Development of a Prototype Cyber Threat Intelligence (CTI) Framework for the Mining Industry

Proposed by:
MUHAMMAD UMER FAROOQ
2925331

Supervisors:

Dr. Helen Robertson (M)

Dr. Ahsan Mahboob

December 10, 2024

A proposal submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in partial fulfilment of the requirements for the degree of Master of Science (Dissertation) in Computer Science

School of Computer Science & Applied Mathematics
University of the Witwatersrand

Abstract

The increasing integration of digital technologies in the mining industry has heightened its vulnerability to cyber threats, which pose significant risks to operations, financial stability, and environmental safety. This research aims to develop a Prototype Cyber Threat Intelligence (CTI) framework specifically tailored for the mining sector, leveraging advanced data analytics and machine learning to improve cyber threat detection and mitigation. The objectives include designing a CTI system that can collect and analyze data from diverse sources such as network logs and open-source intelligence feeds, identifying patterns and anomalies indicative of cyber threats. A mixed-methods approach will be employed, combining a systematic literature review, empirical testing, and stakeholder consultations to ensure the framework is both technically robust and ethically sound. The methodology integrates machine learning models for threat classification and anomaly detection while considering privacy and regulatory concerns. The expected outcomes include enhanced resilience to cyber threats, reduced operational disruptions, and improved protection of critical infrastructure. This research will contribute to the broader field of cybersecurity in critical infrastructure, offering practical insights into the application of CTI systems within the mining industry.

Declaration

I, Muhammad Umer Farooq, hereby declare the contents of this research proposal to be my own work. This proposal is submitted for the Master of Science by Dissertation in Computer Science at the University of the Witwatersrand. This work has not been submitted to any other university, or for any other degree.

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Dr. Helen Robertson from the School of Computer Science and Mathematics, and Dr. M. Ahsan Mahboob, the Head of Sibanye Stillwater Digital Mining Laboratory (DigiMine), for their invaluable guidance and support throughout my research journey. I would like to thank and acknowledge the financial support provided by the Sibanye-Stillwater Digital Mining Laboratory (DigiMine), Wits Mining Institute (WMI).

Dr. Robertson's expertise and insightful feedback were instrumental in refining my work, and her encouragement was a constant source of motivation. I am equally grateful to Dr. Mahboob for his mentorship and for providing me with the opportunity to collaborate within the DigiMine lab, where his leadership and profound knowledge in the field greatly enhanced my research experience.

Their unwavering support, constructive feedback, and encouragement have been crucial to the successful completion of this proposal. I am deeply appreciative of their time and efforts, and I feel privileged to have worked under their guidance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Mining Industry plays an important role for the global economy. It depends on digital technologies to improve operations, safety, and sustainability. The reliance on technology exposes the industry to significant cyber threats that may disrupt operations, compromise safety, and cause financial loss and environmental harm Lenka *et al.* [2023]. Due to these risks, the need for effective cyber threat detection in mining has become more urgent than ever. Traditional and reactive cybersecurity measures often fail to address the complex and evolving threats Liu *et al.* [2022] specific to the mining sector. This has led to the growing importance of Cyber Threat Intelligence (CTI) System for the mining sector, which proactively detect, analyze, and respond to potential threats using intelligence feeds Webb *et al.* [2014]. These systems collect and analyze data from multiple sources, helping organizations identify and mitigate risks before they cause damage. Despite these potential threats, the effective and intelligent cyber threat detection system in mining has never been implemented. Traditional reactive cybersecurity measures used in the sector often fall short of addressing the sophisticated and evolving threats unique to the mining sector. This research aims to fill these gaps by developping tailored Cyber Threat Intelligence System Framework specifically for the mining industry. The CTI Framework will use advanced data analytics and machine learning algorithms to detect patterns and anomalies in threats data and it will improve the industry's ability to defend against cyber-attacks. Additionally, this study will examine the ethical implications of implementing such a system, ensuring it meets both technical and ethical standards. By designing a CTI system tailored to the mining sector, this research will address a critical need and contribute to the broader field of cybersecurity. In Figure: 1.1, World Economic Forum Global Risks Perception Survey 2023-2024 indicating a wide subset of the global population to potential digital and physical exploitation.

1.1 Problem Statement

The mining industry remains one of the most strategic industries in the global economy given that it supplies essential raw materials for several other industries. Since most of the mining companies are now leveraging on digital technologies to improve on pro-

Figure 1.1: World Economic Forum Global Risks Analysis Survey 2023-2024

duction, safety and overall performance, they are opening up numerous cybersecurity risks. There is an increasing use of digital control systems including SCADA systems, automation and remote monitoring systems that create new attack surfaces that can be targeted by cyber criminals. These risks are even amplified by the use of outdated equipment which were not developed with the present day cyber threats in mind.

The mining operations are most at risk of SCADA system infiltration, which manages and manages the essential equipment, energy consumption, and safety, among others. Due to the fact that SCADA systems are often installed in remote places with limited or unreliable connectivity, they are considered as prime targets for cyber attacks. It is therefore possible for cybercriminals to take advantage of such vulnerabilities to intrude into the systems, interfere with normal operation or even cause destruction of mining equipment which is a great risk to the lives of miners as well as the environment.

Another challenge that is associated with cybersecurity is the legacy systems that are still in use in the older mining operations. These systems are often aged and do not receive security updates frequently hence are vulnerable to known hacks. In addition, many mining companies still employ these legacy systems in systems. their operations, which Several makes large-scale it cyber hard incidents to that adopt have new occurred cybersecurity at measures mining without companies causing demonstrate an how overhaul crucial in it their is to address such risks. Evolution Mining, an Australian gold miner, was hit by ransomware in August 2024 which affected the company's IT systems. The attack was quickly contained, however it highlighted the fact that the company was at high risk of cyber incidents. For instance, Sibanye-Stillwater Cameron [2024] also fell prey to a cyberattack in July 2024 that led to IT system shutdowns and forced some processes to be handled manually. Norsk Hydro, a pan international metals and mining company, was hacked by ransomware in 2020 that led to plant shutdowns and financial losses in the millions of dollars Ravichandran *et al.* [2024]. Also, a cyber attack on a South American mining in company 2017 led to the loss of sensitive geological information while an Australian mining company was breached in 2019 which led to unauthorized access that affected automated processes and put the safety of workers at risk.

The cases described earlier demonstrate how complex the contemporary threat environment is for the mining industry and, thus, the necessity of developing and applying effective garrison approaches that will enhance protection of systems and networks that are fundamental to the operation of the enterprise, including SCADA, legacy systems, and remote environments. Cyber threats do not only lead to financial loss and disruption of business but also have the potential of endangering the lives of workers, stealing vital information, and even causing pollution. Hence, there is a need for advanced cybersecurity framework that address the peculiar challenges of mining industry.

Table 1.1: Cybersecurity Incidents in Mining Sector

Company	**Country**	**Date**	**Type of Attack**	**Impact**
Evolution Mining Limited [2024]	Australia	August 2024	Ransomware	IT systems compromised, though quickly contained; incident highlighted vulnerability to cyber threats.
Sibanye-Stillwater Cameron [2024]	South Africa	July 2024	Cyberattack (Type Undisclosed)	Temporary IT system outages; forced manual processing in operations, causing delays and increased operational costs.
Norsk Hydro Ravichandran *et al.* [2024]	Norway	March 2019	Ransomware	Major operational shutdowns across global operations; cost company over $70 million to recover; IT and production disruptions led to delays and losses.
Australian Mining Company	Australia	2019	Unauthorized Access	Unauthorized access to automated systems disrupted production; raised safety risks for workers and operational delays.
Vedanta Resources	India	April 2023	Ransomware	Systems at one of its smelters breached, causing partial shutdowns and impacting the supply chain; ransom demanded in cryptocurrency.
Vale S.A.	Brazil	November 2020	Phishing & Malware Attack	Phishing led to malware spread; disrupted supply chain and logistics for several days; company tightened security protocols post-incident.

