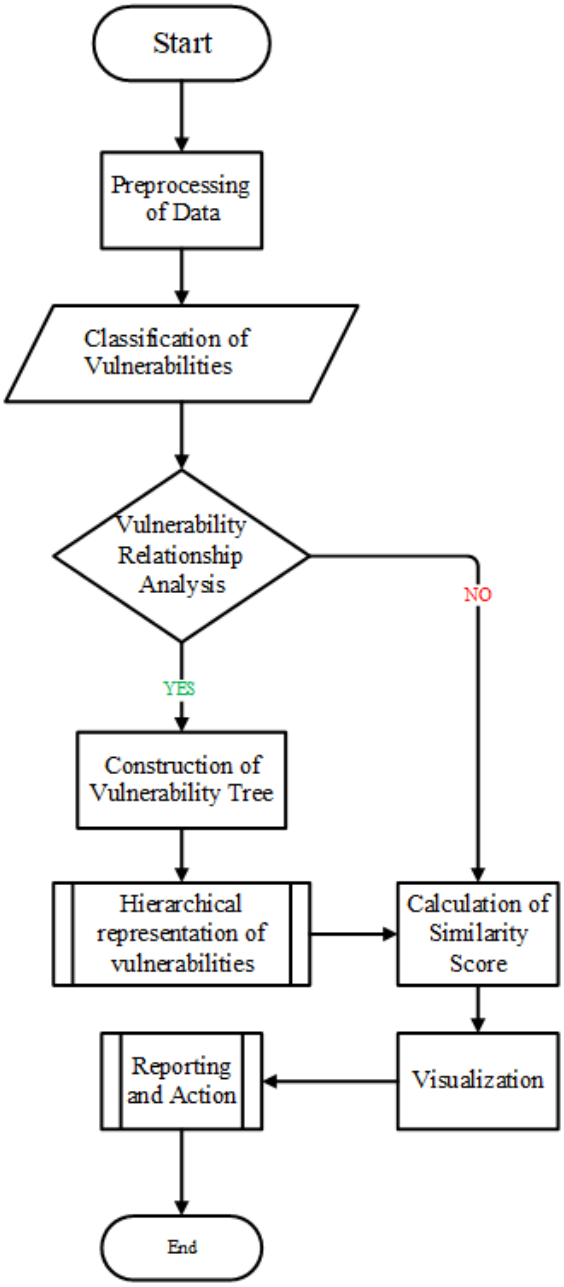


Figure 7-4 The Process of Generating Vulnerability Tree.

The methodology is illustrated in Figure 7.4 and has the following main steps:

1. **Initial Data Collection**
   - Input: Vulnerability data (CVE databases)
   - Description: Collect data on vulnerabilities for analysis.
2. **Preprocessing of Data**
   - Tools: NLP, Data Cleaning
   - Description: Clean and extract usable information from raw data.
3. **Identification of Vulnerabilities**
   - Tools: Automated scanning tools, Manual inspection
   - Description: Identify potential vulnerabilities in the pre-processed data.
4. **Classification of Vulnerabilities**
   - Criteria: Severity, Type, Affected Components
   - Description: Organize vulnerabilities into categories based on predefined criteria.
5. **Vulnerability Relationship Analysis** (Yes/No decision)
   - Yes: Continue to step 6
   - No: Go to step 7 (if relationship analysis not needed)
6. **Construction of Vulnerability Tree**
   - Output: Hierarchical representation of vulnerabilities
   - Description: Build a tree structure showing connections between vulnerabilities, highlighting dependencies and impact flow.
7. **Calculation of Similarity Score**
   - Tools: Similarity algorithms
   - Description: Calculate similarity scores between vulnerabilities to identify shared characteristics or mitigation strategies.
8. **Visualization**
   - Input: Vulnerability tree, Similarity scores
   - Output: Visual diagrams, charts
   - Description: Create visualizations using the tree and scores to understand vulnerability extent and nature.
9. **Reporting and Action**

- Input: analysed data, Vulnerability tree, Similarity scores
- Output: Detailed reports, Action plans
- Description: Generate reports based on the analysis and propose actionable steps based on the tree and scores.

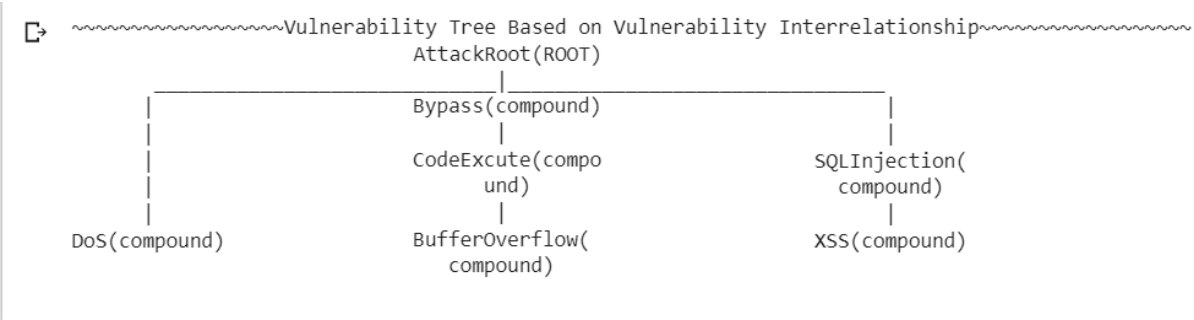


Figure 7-5. Vulnerabilities Tree

Figure 7-5. indicates a visual representation of the vulnerabilities tree that we have created using the NLTK in the context of the link between different vulnerabilities. The figure shows the relationship between different vulnerabilities that we have identified in our analysis, such as user interaction model and CVE description. By analysing this vulnerability tree, we can gain a better understanding of the root causes and potential impacts of different vulnerabilities, as well as the interconnections between them. For instance, the figure shows that a system vulnerability can lead to a data breach, which in turn can be exploited by an insider threat. Similarly, an insider threat can exploit a vulnerability in a system to gain unauthorised access to sensitive data. These interconnections highlight the complexity and multi-layered nature of the link different vulnerabilities and emphasize the importance of taking a holistic and proactive approach to addressing these vulnerabilities.

Overall, Figure 7-4. provides a valuable visual representation of our analysis and can be used as a tool for identifying and prioritizing vulnerabilities, as well as for developing effective strategies for mitigating these vulnerabilities.

7.3 Summary

The focus of this research is to assess the viability of utilizing machine learning and natural language processing to predict the CVSS score, determine the interrelationships among vulnerabilities, and generate a vulnerability tree. CVSS is a widely used metric for measuring the severity of security vulnerabilities. The main objective is to evaluate whether the use of machine

learning and NLP for bio-cybersecurity is reliable enough for cybersecurity engineers. Several methods and tools can be used to predict cybersecurity risks and vulnerabilities, depending on the specific context and objectives of the analysis. A common approach is to use machine learning and data analytics techniques to identify patterns and anomalies in network traffic, system logs, and other data sources that may indicate potential security threats. This involves utilizing algorithms such as clustering, KNN, various classifications, and anomaly detection to generate predictive models and alerts to enable security teams to respond more proactively to emerging risks. The results of the research are a significant breakthrough in the field of cybersecurity, as they highlight the potential of BERT as an effective NLP model in predicting various models with a high degree of accuracy. The precision and recall scores ranging from 0.94 to 0.99 reflect the remarkable accuracy of the model. This demonstrates that machine learning and NLP techniques can be highly advantageous tools for cybersecurity professionals to anticipate potential threats accurately. Additionally, the study identified a significant number of different vulnerabilities that could pose a risk to our system, underscoring the importance of having robust cybersecurity measures in place. By using BERT to analyse and predict potential threats, cybersecurity experts can stay ahead of the curve and take proactive measures to prevent potential security breaches. Furthermore, the results of the research were used to generate a sample vulnerability tree, which provides valuable insights into potential vulnerabilities that could arise in our system. This information can be used by cybersecurity professionals to develop targeted strategies to address these vulnerabilities and minimize the risk of a security breach.

Finally, the research highlights the potential of machine learning and NLP techniques in the field of cybersecurity. The accuracy of BERT in predicting different models and identifying potential vulnerabilities demonstrates the immense value of these tools in keeping our systems secure. With continued research and development, these technologies can be used to develop even more effective cybersecurity measures and keep our sensitive information safe.

Chapter 8

Discussion

8.1 Overview

This thesis has explored the potential of leveraging bio-inspired approaches to enhance cybersecurity measures, specifically through the development of a "cyber immune" technology designed to predict and mitigate cyber vulnerabilities. Drawing upon methodologies from natural language processing and machine learning, and analysing Common Vulnerabilities and Exposure reports, this research aimed to bridge the gap between biological security and cyber defence mechanisms.

8.2 Interpretation of Findings

The findings of this study (as showed in Chapter 7) underscore the high degree of accuracy (>94%) achieved in predicting potential cyber vulnerabilities, indicating a promising avenue for the proactive identification and management of cybersecurity threats. This success aligns with the hypothesis that biological principles can inform and enhance cybersecurity strategies. However, the research also revealed challenges in achieving a comprehensive interrelationship mapping between different vulnerabilities, with just over 53% similarity metric. This suggests that while bio-inspired models hold significant promise, their complexity and the dynamic nature of cyber threats pose substantial challenges to their implementation.

8.3 Methodological Evaluation

The choice of NLP and ML was predicated on their proven capability in pattern recognition and

prediction. These techniques in analysing CVE reports represents a novel approach to understanding and predicting cyber vulnerabilities. This methodology's strength lies in its ability to process and analyse vast datasets beyond human capability, offering granular insight into potential vulnerabilities. However, the study faced limitations, particularly in the data's quality and completeness, which could impact the predictive model's accuracy. Furthermore, the rapidly evolving nature of cyber threats means that models must continuously adapt to remain effective.

8.3.1 Practical Implications

For cybersecurity professionals, applying this research could revolutionize vulnerability management, moving from a reactive to a proactive stance. By integrating predictive bio-inspired models into cybersecurity frameworks, organizations can enhance their defencive capabilities, potentially mitigating threats before they come to light. One of the main aims derived from this research includes advocating for the incorporation of bio-inspired cybersecurity solutions in cybersecurity strategies and encouraging the development of standards for the implementation of such technologies.

8.4 Validation and Evaluation of Findings

The suggested bio-inspired cybersecurity framework, particularly through the use of NLP and ML techniques, was conducted to ensure the accuracy and applicability of the research findings. Our model demonstrated superior performance in predicting potential vulnerabilities, achieving an accuracy rate exceeding 94%. When juxtaposed with traditional models, which typically is around an 80% accuracy threshold [187], the enhancement in predictive capabilities was significant. Moreover, the integration of GS and GI methods has set a new benchmark for model explainability within cybersecurity applications. This was particularly evident when assessing the model's ability to not only predict vulnerabilities but also provide actionable insights into their nature and potential mitigation strategies. Furthermore, our findings demonstrate the importance of adopting a more proactive and predictive approach to cybersecurity, one that is inspired by the adaptability and resilience of biological systems. However, the transition to such models will require a significant shift in cybersecurity practices, emphasizing the importance of interdisciplinary collaboration and continuous innovation. The results also contribute to the academic discourse on bio-inspired cybersecurity, providing a solid foundation for future research endeavors in this promising field.

8.5 Stakeholder Influence Map

An integral component of our research involves understanding and illustrating the influence and benefits of our proposed framework and model across a wide range of stakeholders. Here is a list of the primary stakeholders who are directly impacted by or have a specific interest in cybersecurity enhancements through bio-inspired approaches:

1. **Network Administrators:**
    - Benefit: Enhanced predictive capabilities for identifying potential vulnerabilities and threats, leading to improved network security and reduced risk of breaches.
    - Influence: High, Direct; Network administrators can implement the model's insights to develop defences, tailor security measures, and optimize incident response strategies.
2. **Cybersecurity Academics and Researchers:**
    - Benefit**:** proposed novel research framework combining bio-inspired concepts with cybersecurity, offering a new avenue for interdisciplinary studies and innovation in security technologies.
    - Influence**:** Medium to High, Direct; Academics can leverage the model to expand the body of knowledge in cybersecurity, develop curriculum, and inspire future research directions**.**
3. **CEOs and Business Leaders:**
    - Benefit: Improved organizational security posture and resilience against cyber threats, contributing to safeguarding company assets and reputation.
    - Influence: Medium, Indirect; While not directly involved in technical implementation, business leaders can use the framework's insights to inform strategic decisions and cybersecurity investments.
4. **Cybersecurity Policy Makers:**
    - Benefit: Insights into emerging vulnerabilities and threat patterns, supporting the development of more effective cybersecurity policies and standards.
    - Influence: Medium, Indirect; Policy makers can utilize the research findings to guide legislative and regulatory efforts, enhancing national and international cybersecurity frameworks

5. **IT Security Teams:**
    - Benefit: Access to a cutting-edge analytical toolset for vulnerability assessment, enabling more proactive and data-driven security management.
    - Influence: High, Direct; IT security teams can directly apply the model and framework to strengthen organizational security infrastructure and response mechanisms.
6. **General Public and End-Users:**
    - Benefit: Increased security of personal data and reduced risk of identity theft and other cyber-crimes, as organizations adopt more robust security measures informed by the research.
    - Influence: Low to Medium, Indirect; The general public benefits from the trickle-down effects of improved cybersecurity practices across industries and sectors.

8.6 Summary

These study findings highlighted the significant contributions of both the theoretical and practical realms of cybersecurity. The bio-inspired model not only enhances predictive accuracy and explainability but also paves the way for a new era of cybersecurity solutions that are dynamic, adaptable, and inherently resilient. Future research can build on this foundation, exploring new possibilities in the application of biological principles to cybersecurity challenges.

Chapter 9

Conclusions and Future Work

9.1 Overview

Chapter 6 serves as the concluding chapter of this thesis. The chapter commences by outlining the challenges and limitations that were faced during the research period. The second half of the chapter presents potential areas for future development, which were either beyond the scope of the study or discovered during the validation experiments. The chapter culminates by highlighting the author's methodology's scientific advancements, as well as providing final comments and conclusions. This study discussed the importance of addressing cyber-attacks, which have become increasingly severe and frequent in recent years. To improve cybersecurity, researchers have turned to the human immune system as a model, looking for ways to detect and deter cyber threats. A "cyber immune" technology is proposed as a means to detect and defend against unknown cyber-attacks. The methodology we have used for predicting possible vulnerabilities is presented through the use of NLP and analysing CVE reports. By examining the field of human biology, we aim to gain valuable insights into bio-cybersecurity, which can be applied to the development of more effective cybersecurity measures.

9.2 Conclusion

The application of biological and living organism-inspired solutions in the field of cybersecurity has shown promise in enhancing the effectiveness and resilience of information security systems. By examining the natural mechanisms of living organisms such as the immune system,

cybersecurity experts can gain insights into adaptive and dynamic approaches to threat detection and mitigation. This can lead to the development of more robust and sophisticated security systems that can adapt and evolve to meet the constantly evolving threat landscape. However, there are still limitations and challenges in the implementation of bio-inspired cybersecurity solutions, such as the complexity of mimicking natural systems and the need for specialized expertise in both cybersecurity and biology. Despite these challenges, the potential benefits and innovative possibilities of this approach make it a promising area for continued research and development.

Overall, the integration of biology and living organisms-inspired concepts and techniques in cybersecurity can provide a valuable paradigm shift in the development of information security systems that can withstand the challenges of the modern digital environment. It is an exciting area of exploration and discovery that has the potential to contribute to the continued evolution of cybersecurity.

9.2.1 Comparison with Existing Models

Our methodology's predictive accuracy exceeds 94%, significantly outperforming traditional cybersecurity models, typically achieve around 80 accuracy rates in vulnerability prediction [187]. Our results indicate a significant improvement in predicting cybersecurity vulnerability by adapting vulnerability interrelationship policy. Table 9.1. shows a comparison of existing works and our approaches.

Table 9.1 Comparison of experimental results between the existing work and our model.

References	Methodology	Accuracy
[188]	ML	66.25%
[189]	Deep Learning and NLP	72% to 79%
[190]	NLP	91.12%
[191]	ML	83.13%
[192]	Deep Learning	60%
Proposed Work	NLP	94%

This advancement underscores the potential of integrating bio-inspired concepts with cutting-edge computational techniques to redefine cybersecurity practices.

9.2.2 Bio-inspired Cybersecurity Framework

Drawing in comparison to the existing works such as work of Mthunzi et al [164], who explored the efficacy of bio-inspired algorithms in cybersecurity, our framework further validates the potential of nature-inspired solutions in cybersecurity. However, our approach extends beyond algorithmic inspiration, combining system-level insights from biological protection mechanisms and vulnerability interrelationships, thereby offering a more holistic and adaptable cybersecurity strategy.

In this study, we showed that using machine learning and NLP for training data, such as the ability which the human immune system has, has a significant potential for improving cybersecurity operations and can be effectively utilized to predict the severity of vulnerabilities efficiently and reliably.

9.3 Problems and Constraints

One limitation was that the use of the human immune system as a model for cybersecurity is still in the early stages of development. It was unclear how well the principles of the immune system can be applied to the cyber environment and whether this approach will lead to more effective cybersecurity measures.

Another constraint was the reliance on technology and algorithms to detect and respond to cyber threats. While machine learning algorithms and natural language processing can be powerful tools, they were not foolproof and could be vulnerable to attacks themselves.

Additionally, there may be ethical and privacy concerns related to the use of cyber immune technology to detect and monitor potential threats. It is important to consider how such technology will be used and what measures will be put in place to protect individual privacy and prevent abuses.

Finally, the use of Common Vulnerabilities and Exposures reports as the basis for predicting possible vulnerabilities was limited by the quality and availability of such reports. The accuracy and completeness of these reports can impact the effectiveness of the proposed methodology.

9.4 Dissemination and Exploitation Plan

- **Dissemination Plan:**
    1. Publish research findings in academic journals, particularly those related to

cybersecurity, biological inspired computing, and artificial intelligence.

2. Present research at relevant conferences and seminars, including those focused on cybersecurity, AI, and biology.
3. Engage with industry partners, particularly those in the technology and cybersecurity sectors, to share research findings and potential applications.
4. Create a website and social media accounts to communicate research findings and updates to a wider audience, including the general public, students, and educators.

- **Exploitation Plan:**
    1. Develop a prototype system that demonstrates the proposed solution for securing the information environment inspired by biology.
    2. Collaborate with industry partners to develop and commercialize the solution, potentially through licensing agreements or joint ventures.
    3. Develop training materials and workshops to educate cybersecurity professionals on the use and implementation of bio-cybersecurity.
    4. Work with policymakers and regulatory bodies to promote the adoption of bio-cybersecurity measures and provide guidance on their implementation.

The dissemination and exploitation plan of this study is to show the benefits of the proposed work and technique for securing the information environment inspired by biology, while also ensuring that the ethical implications and potential risks of such an approach are considered and addressed.

9.5 Future Development

Future work for this study involves expanding the research on bio-inspired cybersecurity to other areas beyond the immune system, such as the learning (training system). Additionally, further studies could be conducted to determine the effectiveness of implementing bio-inspired cybersecurity measures in real-world scenarios, and to compare their performance to traditional cybersecurity measures. Finally, research could be done to investigate how bio-inspired cybersecurity can be used to address emerging threats such as those posed by the Internet of Things (IoT) and cloud computing.

1. Further exploration of bio-inspired approaches to cybersecurity, including investigating other biological systems that could provide inspiration for developing secure information

systems.