Anglo American	United Kingdom / Global	May 2021	Ransomware	Ransomware attack on IT infrastructure affected mining operations globally; significant recovery costs incurred and data backups improved post-incident.
Newmont Corporation	United States	February 2022	Phishing Attack	Spear phishing campaign targeting executives, led to exposure of sensitive corporate data; company enhanced email filtering and employee training.
Rio Tinto	Global	March 2023	Data Breach	Attack exposed confidential project information; data breach led to increased competitive risks and raised concerns about insider threats.

1.2 The Importance of CTI for the Mining Industry

Due to the current trends of adopting digital technologies in the mining industry with an aim of improving on efficiency, safety and sustainability the industry is increasingly being exposed to cyber threats Lenka *et al.* [2023]. Cyber Threat Intelligence (CTI) is vital in dealing with such risks since it helps in understanding the possible risks and weaknesses of the system. While traditional reactive security measures are often sufficient in meeting standard cybersecurity needs, CTI provides proactive strategy which enables mining firms to recognise, assess and neutralize threats that may occur. Mining industry appropriate CTI frameworks can be useful in the development of mining industry specific cybersecurity solutions. These needs are usually attributed by the challenges that are characteristic of the industry which include working in remote areas, the use of outdated technology, and reliance on third party contractors. Thus, CTI helps mining companies to obtain real-time knowledge on the threats that exist in the market, and, thus, to protect such critical aspects as infrastructure, data, and operation continuity against the increasing cyber risks.

1.3 Research Questions

The following questions are formulated to guide the research for development of a Cyber Threat Intelligence (CTI) system tailored for the mining industry:

1. What are the key components and design principles required to develop a prototype Cyber Threat Intelligence (CTI) framework tailored to the unique operational challenges of the mining industry?

2. How can machine learning techniques be effectively applied to publicly available or synthetic datasets to identify patterns and anomalies relevant to mining-specific cyber threats?
3. What ethical, regulatory, and operational considerations are critical for the deployment of a CTI framework in the mining industry, and how can these be incorporated into the prototype design to ensure compliance and stakeholder acceptance?

1.4 Research Aims and Objectives

The adoption of digital technologies in the mining industry is essential for advancing productivity and safety. But it has also led it to significant cyber vulnerabilities. Cyber threats continuing to evolve. So, the need for a proactive approach to identify, analyze, and mitigate risks has become essential. The research aims and objectives focus on developing a robust CTI system Framework.

1.4.1 Research Aims

The primary aim of this research is to develop an effective Cyber Threat Intelligence (CTI) framework for the mining industry, designed to enhance cybersecurity resilience. This research aims to mitigate the increasing cyber threat risks within the mining industry, which has seen a sharp rise in digital vulnerabilities as it adopts more advanced technologies by developing a CTI Framework tailored to mining industry. This aim involves creating a specialized CTI system that can detect and respond to cyber threats, ensuring the safety, security, and continuity of mining operations in an increasingly digitalized environment.

1.4.2 Research Objectives

The objectives of this research are as follows:

- Develop a tailored Cyber Threat Intelligence (CTI) framework for the mining industry by identifying its unique cybersecurity challenges, such as remote operations, legacy systems, and integration with critical infrastructure, and by establishing design principles to address these challenges.
- Apply machine learning techniques to publicly available or synthetic datasets to develop models capable of detecting patterns and anomalies indicative of cyber threats in mining operations, followed by rigorous validation to ensure relevance and accuracy in a mining-specific context.
- Integrate ethical, regulatory, and operational considerations into the design of the CTI framework, ensuring compliance with industry regulations (e.g., GDPR, mining-specific standards), addressing privacy concerns, and incorporating stakeholder feedback to ensure practical implementation and acceptance.

1.5 Limitations

The primary limitation of this research lies in the generalizability of its findings across the diverse operational environments within the mining industry. Although the study aims to develop a robust Cyber Threat Intelligence (CTI) framework tailored specifically for mining, it relies on particular datasets and threat scenarios that may not fully encompass all real-world contexts. Variations in mining operations, network infrastructure, and data types across different locations and organizations may affect the performance and applicability of the proposed CTI framework. Furthermore, the effectiveness of the suggested data analysis and machine learning methods in detecting threats and improving model interpretability may differ depending on the complexity and diversity of cyber threat data available at various mining companies. Therefore, while this research seeks to enhance cybersecurity practices within the mining industry, its broader application across the industry may be limited.

1.6 Overview

The digital transformation of the mining industry has brought significant improvements in operational efficiency, safety, and environmental sustainability. Technologies such as IoT, data analytics, and automation have optimized mining processes, but they have also introduced new cybersecurity vulnerabilities. These digital advancements have made mining operations more susceptible to sophisticated cyber threats, which could disrupt production, compromise sensitive data, cause financial losses, and even threaten worker safety. Given the importance of mining infrastructure, it is crucial to adopt a proactive approach to cybersecurity to safeguard both the safety and continuity of operations.

The conventional cybersecurity measures are usually not sufficient for the unique conditions of the mining industry, rare earth materials, and the associated infrastructure, remote locations, and outdated equipment. In light of these gaps, this paper proposes the development of a Cyber Threat Intelligence (CTI) and framework machine for learning the to mining identify, industry. fence, The and framework respond is to intended the to threats use in advanced order data to mining increase methods the industry's cyber defense capabilities.

The suggested CTI framework will use various sources of information such as network logs and OSINT to identify patterns that could be considered as cyber threats. The framework is designed to provide real time monitoring and rapid response and has been developed with the understanding of the operational and technical considerations of the mining sector. Also, this study will also discuss on the ethical implications of using such a system, to ensure that the system is in compliant with industry best practices, legislations and regulations concerning privacy and data integrity.

The purpose of this paper is to design a CTI framework appropriate for the mining industry so that a specific solution for protecting crucial mining facilities can be developed in the framework of this research. The expected results of the work are as follows: Improved capability in identifying threats, lower chance of incidents that might lead to

interruption of operations and enhanced protection of information and assets. This research not only enhances the security of the mining industry but also provides a case study of how CTI is used in other sectors that are also experiencing similar digitalization problems in the critical infrastructure.

Chapter 2

Background and Literature Review

2.1 Introduction

This chapter discusses the background and prior research related to Cyber Threat Intelligence (CTI) systems in detail. This section provides the foundational understanding of CTI frameworks and situates the current research within the context of existing literature. The review covers the evolution of cybersecurity challenges in the mining industry, key components of CTI frameworks, and the use of data analytics and machine learning techniques to identify cyber threats. The chapter also discusses the ethical considerations and regulatory requirements relevant to developing a CTI system in a highly specialized industrial environment.

2.1.1 Cyber Threat Intelligence (CTI)

Cyber Threat Intelligence (CTI) refers to the collection, analysis, and dissemination of information related to current and potential cyber threats targeting an organization, system, or infrastructure Kotsias *et al.* [2023]. It involves gathering data from various sources such as open-source intelligence (OSINT), network logs, threat feeds, and internal security events. This data is then analyzed to identify patterns, trends, and specific tactics, techniques, and procedures (TTPs) used by cyber adversaries. The ultimate goal of CTI is to provide actionable insights that allow organizations to proactively defend against threats, identify vulnerabilities, and anticipate future attacks, thereby enhancing overall cybersecurity posture Webb *et al.* [2014].