2. Testing and validating the proposed solution in real-world settings to determine its effectiveness in securing information environments.
3. Developing a comprehensive framework for evaluating the security of information systems that incorporates bio-inspired approaches, as well as traditional cybersecurity techniques.
4. Collaborating with experts in biology, bioengineering, and related fields to explore the potential for developing new technologies or materials that could improve the security of information systems based on biological principles.
5. Proposing a new model that is able to find vulnerability interrelationships automatically.

Reference

[1] Jowitt, T. 2014. White House Advisory Group: Governments Have Five Years To Secure IoT. TechWeek Europe, http://www.techweekeurope.co.uk/e-regulation/governmentssecure-iot-156149.

[2] McClimans, F., Fersht, P., Snowdon, J., (2016), "The State of Cybersecurity and Digital Trust, 2016", HfS Research &Accenture, Ltd [3]. (McClimans, F., et al. "The state of Cybersecurity and digital trust 2016." HfS Research &Accenture, Ltd (2016).)

[3] Information Security Breaches, GCHQ (2014), www.gov.uk/government/publications/information-securitybreaches-survey-2014.

[4] Bou-Harb, E.; Fachkha, C.; Pourzandi, M.; Debbabi, M.; Assi, C. Communication security for smart grid distribution networks. IEEE Commun. Mag. 2013, 51, 42–49.

[5] National Institute of Standards and Technology, The Smart Grid Interoperability Panel, Cyber Security Working Group. Guidelines for smart grid cyber security. Available online: http://www.nist.gov/smartgrid/upload/nistir-7628_total.pdf (accessed on 10 January 2015).

[6] Ismail, Z.; Leneutre, J.; Bateman, D.; Chen, L. A game theoretical analysis of data confidentiality attacks on smart-grid AMI. IEEE J. Sel. Areas Commun. 2014, 32, 1486–1499.

[7] B. Kordy, S. Mauw, S. Radomirovic, P. Schweitzer, Foundations of attack–defence trees, LNCS. Springer, Heidelberg, vol. 6561, pp. 80-95, 2011.

[8] A. Roy, D. Kim and K.S. Trivedi, Cyber security analysis using attack countermeasure trees, Proc. of Cyber Security and Information Intelligence Research Workshop (CSIIRW2010), ACM, textit Oak Ridge, TN, USA, 2010.

[9] Akerkar, R.A.; Sajja, P.S. KnowledgeBased Systems; Jones & Bartlett Publishers: Toronto, ON, Canada, 2010

[10] Batista, L.O.; de Silva, G.A.; Araujo, V.S.; Araujo, V.J.S.; Rezende, T.S.; Guimaraes, A.J.; Souza, P.V.D.C. Fuzzy neural networks to create an expert system for detecting attacks by sql injection. Int. J. Forensic Comput. Sci. 2018, 1, 8–21

[11] Dheir, I.; Abu-Naser, S.S. Knowledge Based System for Diagnosing Guava Problems. Int. J. Acad. Inf. Syst. Res. (IJAISR) 2019, 3, 9–15.

[12] Atymtayeva, L.; Kozhakhmet, K.; Bortsova, G. Building a Knowledge Base for Expert System in Information Security. In Soft Computing in Artificial Intelligence; Advances in Intelligent Systems and Computing Series.; Springer: Cham, Switzerland, 2014; pp. 57–76.

[13] Pinto, F.J. Application of the Bayesian Model in Expert Systems. In Proceedings of the International Symposium on Distributed Computing and Artificial Intelligence, Ávila, Spain, 26–28 June 2019; Springer: Cham, Switzerland, 2019; pp. 117–124.

[14] Ingoldsby, T.R. Attack Tree-Based Threat Risk Analysis; Amenaza Technologies Limited: Calgary, AB, Canada, 2010; pp. 3–9.

[15] Lohner, B. Attack-Defence-Trees and other Security Modeling Tools. Network 2018, 97, 97–103.

[16] Huistra, D.J. Automated gEneration of Attack Trees by Unfolding Graph Transformation Systems. Master's Thesis, University of Twente, Enschede, The Netherlands, 2016.

[17] Gadyatskaya, O.; Jhawar, R.; Kordy, P.; Lounis, K.; Mauw, S.; Trujillo-Rasua, R. Attack trees for practical security assessment: Ranking of attack scenarios with ADTool 2.0. In Proceedings of the International Conference on Quantitative Evaluation of Systems, Glasgow, UK, 10–12 September 2016; pp. 159–162.

[18] Vigo, R.; Nielson, F.; Nielson, H.R. Automated generation of attack trees. In Proceedings of the 2014 IEEE 27th Computer Security Foundations Symposium, Vienna, Austria, 19–22 July 2014; pp. 337–350.

[19] Tentilucci, M.; Roberts, N.; Kandari, S.; Johnson, D.; Bogaard, D.; Stackpole, B.; Markowsky, G. Crowdsourcing Computer Security Attack Trees. In Proceedings of the 10th Annual Symposium on Information Assurance (ASIA'15), Albany, NY, USA, 2–3 June 2015; pp. 19–23.

[20] Attack Tree Translator and Analyser ATTop. University of Twente—Formal Methods and Tools. Available online: https://github.com/utwente-fmt/attop (accessed on 18 March 2020).

[21] Abdo, H.; Kaouk, M.; Flaus, J.M.; Masse, F. A safety/security risk analysis approach of Industrial Control Systems: A cyber bowtie—Combining new version of attack tree with bowtie analysis. Comput. Secur. 2018, 72, 175–195.

[22] Jhawar, R.; Kordy, B.; Mauw, S.; Radomirovi´c, S.; Trujillo-Rasua, R. Attack trees with sequential conjunction. JIFIP Adv. Inf. Commun. Technol. 2015, 455, 339–353.

[23] Kumar, R.; Schivo, S.; Ruijters, E.; Yildiz, B.M.; Huistra, D.; Brandt, J.; Stoelinga, M. Effective Analysis of Attack Trees: A Model-Driven Approach. In Fundamental Approaches to Software Engineering; Russo, A., Schurr, A., Eds.; Springer: Cham, Switzerland, 2018; pp. 56–73.

[24] Bogaard, D.; Johnson, D. Producing and Evaluating Crowdsourced Computer Security Attack Trees. In Proceedings of the 2016 IEEE Symposium on Technologies for Homeland Security (HST), Waltham, MA, USA, 10–11 May 2016; pp. 1–4.

[25] Ahmed, R., 2016. Security metrics and the risks: an overview. IJCTT, Volume 41(4), pp. 106-112.

[26] Ahsan, M. G. R. a. D. A., 2018. Smote implementation on phishing data to enhance cybersecurity.. s.l., IEEE., pp. 0531-0536 .

[27] Bhopi, S. a. D. N., 2015. Study of dynamic defence technique to overcome drawbacks of moving target defence.. s.l., IEEE, pp. 637-641.

[28] Bitam, S. Z. S. a. M. A., 2016. Bio-inspired cybersecurity for wireless sensor networks. IEEE Communications Magazine, Volume 54(6), pp. 68-74.

[29] Buczak, A. a. G. E., 2015. A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications surveys & tutorials, Volume 18(2), pp. 1153-1176.

[30] Carley, K. C. G. A. N. a. L. H., 2018. Social cyber-security. s.l., s.n., pp. 389-394.

[31] Crowther, G., 2017. The cyber domain. The Cyber Defence Review, Volume 2(3), pp. 63-78.

[32] Daras, N. a. R. M. e., 2015. Computation, cryptography, and network security. s.l.:Springer.

[33] Du, J. & Chao, S., 2010. A study of information security for M2M of IOT.. s.l., IEEE, pp. V3-576.

[34] Fink, G. H. J. M. A. a. F. E., 2014. Defence on the move: ant-based cyber defence. ,. IEEE Security & Privacy, Volume 12(2), , pp. 36-43.

[35] Fink, G. O. C. H. J. M. A. F. E. a. C. M., 2011. Bio-Inspired Enterprise Security. s.l., IEEE, pp. 212-213.

[36] Forouzan, B. a. M. D., 2015. Cryptography and network security. s.l.:Mc Graw Hill Education (India) Private Limited..

[37] Guthikonda, A. A.-S. E. F. A. a. R. M., 2017. Bio-inspired innovations in cyber security. s.l., IEEE, pp. 105-109.

[38] Hariri, S. T. C. S. P. A.-M. F. a. B. E., 2015. Dddas-based resilient cyber battle management services (d-rcbms). s.l., s.n., pp. 65-65.

[39] Jhawar, R. K. B. M. S. R. S. a. T.-R. R., 2015. Attack trees with sequential conjunction.. s.l., Springer, Cham., pp. 339-353.

[40] Korczynski, M. H. A. H. J. H. H. R. S. a. F. N., 2016. Hive oversight for network intrusion early warning using DIAMoND: a bee-inspired method for fully distributed cyber defence.. IEEE Communications Magazine, Volume 54(6), pp. 60-67.

[41] Kordy, B. M. S. R. S. a. S. P., 2014. Attack–defence trees. Journal of Logic and Computation, Volume 24(1), pp. 55-87.

[42] Kordy, B. P.-C. L. a. S. P., 2014. DAG-based attack and defence modeling: Don't miss the forest for the attack trees. Computer science review, Volume 13, pp. 1-38.

[43] Kumar Kar, A., 2016. Bio inspired computing – A review of algorithms and scope of applications,. Expert Systems with Applications,, 59(0957-4174,), pp. 20-32.

[44] Kumar, J. a. P. D., n.d. A survey on internet of things. Security and privacy issues. International Journal of Computer Applications, Volume 90(11), p. 2014.

[45] Kumar, S., 2015. Review on network security and cryptography. International Transaction of Electrical and Computer Engineers System, Volume 3(1), pp. 1-11.

[46] Lin, H., 2019. The existential threat from cyber-enabled information warfare.. Bulletin of the Atomic Scientists, Volume 75(4), pp. 187-196.

[47] Liu, Q. a. F. M., 2021. Bio-inspired photonics–marine hatchetfish camouflage strategies for RF steganography. Optics Express, Volume 29(2), pp. 2587-2596.

[48] Nicholson, L. B., 2016. The immune system. Essays in biochemistry, Volume 60(3), pp. 275-301.

[49] Okamoto, T. T. M., 2016. Toward an artificial immune server against cyber attacks. Artif. Life Robot. , Volume 21(3), p. 351–356.

[50] Parham, P., 2014. The Immune System. New York: Garland Science.

[51] Parn, E. a. E. D., 2019. Cyber threats confronting the digital built environment. Engineering. Construction and Architectural Management.

[52] Rauf, U., 2018. A taxonomy of bio-inspired cyber security approaches: existing techniques and future directions. Arabian Journal for Science and Engineering, Volume 43(12), pp. 6693-6708.

[53] Sinha, S. a. T. S., 2014. Study on agents based meta-heuristic approach for cyber security defence mechanism.. IITM J Manag IT, Volume 5(1), pp. 63-66.

[54] Sneps-Sneppe, M. S. V. a. N. D., 2018. On cyber-security of information systems. s.l., Springer, pp. 201-211.

[55] Sommestad, T. E. M. a. H. H., 2012. The cyber security modeling language: A tool for assessing the vulnerability of enterprise system architectures. IEEE Systems Journal, Volume 7(3), pp. 363-373.

[56] Tarao, M. a. O. T., 2016. Toward an artificial immune server against cyber attacks: enhancement of protection against DoS attacks. Procedia Computer Science, Volume 96, pp. 1137-1146.

[57] Vidalis, S. a. J. A., 2003. Using vulnerability trees for decision making in threat assessment, s.l.: University of Glamorgan, School of Computing, Tech. Rep CS-03-2..

[58] Wheelus, C. B.-H. E. a. Z. X., 2018. Tackling class imbalance in cyber security datasets.. s.l., IEEE, pp. 229-232.

[59] Wlodarczak, P., 2017. Cyber Immunity-A Bio-Inspired Cyber Defence System. Bioinformatics and Biomedical Engineering, Volume 2, pp. 199-208.

[60] Zhang, Z. N. M. W. M. a. W. Z., 2016. Bio-inspired RF steganography via linear chirp radar signals.. IEEE Communications Magazine, Volume 54(6), pp. 82-86.

[61] Zhang, Z. Q. Y. W. Z. N. M. E. J. a. W. M., 2017. RF steganography via LFM chirp radar signals.. IEEE Transactions on Aerospace and Electronic Systems, Volume 54(3), pp. 1221-1236.

[62] Diane R. Murphy and Richard H. Murphy. 2013. Teaching Cybersecurity: Protecting the Business Environment. In Proceedings of the 2013 on InfoSecCD '13: Information Security Curriculum Development Conference (InfoSecCD '13). Association for Computing Machinery, New York, NY, USA, 88–93.

[63] R. Sabillon, J. Serra-Ruiz, V. Cavaller and J. Cano, "A Comprehensive Cybersecurity Audit Model to Improve Cybersecurity Assurance: The CyberSecurity Audit Model (CSAM)," 2017 International Conference on Information Systems and Computer Science (INCISCOS), Quito, 2017, pp. 253-259, doi: 10.1109/INCISCOS.2017.20.

[64] Eugen, P. and Petruţ, D., 2018. Exploring the new era of cybersecurity governance. Ovidius University Annals, Economic Sciences Series, 18(1), pp.358-363.

[65] Slupska, J., 2020. War, Health and Ecosystem: Generative Metaphors in Cybersecurity Governance. Philosophy & Technology, pp.1-20

[66] Bryson, J.M., 2014. What to do when stakeholders matter: stakeholder identification and analysis techniques. Public management review, 6(1), pp.21-53.

[67] Wang, W., Liu, W. and Mingers, J., 2015. A systemic method for organisational stakeholder identification and analysis using Soft Systems Methodology (SSM). European Journal of Operational Research, 246(2), pp.562-574

[68] Murphy, D. R., & Murphy, R. H. (2013, October). Teaching cybersecurity: Protecting the business environment. In Proceedings of the 2013 on InfoSecCD'13: Information Security Curriculum Development Conference (pp. 88-93).

[69] Hadlington, L. (2017). Human factors in cybersecurity; examining the link between Internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviours. Heliyon, 3(7), e00346.

[70] Corallo, A., Lazoi, M., & Lezzi, M. (2020). Cybersecurity in the context of industry 4.0: A structured classification of critical assets and business impacts. Computers in industry, 114, 103165.

[71] Mosteanu, N. R. (2020). Artificial Intelligence and Cyber Security–A Shield against Cyberattack as a Risk Business Management Tool–Case of European Countries. Quality-Access to Success, 21(175).

[72] Chertoff, M., & Simon, T. (2015). The impact of the dark web on internet governance and cyber security.

[73] Shackelford, S. J., & Craig, A. N. (2014). Beyond the new digital divide: Analysing the evolving role of national governments in internet governance and enhancing cybersecurity. Stan. J. Int'l L., 50, 119.

[74] Fritz, M. M., Rauter, R., Baumgartner, R. J., & Dentchev, N. (2018). A supply chain perspective of stakeholder identification as a tool for responsible policy and decision-making. Environmental Science & Policy, 81, 63-76.

[75] Wang, W., Liu, W., & Mingers, J. (2015). A systemic method for organisational stakeholder identification and analysis using Soft Systems Methodology (SSM). European Journal of Operational Research, 246(2), 562-574.

[76] Babar, M. I., Ghazali, M., Jawawi, D. N., & Zaheer, K. B. (2015). StakeMeter: Value-Based stakeholder identification and quantification framework for value-based software systems. PloS one, 10(3), e0121344

[77] Zhou, Y., Liu, S., Siow, J., Du, X., & Liu, Y. (2019). Devign: Effective vulnerability identification by learning comprehensive program semantics via graph neural networks. arXiv preprint arXiv:1909.03496.

[78] Lee, G., Jun, K. S., & Chung, E. S. (2015). Group decision-making approach for flood vulnerability identification using the fuzzy VIKOR method. Natural Hazards and Earth System Sciences, 15(4), 863-874.

[79] Wijayasekara, D., Manic, M., & McQueen, M. (2014, October). Vulnerability identification and classification via text mining bug databases. In IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society (pp. 3612-3618). IEEE.

[80] Zhou, Y., Liu, S., Siow, J., Du, X., & Liu, Y. (2019). Devign: Effective vulnerability identification by learning comprehensive program semantics via graph neural networks. arXiv preprint arXiv:1909.03496.

[81] Kadivar, Mehdi. "Cyber-attack attributes." Technology Innovation Management Review 4.11 (2014).

[82] Bozorgi, Mehran, et al. “Beyond Heuristics.” Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '10, 2010, doi:10.1145/1835804.1835821.

[83] Guthikonda A, Al-Shaer E, Farooq A, Raja MY. Bio-inspired innovations in cyber security. In2017 14th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT) 2017 Oct 9 (pp. 105-109). IEEE.

[84] Spiering MJ. Primer on the immune system. Alcohol research: current reviews. 2015;37(2):171.

[85] Khazaei, Atefeh, et al. “An Automatic Method for CVSS Score Prediction Using Vulnerabilities Description.” Journal of Intelligent & Fuzzy Systems, vol. 30, no. 1, 2015, pp. 89–96., doi:10.3233/ifs-151733.

[86] Farmer JD, Packard NH, Perelson AS. The immune system, adaptation, and machine learning. Physica D: Nonlinear Phenomena. 1986 Oct 1;22(1-3):187-204.

[87] Elbaz, Clément, et al. “Fighting N-Day Vulnerabilities with Automated CVSS Vector Prediction at Disclosure.”Proceedings of the 15th International Conference on Availability, Reliability and Security, 2020, doi:10.1145/3407023.3407038.

[88] Tarao M, Okamoto T. Toward an artificial immune server against cyber attacks: enhancement of protection against DoS attacks. Procedia Computer Science. 2016 Jan 1;96:1137-46.

[89] Yin, Jiao, et al. “Apply Transfer Learning to Cybersecurity: Predicting Exploitability of Vulnerabilities by Description.” Knowledge-Based Systems, vol. 210, 2020, p. 106529., doi:10.1016/j.knosys.2020.106529.

[90] Igbe, Obinna. "Artificial immune system based approach to cyber attack detection." PhD diss., The City College of New York, 2019.

[91] Wu, Zhengxuan, and Desmond C. Ong. “On Explaining Your Explanations of BERT: An Empirical Study with Sequence Classification.” 2021, doi:arXiv:2101.00196.

[92] Urbanska M, Ray I, Howe AE, Roberts M. Structuring a vulnerability description for comprehensive single system security analysis. Rocky Mountain Celebration of Women in Computing, Fort Collins, CO, USA. 2012 Nov.

[93] Li, Jiwei, et al. “Visualizing and Understanding Neural Models in NLP.” Proceedings of the 2016 Conference othe North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016, doi:10.18653/v1/n16-1082.

[94] Rauf U. A taxonomy of bio-inspired cyber security approaches: existing techniques and future directions. Arabian Journal for Science and Engineering. 2018 Dec;43(12):6693-708.

[95] Kindermans, Pieter-Jan, et al. . “The (Un)Reliability of Saliency Methods.” 2017, doi:arXiv1711.00867.

[96] Abraham C, Chatterjee D, Sims RR. Muddling through cybersecurity: Insights from the US healthcare industry. Business horizons. 2019 Jul 1;62(4):539-48.