CTI is typically categorized into three types based on the level of detail and use:

- **Tactical CTI:** Focuses on immediate, actionable threat data, such as specific malware signatures, IP addresses, or domain names used by cybercriminals. It provides information that can be directly implemented into security systems (e.g., firewalls or antivirus software).
- **Operational CTI:** Provides insights into attack campaigns, including the strategies and methods used by attackers, and is useful for identifying trends or clusters of attacks across an organization or sector.

- **Strategic CTI:** Focuses on high-level intelligence regarding the cyber threat landscape, helping organizations understand the motivations, capabilities, and intent of threat actors. It is more relevant for senior leadership to make informed decisions about long-term security and risk management.

2.1.2 Distinction from Related Cybersecurity Tools

While Cyber Threat Intelligence (CTI) focuses on proactive threat identification and mitigation through data-driven insights, other cybersecurity tools like antivirus software typically focus on reactive measures, protecting systems from known threats.

Here's how CTI differs from tools like antivirus software:

Focus and Purpose:

- **CTI:** Aims to proactively identify and mitigate threats before they affect the system. It provides a broader strategic view of threats, identifying emerging risks and long-term trends.
- **Antivirus:** Primarily focused on detecting and removing known malware or malicious files based on predefined signatures. It reacts to threats once they've entered the system, offering defense through signature-based detection.

Scope:

- **CTI:** Covers a wide range of cyber threats, including malware, ransomware, phishing attacks, Advanced Persistent Threats (APTs), and insider threats. CTI gathers information on threat actors' tactics, techniques, and procedures (TTPs) to predict and prevent attacks.
- **Antivirus:** Primarily targets known types of malware. Antivirus software typically focuses on file-based threats, often using signature databases to identify malicious code.

Data Sources:

- **CTI:** Involves gathering data from various sources, such as open-source intelligence (OSINT), threat feeds, network logs, and industry-specific reports. It often integrates machine learning and data analytics to recognize patterns and anomalies in threat data.
- **Antivirus:** Uses a signature-based approach that compares files and behaviors against a database of known malware signatures. Antivirus tools may also use heuristic analysis to detect unknown threats based on behavior, but the scope of analysis is generally more limited.

Proactivity vs. Reactivity:

- **CTI:** Is inherently proactive, aiming to anticipate and prepare for threats before they occur. It provides security teams with actionable intelligence that helps them adjust defenses and response strategies.
- **Antivirus:** Is reactive, meaning it relies on detecting threats after they have been introduced to the system. It typically requires regular updates to its signature database to remain effective against new threats.

Integration:

- **CTI:** Integrates with various security systems (firewalls, intrusion detection/prevention systems, SIEM tools, etc.) to improve overall security strategy. It can provide security teams with insights to update or fine-tune defenses based on emerging threats.
- **Antivirus:** Operates as an individual tool, designed to be installed on endpoints and monitor for specific, known threats, often without direct integration into the broader security ecosystem of an organization.

2.2 Background

The mining industry has become integrated with the digital ecosystem in a way that has led to increased operations automation, fully better dependent safety on standards, the and digital better technologies overall which, performance. in However, turn, this makes has the also operations made prone the to mining advanced cyber threats. These threats are usually sophisticated; they may range from ransomware that can freeze operations to Advanced Persistent Threats (APTs) for intellectual property theft (IP) as reported by gately2023russian.

2.2.1 Industry 4.0 and Digitalization

Industry 4.0 has brought about digital changes across different industries including mining through the use of IoT, machine learning, cloud and big data. In mining, real time monitoring systems, autonomous mining equipment and predictive maintenance are fast becoming standard practice. However, this transformation poses threat to the critical infrastructure and exposes it to cyber threats [Wang and Lu 2013; Sajid *et al.* 2016].

2.2.2 Cyber Threat Landscape in Mining

The mining industry is at high risk of cyber threats because of the importance of mining operations. Some of the examples of past cyber incidents comprise of ransomware which targeted Evolution Mining in 2024 and constricted the operations of the company furthermore by stealing the geological data or attacking the automated systems. The

cases such as the ransomware attack on Evolution Mining in 2024 and Norsk Hydro in 2020 have shown the vulnerability and the consequences on productivity and financial performance.

2.2.3 SCADA Systems and Vulnerabilities

Supervisory Control and Data Acquisition (SCADA) systems are used in mining for managing and controlling the physical processes. However, the traditional SCADA systems have weaknesses that cannot withstand the threats that are associated with the integration of cloud may and cause IoT. irreparable This loss is in where mining the operations need [Wang and Lu 2013]. for an enhanced CTI system becomes evident since such vulnerabilities

2.2.4 Some Considerations in Respecting Cyber Security in Mining Operations

- **Remote and Harsh Environments:** This is because most of the mining operations are carried out in remote places with poor or no connection at all, which makes it difficult to implement and sustain efficient cybersecurity measures.
- **Legacy Systems and Modernization:** It is still possible to encounter legacy facilities in mining facilities, and it is hard to protect such systems, and the process of upgrading these without creating a disturbance to the operations is a big challenge.
- **Supply Chain Vulnerabilities:** The mining industry is highly dependent on third-party vendors and contractors, which increases the security risks and requires extensive supply chain security precautions.

2.3 Related Work

2.3.1 Cyber Threat Intelligence (CTI) Frameworks

CTI frameworks have emerged as crucial solutions for enhancing the proactive capabilities of cybersecurity defenses. They leverage data collection, analysis, and dissemination to provide actionable insights into emerging threats. Various researchers have proposed different components for CTI frameworks, each tailored to address specific cybersecurity needs.

Existing Framework Comparison

Existing Cyber Threat Intelligence (CTI) frameworks vary in structure, focus, and capabilities, addressing different facets of cybersecurity needs across industries. A comparison of prominent CTI frameworks reveals a variety of approaches to data collection, analysis, and intelligence dissemination, which contribute uniquely to threat intelligence development.

MITRE ATT&CK Framework Georgiadou *et al.* **[2021]** *Focus:* Threat tactics, techniques, and procedures (TTPs) Shahi [2018].
Components: Provides a detailed matrix of attack techniques that allow organizations to understand the methods and behaviors of adversaries.
Adoption: Widely adopted across multiple sectors, MITRE ATT&CK assists in understanding attacker behavior and mapping defensive measures to specific threat actions.

Diamond Model of Intrusion Analysis Caltagirone *et al.* **[2013]** *Focus:* Links threat activity to the context, including attacker, infrastructure, and victim relationships.
Components: Emphasizes understanding the relationships between adversaries and their targets by examining the four core elements—adversary, infrastructure, capability, and victim.
Adoption: Popular in critical infrastructure sectors, it aids in identifying threat campaigns and understanding complex threat vectors.

Lockheed Martin Cyber Kill Chain Naik *et al.* **[2022]** *Focus:* The stages of cyber attacks, from initial intrusion to data exfiltration.
Components: Includes stages like reconnaissance, weaponization, delivery, exploitation, and actions on objectives.
Adoption: Utilized by security teams to map out attack stages and identify gaps in defenses for preemptive countermeasures.

STIX/TAXII Protocols Provatas *et al.* **[2023]** *Focus:* Standardized threat data formatting and sharing.
Components: Structured Threat Information Expression (STIX) provides a common language for sharing threat intelligence, while Trusted Automated Exchange of Indicator Information (TAXII) facilitates secure and automated data exchange.
Adoption: Primarily used by organizations seeking interoperability in threat data sharing, enhancing collaboration across different cybersecurity platforms.