[97] Interesting Engineering, “Cyber Attacks More Likely to Bring Down F-35 Jets Than Missile”, Fabienne Lang, Feb 25, 2021

[98] Rodriguez C, Zamanirad S, Nouri R, Darabal K, Benatallah B, Al-Banna M. Security vulnerability information service with natural language query support. InInternational Conference on Advanced Information Systems Engineering 2019 Jun 3 (pp. 497-512). Springer, Cham.

[99] PRNewsWires, “(ISC)2 Survey Finds Cybersecurity Professionals Have Increasing Level of Concern About SolarWinds Incident”, Mar 29, 2021

[100] Fieber, T.J., 2015. The Iranian computer network operations threat to US critical infrastructures (Doctoral dissertation, Utica College).

[101] Durkota, K., Lisý, V., Kiekintveld, C., Bošanský, B. and Pěchouček, M., 2016. Case studies of network defence with attack graph games. IEEE Intelligent Systems, 31(5), pp.24-30.

[102] Hu, J., Shen, L., Albanie, S., Sun, G. and Vedaldi, A., 2018. Gather-excite: Exploiting feature context in convolutional neural networks. Advances in neural information processing systems, 31.

[103] Robinson, M., Jones, K. and Janicke, H., 2015. Cyber warfare: Issues and challenges. Computers & security, 49, pp.70-94.

[104] Bellaby, R.W., 2016. Justifying cyber-intelligence?. Journal of Military Ethics, 15(4), pp.299-319.

[105] Whyte, C., 2016. Ending cyber coercion: Computer network attack, exploitation and the case of North Korea. Comparative strategy, 35(2), pp.93-102.

[106] Villalón-Huerta, A., Ripoll-Ripoll, I. and Marco-Gisbert, H., 2021. CNA tactics and techniques: a structure proposal. Journal of Sensor and Actuator Networks, 10(1), p.14.

[107] Lugo, A.D.P., 2020. Computer Network Attack and Exploitation in Criminal Investigations (Doctoral dissertation, Utica College).

[108] Bennett, K.B., Bryant, A. and Sushereba, C., 2018. Ecological interface design for computer network defence. Human factors, 60(5), pp.610-625.

[109] Urias, V.E., Stout, W.M. and Loverro, C., 2015, September. Computer network deception as a moving target defence. In 2015 International Carnahan Conference on Security Technology (ICCST) (pp. 1-6). IEEE.

[110] Villalón-Huerta, A., Ripoll-Ripoll, I. and Marco-Gisbert, H., 2021. CNA tactics and techniques: a structure proposal. Journal of Sensor and Actuator Networks, 10(1), p.14.

[111] Lugo, A.D.P., 2020. Computer Network Attack and Exploitation in Criminal Investigations (Doctoral dissertation, Utica College).

[112] Ling, C., 2017. Information Asset Management for Cyber. The cyber risk handbook: creating and measuring effective cybersecurity capabilities, pp.281-288.

[113] Dekker, S.W., 2015. The danger of losing situation awareness. Cognition, Technology & Work, 17, pp.159-161.

[114] Stanton, N.A., Salmon, P.M., Walker, G.H., Salas, E. and Hancock, P.A., 2017. State-of-science: situation awareness in individuals, teams and systems. Ergonomics, 60(4), pp.449-466.

[115] Lees, D.W., 2016. Understanding effects of Operations Security (OPSEC) awareness levels of military spouses through the lenses of training and program management: A qualitative study (Doctoral dissertation, Creighton University).

[116] Theohary, C.A., 2020, December. Defence primer: Information operations. LIBRARY OF CONGRESS WASHINGTON DC.

[117] van Hardeveld, G.J., Webber, C. and O'Hara, K., 2018. Expert perspectives on the evolution of carders, cryptomarkets and operational security. In 10th ACM Conference on Web Science, Amsterdam, the Netherlands (pp. 6-10).

[118] Wanjohi, D.M., 2019. Information Security Research Project.

[119] Sinno, A.M., 2017. Information security awareness in Lebanese community.(c2017) (Doctoral dissertation, Lebanese American University).

[120] Mbabazi, S., 2018. Examining information security controls in the human resource department of Kampala International University, Uganda from 2013-2018 (Doctoral dissertation, Kampala International University).

[121] Etesami, S.R. and Başar, T., 2019. Dynamic games in cyber-physical security: An overview. Dynamic Games and Applications, 9(4), pp.884-913.

[122] Tan, J., Yang, J., Wu, S., Chen, G. and Zhao, J., 2021. A critical look at the current train/test split in machine learning. arXiv preprint arXiv:2106.04525.

[123] Singh, V., Pencina, M., Einstein, A.J., Liang, J.X., Berman, D.S. and Slomka, P., 2021. Impact of train/test sample regimen on performance estimate stability of machine learning in cardiovascular imaging. Scientific Reports, 11(1), p.14490.

[124] Feutrill, A., Ranathunga, D., Yarom, Y. and Roughan, M., 2018, November. The effect of common vulnerability scoring system metrics on vulnerability exploit delay. In 2018 Sixth International Symposium on Computing and Networking (CANDAR) (pp. 1-10). IEEE.

[125] Holm, H. and Afridi, K.K., 2015. An expert-based investigation of the common vulnerability scoring system. Computers & Security, 53, pp.18-30.

[126] Krumay, B., Bernroider, E.W. and Walser, R., 2018. Evaluation of cybersecurity management controls and metrics of critical infrastructures: A literature review considering the NIST cybersecurity framework. In Secure IT Systems: 23rd Nordic Conference, NordSec 2018, Oslo, Norway, November 28-30, 2018, Proceedings 23 (pp. 369-384). Springer International Publishing.

[127] Al-Shaer, R., Spring, J.M. and Christou, E., 2020, June. Learning the associations of mitre att & ck adversarial techniques. In 2020 IEEE Conference on Communications and Network Security (CNS) (pp. 1-9). IEEE.

[128] Wang, J., Neil, M. and Fenton, N., 2020. A Bayesian network approach for cybersecurity risk assessment implementing and extending the FAIR model. Computers & Security, 89, p.101659.

[129] Groš, S., 2021, June. A critical view on CIS controls. In 2021 16th International Conference on Telecommunications (ConTEL) (pp. 122-128). IEEE.

[130] Donaldson, S.E., Siegel, S.G., Williams, C.K., Aslam, A., Donaldson, S.E., Siegel, S.G., Williams, C.K. and Aslam, A., 2015. Cybersecurity frameworks. Enterprise Cybersecurity: How to Build a Successful Cyberdefence Program Against Advanced Threats, pp.297-309.

[131] Sayres, R., Taly, A., Rahimy, E., Blumer, K., Coz, D., Hammel, N., Krause, J., Narayanaswamy, A., Rastegar, Z., Wu, D. and Xu, S., 2019. Using a Machine learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy. Ophthalmology, 126(4), pp.552-564.

[132] Alaparthi, S. and Mishra, M., 2020. Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey. arXiv preprint arXiv:2007.01127.

[133] Dodiya, B., Singh, U.K. and Gupta, V., 2021. Trend Analysis of the CVE Classes Across CVSS Metrics. International Journal of Computer Applications, 975, p.8887.

[134] Younis, A., Malaiya, Y.K. and Ray, I., 2016. Evaluating CVSS base score using vulnerability rewards programs. In ICT Systems Security and Privacy Protection: 31st IFIP TC 11 International Conference, SEC 2016, Ghent, Belgium, May 30-June 1, 2016, Proceedings 31 (pp. 62-75). Springer International Publishing.

[135] Osaci, M., Cristea, A.D., Ghiuzan, D. and Berdie, D.A., 2018. SAP authorization based on the four eyes principle. Annals of the Faculty of Engineering Hunedoara, 16(2), pp.43-46.

[136] Abu Alfeilat, H.A., Hassanat, A.B., Lasassmeh, O., Tarawneh, A.S., Alhasanat, M.B., Eyal Salman, H.S. and Prasath, V.S., 2019. Effects of distance measure choice on k-nearest neighbor classifier performance: a review. Big data, 7(4), pp.221-248.

[137] Senders, J.T., Staples, P.C., Karhade, A.V., Zaki, M.M., Gormley, W.B., Broekman, M.L., Smith, T.R. and Arnaout, O., 2018. Machine learning and neurosurgical outcome prediction: a systematic review. World neurosurgery, 109, pp.476-486.

[138] Schepens, J., Van der Slik, F. and Van Hout, R., 2016. The L2 impact on acquiring Dutch as a L3: The L2 distance effect. Mixed effects regression models in linguistics. Berlin, Germany: Springer.

[139] Hardeniya, N., Perkins, J., Chopra, D., Joshi, N. and Mathur, I., 2016. Natural language processing: python and NLTK. Packt Publishing Ltd.

[140] Hassan, M., Mylonas, A. and Vidalis, S., 2016. WHEN BIOLOGY MEETS CYBER-SECURITY. Journal of Information System Security, 12(3).

[141] Worldwide Data Created from 2010 to 2025, Statista, https://www.statista.com/statistics/871513/worldwide-data-created/, 2023.

[142] Markowsky, L. and Markowsky, G., 2015, September. Scanning for vulnerable devices in the Internet of Things. In 2015 IEEE 8th International conference on intelligent data acquisition and advanced computing systems: technology and applications (IDAACS) (Vol. 1, pp. 463-467). IEEE.

[143] Bhuyan, M.H., Bhattacharyya, D.K. and Kalita, J.K., 2011. Surveying port scans and their detection methodologies. The Computer Journal, 54(10), pp.1565-1581.

[144] Kaushik, K., Punhani, I., Sharma, S. and Martolia, M., 2022, December. An Advanced Approach for performing Cyber Fraud using Banner Grabbing. In 2022 5th International Conference on Contemporary Computing and Informatics (IC3I) (pp. 298-302). IEEE.

[145] Kondo, T.S. and Mselle, L.J., 2014. Penetration testing with banner grabbers and packet sniffers. Journal of Emerging Trends in computing and information sciences, 5(4), pp.321-327.