OpenIOC Janoti *et al.* *Focus:* Rapid identification of indicators of compromise (IOCs).
Components: Offers a framework for describing, archiving, and sharing threat intelligence specific to forensic artifacts and attack indicators.
Adoption: Commonly used for incident response to detect and isolate threats based on predefined IOCs.

CTI Framework Adoption in Other Critical Infrastructures Kayode-Ajala **[2023]**

Cyber threat intelligence frameworks have been increasingly adopted across various critical infrastructure sectors, such as energy Gong and Lee [2021], healthcare, finance, and government, where cybersecurity is paramount to ensure continuity and protect against disruptive attacks.

Energy Sector Gong and Lee **[2021]** *Application:* The energy sector, particularly in Supervisory Control and Data Acquisition (SCADA) systems, uses frameworks like

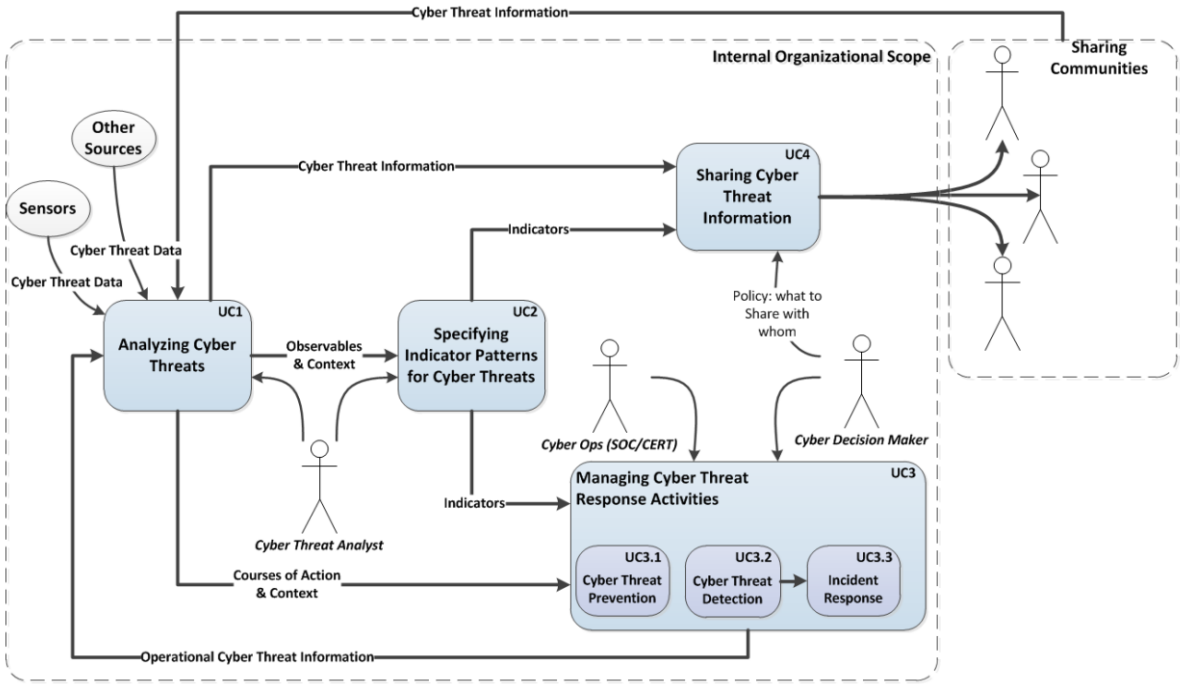

Figure 2.1: [Barnum 2012] Core Use Cases Targeted by STIX

MITRE ATT&CK and the Diamond Model to detect and respond to attacks that could affect power grid stability.
Challenges: Due to the complexity and real-time nature of SCADA systems, CTI frameworks in this sector often incorporate advanced analytics and machine learning for anomaly detection to ensure rapid threat mitigation.
Examples: Many energy companies have adopted the Cyber Kill Chain model to prevent and disrupt threats across the different stages of cyberattacks targeting power grids.

Healthcare Sector Krauss and Papesh [2022] *Application:* CTI frameworks like STIX/TAXII are widely used in healthcare for secure threat intelligence sharing, especially for protecting patient data and healthcare services.
Challenges: Healthcare networks, which contain sensitive patient information and are increasingly targeted by ransomware, rely on frameworks with strong data privacy and real-time response capabilities.
Examples: Integration with ISACs Bugiardini *et al.* [2016] for the healthcare sector provides enhanced collaboration and intelligence sharing to keep up with rapid threat evolution.

Financial Sector bin Mohd Aziz [2024] *Application:* The financial industry uses frameworks like MITRE ATT&CK to monitor and respond to fraud, phishing, and advanced persistent threats (APTs) targeting financial institutions.
Challenges: Given the sector's high vulnerability to fraud and data theft, CTI frameworks are often tailored to identify threat vectors unique to financial data and transactions, such as money laundering and payment fraud.

Examples: Financial institutions participate in FS-ISAC, a sector-specific information-sharing platform that allows them to stay ahead of cross-border cyber threats.

Government and Defense bin Mohd Aziz [2024] *Application:* Governments and defense sectors use comprehensive CTI frameworks like MITRE ATT&CK and Diamond Model to protect national security infrastructure.
Challenges: These sectors face complex threats from state-sponsored actors, necessitating robust intelligence sharing and high-level threat attribution capabilities.
Examples: Government agencies often utilize frameworks integrated with global intelligence alliances and automated response systems to mitigate sophisticated attacks on critical infrastructure.

Components of a CTI Framework

CTI Data Collector *Role and Functionality:* The data collector component is responsible for aggregating raw cyber threat data from multiple sources. These include OSINT feeds, network logs, vendor-provided threat intelligence, and even data from the dark web. The data collected must be diverse and comprehensive to ensure accurate threat detection.

Literature Examples:

- [Lee and Shon 2016]: Developed an OSINT-focused data collection strategy that ensures timely and relevant threat information. Their framework emphasizes preparing and implementing an OSINT plan before gathering and analyzing data from open sources. In Fig. 2.4, we see the OSINT-based Cyber Threat Inspection Framework proposed by Lee and Shon [2016], designed specifically for critical infrastructures.
- [Ryandy *et al.* 2020]: Outlined a systematic approach to data collection, emphasizing the importance of processing, analysis, and evaluation to ensure that the collected data is usable for threat intelligence purposes.
- [Tundis *et al.* 2022]: Focused on feature selection and OSINT source identification, demonstrating that a well-designed data collection system can significantly impact the quality of threat intelligence.

Analysis Medium *Role and Functionality:* This component transforms raw data into actionable intelligence through pre-processing, correlation analysis, pattern detection, and anomaly recognition. The analysis medium uses algorithms and heuristics to classify and prioritize threats.

Literature Examples:

- [Kim *et al.* 2016]: Emphasized the importance of structured data analysis, correlation techniques, and the use of YARA rules for malware detection. Their research shows how the analysis medium can derive meaningful insights from large datasets.