[146] Tennina, S., Gaddour, O., Koubâa, A., Royo, F., Alves, M. and Abid, M., 2016. Z-Monitor: A protocol analyser for IEEE 802.15. 4-based low-power wireless networks. Computer Networks, 95, pp.77-96.

[147] Talasila, P., Kakrambe, M., Rai, A., Santy, S., Goveas, N. and Deshpande, B.M., 2018, January. BITS Darshini: A Modular, Concurrent Protocol Analyser Workbench. In Proceedings of the 19th International Conference on Distributed Computing and Networking (pp. 1-10).

[148] Vinzenz, N. and Oka, D.K., 2021. Integrating fuzz testing into the cybersecurity validation strategy (No. 2021-01-0139). SAE Technical Paper.

[149] Foreman, P., 2019. Vulnerability management. CRC Press.

[150] Franco, J., Aris, A., Canberk, B. and Uluagac, A.S., 2021. A survey of honeypots and honeynets for internet of things, industrial internet of things, and cyber-physical systems. IEEE Communications Surveys & Tutorials, 23(4), pp.2351-2383.

[151] Sokol, P., Míšek, J. and Husák, M., 2017. Honeypots and honeynets: issues of privacy. EURASIP Journal on Information Security, 2017, pp.1-9.

[152] Khan, R., Maynard, P., McLaughlin, K., Laverty, D. and Sezer, S., 2016, October. Threat analysis of blackenergy malware for synchrophasor based real-time control and monitoring in smart grid. In 4th International Symposium for ICS & SCADA Cyber Security Research 2016 4 (pp. 53-63).

[153] Rani, S.K., Soundarya, B.C., Gururaj, H.L. and Janhavi, V., 2021, October. Comprehensive Analysis of Various Cyber Attacks. In 2021 IEEE Mysore Sub Section International Conference (MysuruCon) (pp. 255-262). IEEE.

[154] Subrahmanian, V.S., Ovelgonne, M., Dumitras, T. and Prakash, B.A., 2015. The global cyber-vulnerability report.

[155] Harzevili, N.S., Shin, J., Wang, J. and Wang, S., 2022. Characterizing and Understanding Software Security Vulnerabilities in Machine Learning Libraries. arXiv preprint arXiv:2203.06502.

[156] Krishna, C.L. and Murphy, R.R., 2017, October. A review on cybersecurity vulnerabilities for unmanned aerial vehicles. In 2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR) (pp. 194-199). IEEE.

[157] Humayun, M., Niazi, M., Jhanjhi, N.Z., Alshayeb, M. and Mahmood, S., 2020. Cyber security threats and vulnerabilities: a systematic mapping study. Arabian Journal for Science and Engineering, 45, pp.3171-3189.

[158] Razaque, A., Amsaad, F., Khan, M.J., Hariri, S., Chen, S., Siting, C. and Ji, X., 2019. Survey: Cybersecurity vulnerabilities, attacks and solutions in the medical domain. IEEE Access, 7, pp.168774-168797.

[159] Ruijters, E. and Stoelinga, M., 2015. Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. Computer science review, 15, pp.29-62.

[160] Senol, Y.E., Aydogdu, Y.V., Sahin, B. and Kilic, I., 2015. Fault tree analysis of chemical cargo contamination by using fuzzy approach. Expert Systems with Applications, 42(12), pp.5232-5244.

[161] Wu, D., Ren, A., Zhang, W., Fan, F., Liu, P., Fu, X. and Terpenny, J., 2018. Cybersecurity for digital manufacturing. Journal of manufacturing systems, 48, pp.3-12.

[162] Ganin, A.A., Quach, P., Panwar, M., Collier, Z.A., Keisler, J.M., Marchese, D. and Linkov, I., 2020. Multicriteria decision framework for cybersecurity risk assessment and management. Risk Analysis, 40(1), pp.183-199.

[163] Faisal, M.M.A. and Chowdhury, M.A.I., 2016, October. Bio inspired cyber security architecture for smart grid. In 2016 International Conference on Innovations in Science, Engineering and Technology (ICISET) (pp. 1-5). IEEE.

[164] Mthunzi, S.N., Benkhelifa, E., Bosakowski, T. and Hariri, S., 2019. A bio-inspired approach to cyber security. In Machine Learning for Computer and Cyber Security (pp. 75-104). CRC Press

[165] Inouye, S.K. and Charpentier, P.A., 1996. Precipitating factors for delirium in hospitalized elderly persons: predictive model and interrelationship with baseline vulnerability. Jama, 275(11), pp.852-857.

[166] Kornecki, A.J., Subramanian, N. and Zalewski, J., 2013, September. Studying interrelationships of safety and security for software assurance in cyber-physical systems: Approach based on bayesian belief networks. In 2013 Federated Conference on Computer Science and Information Systems (pp. 1393-1399). IEEE.

[167] Téglásy, B.Z., Gran, B.A., Katsikas, S., Gkioulos, V. and Lundteigen, M.A., 2020. Clarification of the Cybersecurity and Functional Safety Interrelationship in Industrial Control Systems: Barrier Concepts and Essential Functions. In Proceedings of the 29th European Safety and Reliability Conference (ESREL).

[168] Searle, R. and Renaud, K., 2023. Trust and vulnerability in the cybersecurity context. In HICSS, Hawaii.

[169] Vorozhtsova, T. and Skripkin, S., 2018, August. Ontological analysis of vulnerabilities in the energy sector. In Vth International workshop" Critical infrastructures: Contingency management, Intelligent, Agent-based, Cloud computing and Cyber security"(IWCI 2018) (pp. 202-206). Atlantis Press.

[170] Dahl, R.E., 2004. Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. Annals of the New York Academy of Sciences, 1021(1), pp.1-22.

[171] J. Doe, "Malware Evolution and Defence: A Survey," in IEEE Transactions on Dependable and Secure Computing, pp. 1-15, 2023.

[172] Cybersecurity and Infrastructure Security Agency (CISA), "Emotet: Understanding the Threat and Its Impact," 2023. [Online]. Available: https://www.cisa.gov/emotet-threat.

[173] S. Smith and R. Jones, "Analysing DDoS Attacks and Defence Mechanisms," in IEEE Security & Privacy, pp. 30-40, 2023.

[174] A. Johnson, "Mitigating Man-in-the-Middle Attacks: Techniques and Technologies," in IEEE Communications Magazine, pp. 50-60, 2023.

[175] B. Lee and C. Kim, "Phishing in the Age of SaaS: A Comprehensive Study," in IEEE Access, pp. 100-110, 2023.

[176] D. Patel, "SQL Injection Attacks and Defence Strategies," in IEEE Transactions on Information Forensics and Security, pp. 200-215, 2023.

[177] E. Thompson, "The Psychology of Password Attacks: Understanding and Preventing Social Engineering," in IEEE Security & Privacy, pp. 70-85, 2023.

[178] Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., & Akin, E. (2023). A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. Electronics, 12(6), 1333.

[179] Strom, B. E., Applebaum, A., Miller, D. P., Nickels, K. C., Pennington, A. G., & Thomas, C. B. (2018). Mitre att&ck: Design and philosophy. In Technical report. The MITRE Corporation.

[180] Alexander, O., Belisle, M., & Steele, J. (2020). MITRE ATT&CK for industrial control systems: Design and philosophy. *The MITRE Corporation: Bedford, MA, USA*, *29*.

[181] Georgiadou, A., Mouzakitis, S., & Askounis, D. (2021). Assessing mitre att&ck risk using a cyber-security culture framework. *Sensors*, *21*(9), 3267.

[182] Types of immunity - active immunity - passive immunity (2022) TeachMePhysiology. Available at: https://teachmephysiology.com/immune-system/immune-responses/types-of-immunity/

[183] C, R. (2022) Difference between innate and acquired immunity (with comparison chart), Bio Differences. Available at: https://biodifferences.com/difference-between-innate-and-acquired-immunity.html?utm_content=cmp-true (Accessed: 12 February 2024).

[184] Alenezi, Rafa, and Simone A. Ludwig. "Explainability of cybersecurity threats data using shap." *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2021.

[185] Jeyaraj, Anand, and Amir H. Zadeh. "Exploration and exploitation in organizational cybersecurity." *Journal of Computer Information Systems* 62.4 (2022): 680-693.

[186] Millstein, F. (2020). *Natural language processing with python: natural language processing using NLTK*. Frank Millstein.

[187] Wang, Jin, et al. "DeepVulSeeker: A novel vulnerability identification framework via code graph structure and pre-training mechanism." *Future Generation Computer Systems* 148 (2023): 15-26.

[188] Alqudhaibi, Adel, et al. "Predicting cybersecurity threats in critical infrastructure for industry 4.0: a proactive approach based on attacker motivations." *Sensors* 23.9 (2023): 4539.

[189] Ziems, Noah, and Shaoen Wu. "Security vulnerability detection using deep learning natural language processing." *IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2021.

[190] Yin, Jiao, et al. "Apply transfer learning to cybersecurity: Predicting exploitability of vulnerabilities by description." *Knowledge-Based Systems* 210 (2020): 106529.

[191] Huang, Jing, et al. "Smart contract vulnerability detection model based on multi-task learning." *Sensors* 22.5 (2022): 1829.

[192] Chen, Yizheng, et al. "Diversevul: A new vulnerable source code dataset for deep learning based vulnerability detection." *Proceedings of the 26th International Symposium on Research in Attacks, Intrusions and Defences*. 2023.

[193] Mantha, B.R. and García de Soto, B., 2021. Assessment of the cybersecurity vulnerability of construction networks. *Engineering, Construction and Architectural Management*, *28*(10), pp.3078-3105.

[194] Syed, R., 2020. Cybersecurity vulnerability management: A conceptual ontology and cyber intelligence alert system. *Information & Management*, *57*(6), p.103334.