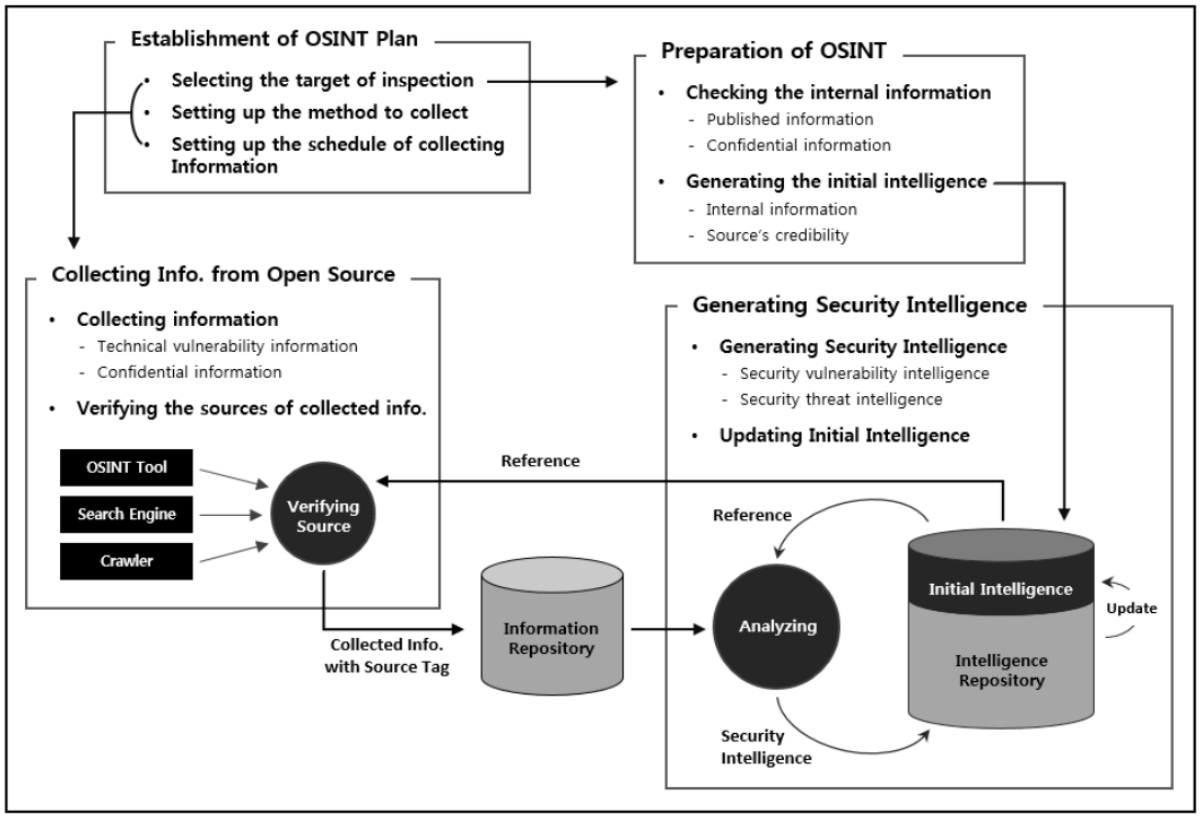

Figure 2.2: [Lee and Shon 2016] proposed Open Source Intelligence base Cyber Threat Inspection Framework for Critical Infrastructures

- [Noor *et al.* 2019]: Proposed an analysis approach that incorporates cyber threat attribution, helping organizations understand the origin and intent behind cyber-attacks.
- [Islam *et al.* 2022]: Integrated network data with a threat detector and alert validation system, highlighting the necessity of real-time data analysis to reduce false positives and improve threat response.

Information Platform *Role and Functionality:* This component acts as the user interface and decision-support system for disseminating analyzed intelligence. It allows for threat information sharing, real-time monitoring, and supports collaborative efforts among stakeholders.

Literature Examples:

- [Böhm *et al.* 2018]: Designed a CTI information platform with features like data filtering, mapping, rendering, and user interaction. Their focus was on making threat intelligence more accessible and actionable for security teams.
- [Kim *et al.* 2016]: Focused on the conversion of data into security rules, making the information platform a critical part of cybersecurity operations. This frame-

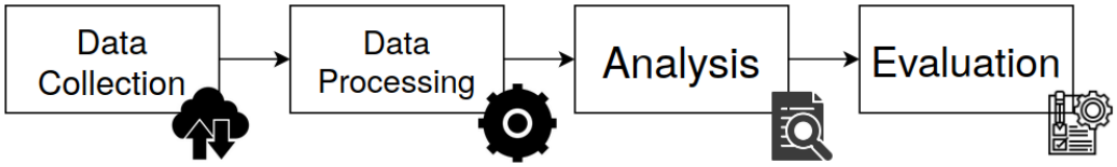

Figure 2.3: Research Framework for Data Collection proposed by [Ryandy *et al.* 2020]

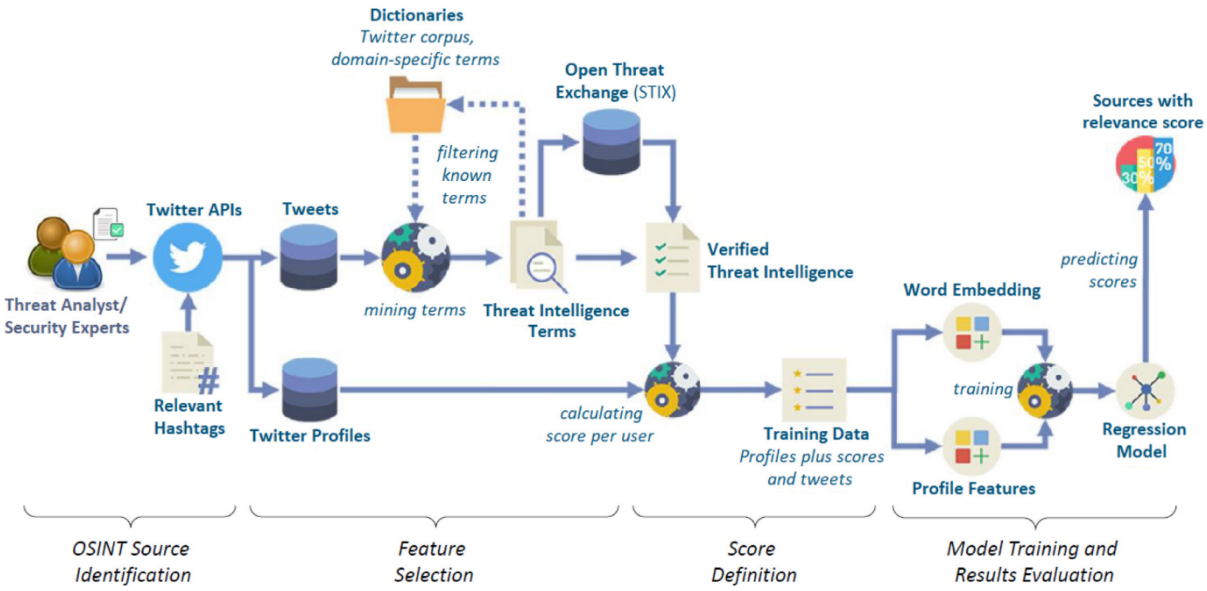

Figure 2.4: Research method overview proposed by [Tundis *et al.* 2022]

work highlighted the use of automated security updates based on the analyzed intelligence.

- [Papastergiou *et al.* 2021]: Proposed a comprehensive information-sharing model that integrates deep and dark web mining, live monitoring, and data protection orchestrators. Their approach, which included the HybridNet and ShareNet components, emphasizes the importance of robust and secure information dissemination.

2.3.2 Advanced Data Analytics and Machine Learning in CTI

The application of advanced data analytics and machine learning in CTI frameworks Naseer [2023] is becoming increasingly prevalent. These technologies enhance the capability to detect patterns, classify threats, and predict future cyber-attacks.

Machine Learning Techniques

Algorithms such as decision trees, support vector machines (SVMs) Deliu *et al.* [2017], and neural networks Yu *et al.* [2023] are commonly used for threat detection. Unsu-

pervised learning methods, like clustering, help in identifying unknown threat patterns, while supervised learning assists in classifying known threats.

Data Fusion and Anomaly Detection

Techniques like time-series analysis and real-time data fusion Song *et al.* [2022] are essential for identifying anomalies in network behavior. For instance, Islam *et al.* [2022] utilized data from both network and business operations to validate alerts and ensure that only genuine threats are flagged.

Case Studies in Industrial Settings

Review studies that have successfully implemented machine learning in critical infrastructure protection, such as detecting anomalies in SCADA systems or preventing ransomware attacks. Discuss how these methods can be adapted to the mining industry.

Collaborative Cybersecurity Initiatives and Intelligence Sharing

- **Industry Partnerships and Alliances:** The role of collaborations among mining companies and cybersecurity organizations in improving threat intelligence capabilities.
- **Information Sharing Platforms:** The significance of platforms like ISACs (Information Sharing and Analysis Centers) that facilitate threat intelligence sharing across the industry.
- **Global Cybersecurity Alliances:** Participation in global threat intelligence networks to stay updated on emerging cross-border cyber threats.

2.3.3 Limitations of CTI Frameworks in Mining-Specific Contexts

While Cyber Threat Intelligence (CTI) frameworks offer significant advantages in improving the cybersecurity posture of various industries, there are specific challenges and limitations when applying these frameworks to the mining sector. These limitations arise due to the unique characteristics and operational needs of mining operations, including their complex, remote environments and reliance on legacy systems.

- **Complex and Heterogeneous IT/OT Environments:** Mining operations often involve a mix of Information Technology (IT) and Operational Technology (OT) systems, which have distinct requirements and vulnerabilities. CTI frameworks that are primarily designed for IT environments may struggle to address the complexities of OT systems, such as Supervisory Control and Data Acquisition (SCADA) systems, which are commonly used in mining operations. The integration of CTI into these diverse environments may be challenging and may require custom solutions to bridge the gap between IT and OT security [Wang and Lu 2013].

- **Limited Connectivity in Remote Locations:** Many mining operations are located in remote and harsh environments where internet connectivity is either limited or unreliable. This makes the deployment of CTI systems challenging, as many CTI tools rely on real-time data feeds and cloud-based analytics. In remote locations, mining companies may face difficulties in implementing robust CTI systems that require continuous data collection and analysis.

- **Lack of Real-Time Threat Data:** CTI systems often depend on timely and accurate threat intelligence data, such as threat feeds, open-source intelligence (OSINT), and logs. However, in the mining industry, due to its global supply chain and fragmented data sources, obtaining real-time threat data may be difficult. Furthermore, many mining companies are not part of established Information Sharing and Analysis Centers (ISACs), which would otherwise provide a channel for sharing threat intelligence across industries.

- **Inadequate Threat Attribution:** One of the critical aspects of CTI is accurately attributing cyberattacks to specific threat actors. In the context of mining, particularly when dealing with attacks on critical infrastructure, it can be challenging to attribute threats reliably due to the involvement of multiple adversaries, including state-sponsored actors, cybercriminal groups, and insiders. Additionally, mining companies may lack the sophisticated tools or expertise to conduct in-depth threat attribution analysis.

- **Difficulty in Applying Machine Learning Models:** While machine learning (ML) and data analytics are commonly used in CTI frameworks to identify patterns and predict future threats, the application of these technologies in the mining industry may face hurdles. Mining companies may not have the required data infrastructure or historical data to train accurate machine learning models. Additionally, mining-specific threat scenarios, such as attacks on equipment or geospatial data, may require specialized models that do not fit within the conventional CTI frameworks.

- **Lack of Skilled Personnel:** The mining industry, especially in remote regions, may face a shortage of cybersecurity professionals with expertise in CTI and the specific challenges of protecting industrial control systems. Mining companies may struggle to build the internal capacity to manage and integrate CTI effectively into their cybersecurity strategies, leading to gaps in threat detection, analysis, and response.

- **Cost and Resource Constraints:** Implementing a robust CTI system, especially one that incorporates machine learning, advanced analytics, and real-time threat intelligence feeds, can be costly. Mining companies, particularly small- to medium-sized operations, may lack the financial resources to invest in comprehensive CTI solutions. This can lead to underinvestment in critical cybersecurity infrastructure, leaving systems vulnerable to cyberattacks.

Future Trends and Emerging Threats

- **Zero Trust Architecture** Stafford [2020]**:** Exploring the adoption of zero trust principles in securing industrial networks.
- **Blockchain Technology** Prakash *et al.* [2022]**:** The potential of blockchain for securing data and communications in mining operations.
- **Quantum Computing Risks** Brooks [2024]**:** How advancements in quantum computing could pose new cybersecurity threats to the mining industry.

Chapter 3

Research Methodology

3.1 Research Design

This research adopts a multi-phase design tailored to address the unique cybersecurity challenges encountered by the mining industry. The study is structured to incorporate both qualitative and quantitative approaches, integrating data from various sources and analyzing it through machine learning techniques to develop a specialized Cyber Threat Intelligence (CTI) framework. The research follows an iterative design that involves data collection, model development, ethical considerations, and system evaluation, as outlined in the following phases (see section: 3.2). Figure 3.1 shows the flowchart for the proposed Cyber Threat Intelligence System Framework for the Mining Industry.

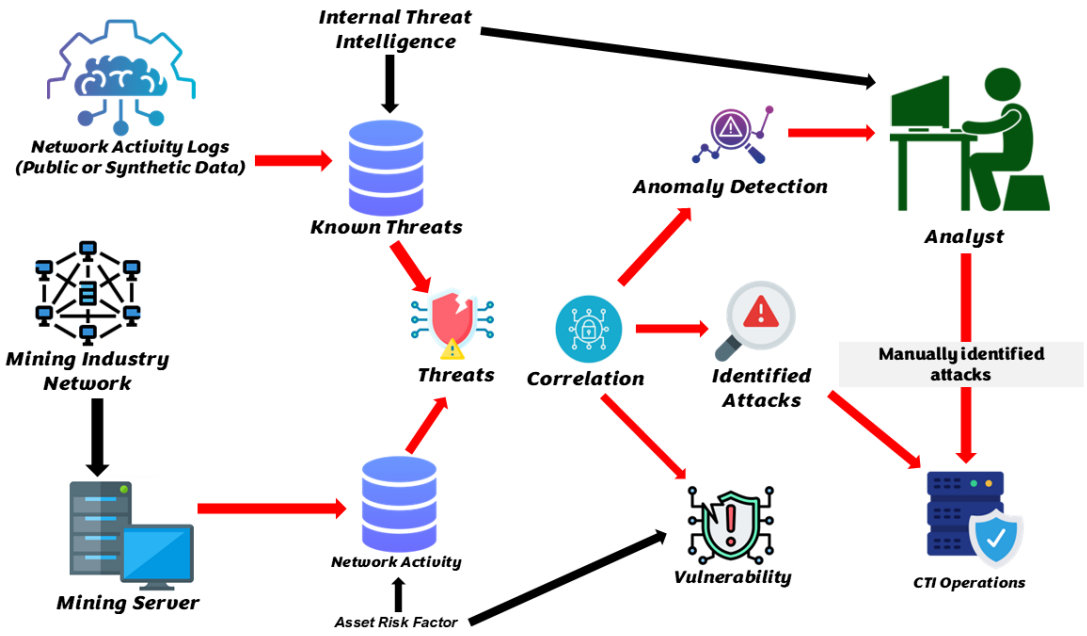

Figure 3.1: Flowchart for the Proposed Cyber Threat Intelligence System Framework for the Mining Industry