[195] Shah, A., Farris, K.A., Ganesan, R. and Jajodia, S., 2022. Vulnerability selection for remediation: An empirical analysis. *The Journal of Defence Modeling and Simulation*, *19*(1), pp.13-22.

[196] Enayaty-Ahangar, F., Albert, L.A. and DuBois, E., 2020. A survey of optimization models and methods for cyberinfrastructure security. *IISE Transactions*, *53*(2), pp.182-198.

Appendix

Table A.1 show some examples of High-Profile Cyber-Attacks

- The first objective of Table 4 is to gather data related to the five attributes that are identified from the definitions of cyber vulnerability and attacks.
- The second objective of Table *4* is to identify any additional attributes of the ten cyber-attacks that may not be explicitly mentioned in the definitions of cyber vulnerability and attacks.

Table A.1. Five attributes of high-profile cyber-attacks [82]

ACK	Actor: Known Target	Actor: Alleged Attacker	Asset Targeted	Motivation	Effect on Targeted Asset	Attack Duration
1	Google (multinational specializing in Internet-related services and products)	Elderwood Gang (large Chinese cyberespionage organization)	Source code repositories that support supply chain functions	Collect valuable proprietary information of businesses	Gmail database was modified to allow extraction of information without detection	28 weeks (Jun to Dec '09)
2	Iran	Israel & US	Nuclear centrifuges controlled by computers at Natanz, Iran	Delay Iran's nuclear R&D program	1,000 centrifuges destroyed.	32 weeks (Nov '07 to Jun '10)
3	New York Times: publisher of American daily newspaper	Hackers who used methods of the Chinese military	Passwords and data of reporters and other employees	Obtain names of people who provided information about relatives of China's prime minister accumulating billions through business dealings	Data of 50 employees copied and uploaded to external server without detection	28 weeks (Oct '12 to Jan '13)
4	Chemical & defence firms in US	Covert Grove (group located in Hebei region in China)	Domain administrator credentials and networks of computers that store information	Collect valuable proprietary information of businesses	Data from 48 companies copied and uploaded to external server without detection	12 weeks (Jul to Sep '11)

5	The Spamhaus Project (not-for-profit that tracks spammers)	CyberBunker (an Internet service provider)	The Spamhaus Project website	Retaliate against Spamhaus for identifying CyberBunker as hosting spammers and asking its upstream service provider to cancel service	Website not available to users	2 weeks (Mar '13)
6	Target Corporation (American discount retailer)	Criminal group	Confidential customer information	Obtain confidential information	Data from 110 million customers copied and uploaded to external server without detected	8 weeks (Nov to Dec '13)
7	TJX Companies (American apparel and home goods company)	Criminal group	Credit and debit card numbers	Obtain confidential information to sell	Data from 94 million customers copied and uploaded to external server without detection	32 weeks (May '06 to Jan '07)
8	Bank customers	Aleksandr Andreevich Panin, a.k.a. "Gribodemon" and "Harderman " (Hacker)	1.4 million computers that store online banking credentials, credit card data, usernames, PINs, and other sensitive information	Obtain confidential information to sell	Sensitive information in 30,000 bank accounts was copied and uploaded to external server without detection	In progress
9	Computer owners	Criminals	Users' data or systems	Demand ransom to restore access	250,000 computers encrypted	In progress
10	Gaming companies	Criminals	Digital certificates for the secure exchange of information over the Internet using the public key infrastructure	Obtain confidential information to sell	Data from up to 30 gaming companies copied and uploaded to external server without detection	In progress

Table A.2. The relationship between security threats and vulnerabilities

	Threat	**Description**	**Vulnerabilities**	**Incidents**
1	DB	Data Breaches	V1, V3, V4, V5, V7	A perpetrator may employ various attack techniques, such as SQL injection, command injection, and cross-site scripting, to compromise a system. Additionally, vulnerabilities in virtualization technology can be exploited to extract sensitive data.
2	IAM	Weak Identity, Credential and Access Management	V1, V3	A potential attacker can take advantage of the absence of multifactor authentication or the use of weak passwords.
3	API	Insecure interfaces APIs	V1	An attacker can exploit vulnerabilities in the usage of APIs such as SOAP and the HTTP protocol. Bugs in APIs can also be leveraged to compromise a system.
4	SV	System Vulnerabilities	V4, V5, V6, V7	A perpetrator can launch an attack by exploiting vulnerabilities in Virtual Machine images, Hypervisors, and Virtual Networks.
5	AH	Account Hijacking	V1	In order to gain access to a system, attackers can utilize the victim's account.
6	MI	Malicious Insiders	V5, V7	A perpetrator can create a VM image that contains malware and then spread it through various means.
7	APT	Advanced Persistent Threats	V1, V4, V5, V6, V7	An attacker can exploit various vulnerabilities in a particular virtual cloud or APIs to implant persistent bugs in the target system, primarily for the purpose of harvesting data.
8	DL	Data Loss	V3, V4, V7	A perpetrator can utilize data-driven attack methods to obtain confidential information from other Virtual Machines that are co-located on the same server. They can also exploit the vulnerabilities in the data backup and storage process to harvest data.
9	IDD	Insufficient Due Diligence	V4, V6	A potential attacker can exploit weaknesses in compliance with rules when using a cloud system, such as the configuration of Virtual Machines and the sharing of data and technology.
10	ANU	Abuse and Nefarious Use of Cloud Services	V4	A perpetrator can launch an attack by utilizing anonymous accounts to access and share a customer's servers and data.
11	DOS	Denial of Service	V1, V2	A perpetrator can request additional IT resources in order to prevent authorized users from accessing cloud services.
12	STV	Shared Technology Vulnerabilities	V4, V6	A perpetrator can exploit virtual networks by sniffing and spoofing, or by taking advantage of the flexible configuration of Virtual Machines and hypervisors.

- **Extra outcomes of this research**

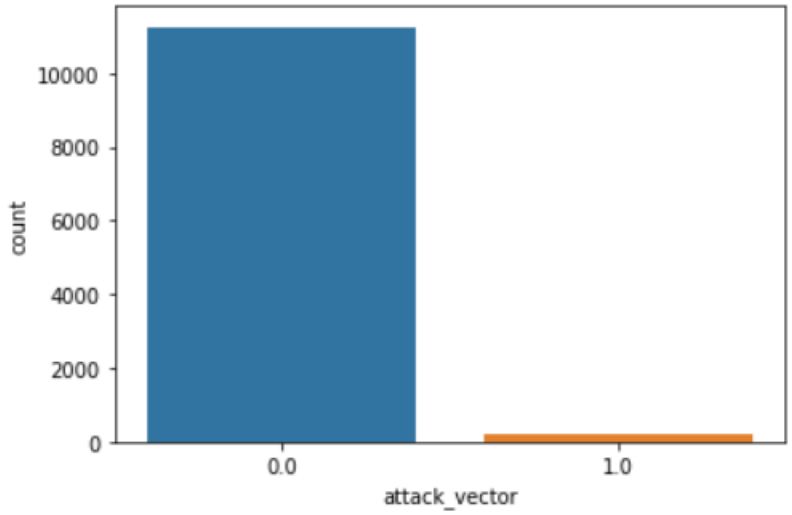


Figure A-1 Distribution of Attack Vector.

In the context of cybersecurity, attack vector typically refers to the method by which a cyber-attack is carried out.

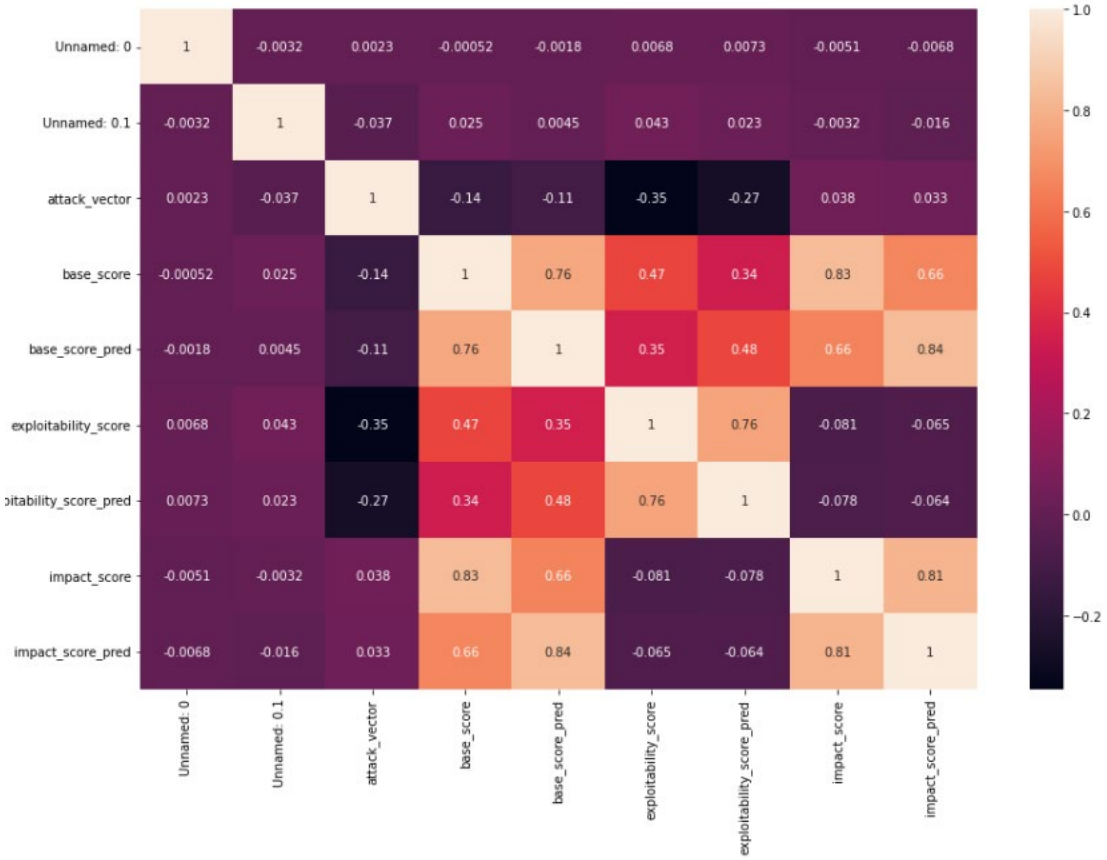


Figure A-2. Heatmap Representing a Correlation Matrix by Considering Attack Vector.