3.2 Methods

The research will follow a comprehensive, multi-phase approach tailored to address the unique cybersecurity challenges encountered by the mining industry. The process will begin with a requirements analysis, which will involve collecting stakeholder opinions and reviewing relevant literature to identify specific needs and gaps in current cybersecurity frameworks. Then, during the framework development phase, a customized architecture will be developed, integrating data from various sources, such as network logs and open-source intelligence feeds. Following this, real-time threat data will be collected and aggregated, which will then be analyzed using advanced machine learning techniques to detect patterns, classify, and cluster the logs and anomalies. Ethical considerations will be carefully examined to ensure compliance with industry regulations and data privacy standards. A prototype of the CTI System will be developed and tested in a controlled environment to evaluate its effectiveness. Then, the system's performance will be assessed, and detailed documentation will be prepared, offering insights and recommendations for practical implementation within the mining industry.

3.2.1 Pre-Modelling Phase

The pre-modelling phase focuses on the initial setup and preparation needed to build an effective Cyber Threat Intelligence (CTI) system. It includes data collection, data preprocessing, and defining the threat intelligence goals specific to the mining industry.

Data Collection

To build an effective CTI framework, a systematic approach to data collection is adopted, emphasizing diverse and comprehensive threat intelligence data relevant to mining. In addition to raw data from various real-time sources such as network logs, open-source intelligence (OSINT) feeds, and vendor-specific threat intelligence reports, we will integrate **synthetic and public datasets** to simulate potential cyber threats and validate the system's performance.

Synthetic and Public Datasets:

Given the challenges of accessing proprietary or industry-specific data, synthetic and public datasets will be heavily utilized in this research. These datasets provide valuable, publicly available data on network traffic, malware behaviors, and common attack strategies, which are critical for training and evaluating machine learning models. Here are some datastets listed below:

- **CICIDS 2017 and 2018:** Canadian Institute for Cybersecurity's Intrusion Detection System datasets, which include traffic data and attack scenarios across various networks, such as botnet attacks, DoS (Denial of Service), and exploitation of vulnerabilities.
- **UNSW-NB15:** A widely used dataset for network traffic and cybersecurity research that includes a broad spectrum of cyberattacks like DoS, web attacks, and exploits.

- **DARPA 1999:** A classic dataset for evaluating intrusion detection systems, which includes labeled attack types such as rootkits and DoS.
- **KDDCup 1999:** Contains simulated network traffic data with a variety of attack types, which is useful for training anomaly detection algorithms.

To build an effective CTI framework, a systematic approach to data collection is adopted, emphasizing diverse and comprehensive threat intelligence data relevant to mining. We will collect raw data from various sources such as network logs, open-source intelligence (OSINT) feeds, and vendor-specific threat intelligence reports. In mining operations, sensor and operational data can also be leveraged to detect abnormal patterns in the network.

Type and Scope of Data Sources

Data will be collected from various sources to capture the full spectrum of potential cyber threats:

- **Network Logs:** Network traffic data, including firewall logs, intrusion detection system (IDS) logs, and packet captures, will be collected from operational networks. This data reveals patterns and potential anomalies in network behavior.
- **Endpoint Data:** Logs from endpoint devices (e.g., workstations, IoT sensors, mobile devices) will be collected to monitor user activity and detect endpoint-based threats. Endpoint detection and response (EDR) data will provide insights into suspicious activities, processes, and applications.
- **Open Source Intelligence (OSINT):** Threat intelligence feeds from OSINT platforms like AlienVault, MISP, and forums, along with social media and dark web sources, will help to track new vulnerabilities, known exploits, and other indicators of compromise (IOCs).
- **Historical Incident Data:** Past cyber incidents within the mining industry, including ransomware attacks and unauthorized access events, will be collected from industry reports and public repositories. This historical data will be used to train and validate the model.

Data Collection Frequency

Data will be collected at varying frequencies to accommodate the real-time nature of threat detection and historical trend analysis:

- **Real-time Collection:** Network and endpoint logs will be captured in real-time to support immediate threat detection and mitigation.
- **Daily Aggregation:** OSINT data will be aggregated daily to capture trends and emerging threats without overwhelming the system.
- **Quarterly Historical Analysis:** Historical incident data will be reviewed and updated quarterly to incorporate any recent threats relevant to mining operations.

Data Preprocessing and Cleaning

Preprocessing ensures that the collected data is clean, consistent, and ready for analysis:

- **Data Cleaning:** We will clean and transform data into a usable format. This will involve handling missing data, removing duplicates, and normalizing data formats. Duplicate entries and irrelevant data points will be removed. Null values will be addressed to improve data quality.
- **Normalization:** Standardization of formats (e.g., timestamp consistency) and normalization of data across sources.
- **Feature Extraction:** Key features for cyber threat detection, such as IP addresses, activity timestamps, and user patterns, will be extracted. Principal Component Analysis (PCA) will reduce noise and dimensionality, improving computational efficiency.
- **Data Labeling**: For supervised learning, historical data will be labeled by domain experts as normal or malicious based on past incidents.
- **Goal Definition**: Specific objectives will be defined, such as detecting ransomware, Advanced Persistent Threats (APTs), or insider threats, which will influence the models and algorithms chosen in the next phase.

3.2.2 Modelling Phase

The modelling phase will involve building machine learning models that will predict and detect cyber threats based on the pre-processed data.

Model Evaluation Metrics

Key metrics are selected to evaluate the model's performance in threat detection, with an emphasis on balancing detection accuracy and operational feasibility:

- **Accuracy**: Overall accuracy is measured as the proportion of correctly classified instances. This metric provides a general measure of performance but can be misleading for imbalanced datasets.
- **Precision**: Indicates the proportion of true positives among all positive predictions, reducing the frequency of false alarms and minimizing alert fatigue among analysts.
- **Recall (Sensitivity)**: Measures the proportion of actual positives correctly identified, critical for minimizing missed threats and maximizing detection rates.
- **F1-Score**: A harmonic mean of precision and recall, providing a balanced evaluation for situations with an uneven class distribution.
- **Area Under the ROC Curve (AUC-ROC)**: AUC-ROC evaluates the model's ability to distinguish between true positives and false positives across different thresh-

olds. This is particularly useful for imbalanced datasets common in threat detection.

Model Selection

- **Model Selection**: Machine learning models such as decision trees, random forests, support vector machines (SVMs), or neural networks will be selected based on the data type and threat scenarios.
- **Feature Engineering**: Relevant features will be engineered to improve the model's ability to detect cyber threats. Time-series analysis can also be used for real-time threat detection in operational environments.
- **Model Training**: Models will be trained using supervised learning on labeled data or unsupervised learning on unlabelled data to detect anomalies. Cross-validation techniques will be used to avoid overfitting.
- **Threat Classification**: The trained model will be used to classify network activity or system behavior into predefined threat categories (e.g., malware, phishing, Denial of Service attacks).

Risk Assessment and Threat Probability

The probability of a cyber threat occurring will be evaluated using Bayesian Probability or Conditional Probability models. These tools estimate threat likelihood given past occurrences or indicators.

Bayesian Probability

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Where:

- $P(A|B)$ is the probability of threat A given evidence B.
- $P(B|A)$ is the probability of evidence B given that threat A has occurred.
- $P(A)$ is the prior probability of threat A.
- $P(B)$ is the total probability of evidence B.