As Figure A-2 indicates:

- **base score** and **impact score** have a high positive correlation (around 0.83), indicating that as the base score of a vulnerability increases, its impact score tends to increase as well.
- **attack vector** has a strong negative correlation with both **exploitability score** and **exploitability score pred** (around -0.35 and -0.27 respectively), suggesting that as the ease of attack increases, the exploitability scores decrease, which might seem counterintuitive without additional context.

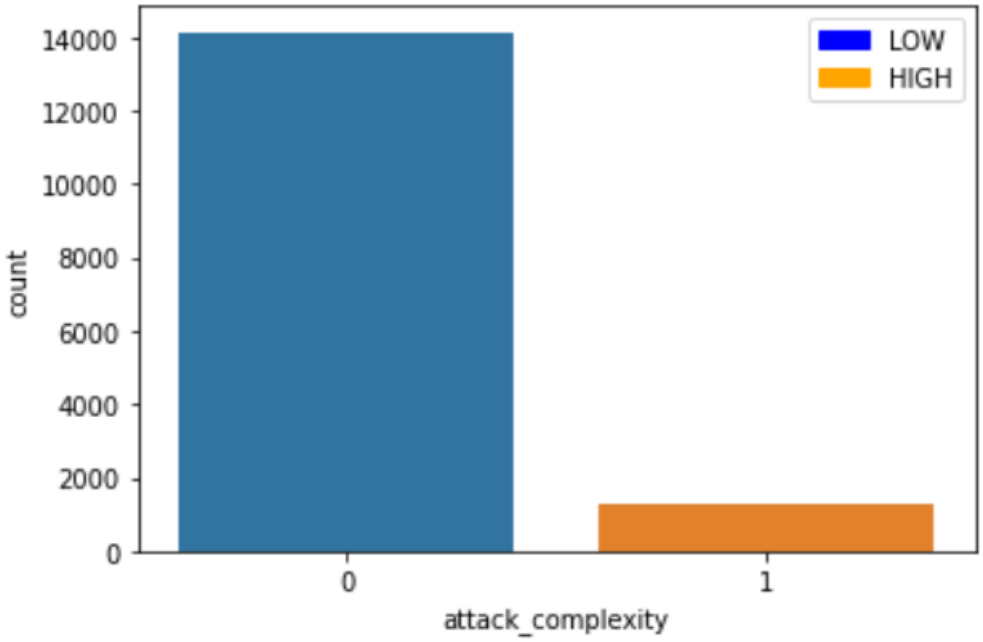


Figure A-3. Distribution of Attack Complexity.

In cybersecurity terms, attack complexity usually refers to the level of difficulty associated with executing a successful attack. A low complexity attack might require minimal skills or resources to carry out, while

a high complexity attack might involve more sophisticated methods and a higher level of skill.

Figure A-4. Heatmap Representing a Correlation Matrix by Considering Attack Vector and Attack Complexity.

As Figure A-4 shows:

- **base score** shows a strong positive correlation with both **exploitability score** (0.47) and **impact score** (0.83), which suggests that as the fundamental severity of a vulnerability increases, both its potential to be exploited and its impact on the system also tend to increase.
- **privileges required** has a moderately negative correlation with **exploitability score** (-0.58), which could imply that the more privileges required for an attack, the less exploitable the vulnerability is, possibly due to increased security measures associated with higher privilege levels.
- Predicted scores for base, exploitability, and impact (**score pred**) show very high correlations with their actual counterparts, indicating that the predictions are closely mirroring the actual scores.

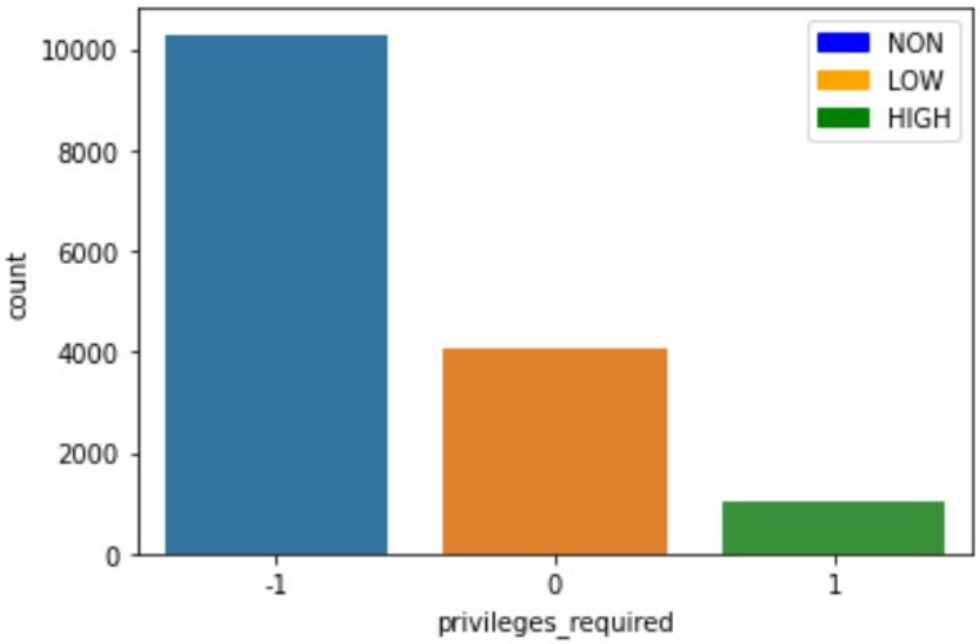


Figure A-5. Distribution of Privileges Required.

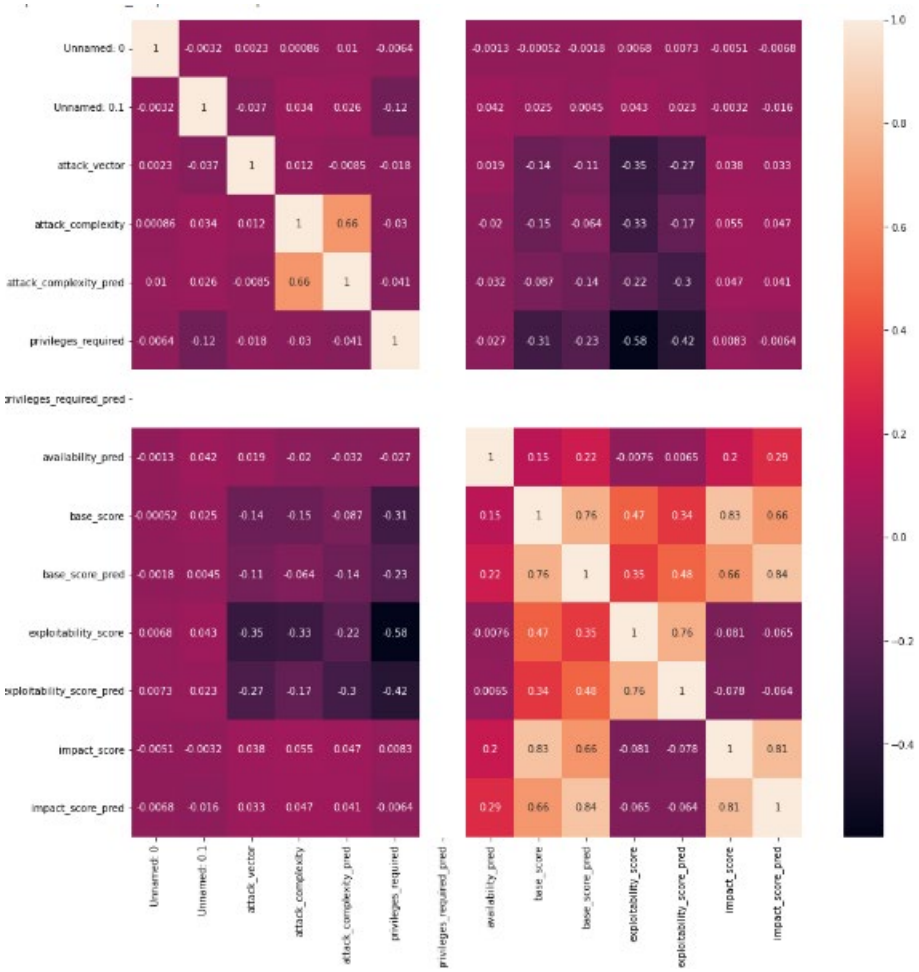


Figure A-6. Heatmap Representing a Correlation Matrix by Considering Attack Vector and Attack Complexity and Privileges Required.

As Figure A-6 illustrates:

- **base score** has a relatively strong positive correlation with **exploitability score** and **impact score**, suggesting that vulnerabilities that are more exploitable and have a higher impact also have a higher base score.
- **attack complexity** has a slightly negative correlation with **privileges required**, which might suggest that attacks that are more complex to carry out could potentially require fewer privileges. However, the correlation is not strong, so this relationship isn't necessarily predictive.
- Predictive scores (**score pred**) have strong positive correlations with their actual scores, which implies that the predictions align well with actual values.

❖ You can find this study outcomes in below link:

Some files and outcomes of this study are too large to be saved on GitHub. They can be downloaded from the Google drive location below:

https://drive.google.com/drive/folders/1zfm6pvpMxXDX6KiKCfVx3cbytALs5jYz?usp=share_link