Expected Risk (ER)

$$ER = P(T) \times C(T)$$

Where:

- $P(T)$ is the probability of a specific threat occurring.
- $C(T)$ is the cost impact of the threat.

Anomaly Detection Using Statistical Analysis

Anomalies in network behavior can indicate a possible cyber threat. Statistical techniques like mean and standard deviation are used to identify unusual behavior in a dataset.

Mean (μ)

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

Standard Deviation (σ)

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2}$$

Z-score for Anomaly Detection

$$Z = \frac{X - \mu}{\sigma}$$

Where Z-score measures how far a data point X is from the mean in terms of standard deviations, which can help identify outliers or anomalies.

Time Series Analysis for Anomaly Detection

Time-series models in CTI are used to detect temporal patterns or changes in network behavior over time.

Autoregressive Model (AR)

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

Where:

- X_t is the value at time t.
- ϕ represents model parameters.
- ϵ_t is white noise.

Moving Average (MA) Model

$$X_t = \mu + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$

Time-series modeling can help identify anomalies by looking for deviations from expected patterns.

Game Theory in Threat Intelligence

Game theory can model attacker-defender interactions, evaluating possible moves and responses in cybersecurity.

Payoff Matrix

Let A and D be attacker and defender strategies, respectively:

	D_1	D_2
A_1	P_{11}	P_{12}
A_2	P_{21}	P_{22}

Where:

- P_{ij} represents the payoff for choosing strategy A_i against D_j. Game theory helps in strategic planning and resource allocation for defense.

Markov Chains for Threat Propagation Analysis

Markov Chains model the likelihood of moving from one state (e.g., security level) to another, useful for understanding threat progression.

Markov Chain Transition Probability

$$P(X_{t+1} = s_j | X_t = s_i) = p_{ij}$$

Where p_{ij} is the probability of transitioning from state s_i to state s_j. This method is useful for predicting future system states based on current security conditions.

3.2.3 Post-Modelling Phase

After building the models, the post-modelling phase focuses on evaluation, validation, and deployment.

- **Model Validation**: Models are validated using test datasets. Key performance metrics such as accuracy, precision, recall, and F1 score are used to measure the model's effectiveness.
- **Model Tuning**: Hyperparameter tuning (e.g., grid search or random search) is used to optimize the model. This process involves adjusting key parameters to improve performance.
- **Deployment Readiness**: Once validated, the model is integrated into existing systems (e.g., Security Information and Event Management systems) for real-time threat detection. The model undergoes stress testing to ensure it can handle real-time data streams.

Test Environment Configuration

The experimental setup defines how the CTI models are tested in a controlled environment before full deployment. The CTI framework will be tested in a controlled environment replicating typical mining network configurations. The test environment includes:

- **Test Environment Setup**: A simulated or controlled mining network is created, including virtual machines, network traffic simulators, and log generators to mimic realistic scenarios.
- **Simulated Network Traffic**: Simulated mining network traffic, including both typical operational data and attack patterns (e.g., DDoS, malware, and insider threats), will be generated to assess the system's response.
- **Data Injection**: Historical data and simulated attack data are injected into the system to evaluate its ability to detect and respond to various threats.
- **Performance Measurement**: The system's performance is measured using metrics such as detection rate, false alarm rate, response time, resource usage (CPU, memory), and scalability under different load conditions. Resource usage (CPU, memory) is also monitored to ensure scalability.
- **Comparison with Existing Systems**: The CTI framework's performance is compared to existing cybersecurity systems in the mining industry to assess improvements.

Benchmarking and System Performance Metrics

- **Detection Rate**: Measures the framework's ability to identify known and unknown threats accurately.
- **False Positive Rate (FPR)**: Low FPR is essential to maintain operational efficiency by avoiding excessive false alarms.
- **Response Time**: Measures the time taken to detect and classify potential threats, crucial for real-time threat mitigation.
- **Resource Utilization**: CPU and memory usage are monitored during peak times to assess scalability and efficiency.

Optimization and Training Models

The optimization and training phase focuses on refining the models to maximize their performance and efficiency.

- **Hyperparameter Optimization**: Techniques such as grid search or Bayesian optimization are used to identify the best hyperparameters (e.g., learning rates, kernel functions) to improve detection accuracy.

- **Continuous Learning**: As new threats emerge, the model is retrained with updated data. Incremental learning or transfer learning methods may be used to update the model without complete retraining.
- **Efficiency Optimization**: The model is optimized for real-time performance through techniques such as model pruning, quantization, or the use of lightweight architectures.
- **Final Model Training**: The final optimized model is trained on the entire dataset to ensure robustness and accuracy. It is then ready for deployment into the CTI system.

Documentation and Reporting

To support effective usage and future updates, comprehensive documentation is maintained throughout the research:

- **Technical Documentation**: Detailed documentation of model architectures, data sources, and preprocessing steps ensures maintainability and scalability.
- **Evaluation Reports**: Performance metrics, validation results, and system performance summaries provide insights into model reliability and areas for improvement.
- **User Guidelines**: A manual for cybersecurity teams covers system monitoring, alert management, and updating procedures, ensuring practical use of the CTI framework.

3.3 Expected Outcomes

3.3.1 Prototype Framework

The first output of this research is the design of a **Cyber Threat Intelligence (CTI) framework** suitable for the mining industry. This framework will meet the special requirements of mining operations, such as remote locations, old equipment, and the risks associated with sensitive structures. The prototype framework will also incorporate real-time threat detection, advanced data processing, and machine learning to help identify and counter cyber threats more effectively. Some of the features will include the ability to interface with mining-related data sources like the SCADA systems, IoT devices, and network logs to make sure that the framework is both comprehensive and applicable to mining operations. Also, the framework will be developed in a manner that will make it flexible and able to be incorporated with other mining systems.

3.3.2 Contributions

This research is anticipated to produce the following significant contributions to the field of cybersecurity with a focus on the mining industry:

- **Improved Cybersecurity Resilience**: By developing a customized CTI framework, this research will enhance the ability of mining organizations to detect, analyze, and respond to cyber threats. This will reduce the risk of operational disruptions, minimize data breaches, and improve the overall security posture.
- **Broader Applicability to Critical Infrastructure**: While the framework is specifically designed for the mining sector, the methodologies and principles used to develop it can be applied to other critical infrastructure industries facing similar cybersecurity challenges, such as energy, healthcare, and transportation. This offers a foundation for cybersecurity improvements across multiple sectors.
- **Practical Implementation of CTI Systems**: The research will provide a clear roadmap for integrating CTI systems into industry-specific contexts, offering practical insights into how such systems can be deployed and maintained. This will address key challenges related to data privacy, regulatory compliance, and stakeholder engagement.
- **Contribution to Industry Standards**: The development of this CTI framework could inform the creation or refinement of cybersecurity standards and best practices within the mining industry, particularly concerning the integration of advanced threat detection technologies.

3.3.3 Limitations

The primary limitation of this research is its reliance on a specific set of data sources and scenarios, which may restrict the generalizability of findings across the varied operational environments within the mining industry. Although the study aims to develop a comprehensive Cyber Threat Intelligence (CTI) framework tailored to mining, it is constructed on data samples and hypothetical scenarios that might not reflect all real-world conditions in diverse mining settings. Differences in operational practices, network architectures, and security postures across various geographical and organizational contexts could affect the performance and adaptability of the CTI framework. Additionally, the complexity and diversity of cyber threat data can lead to varying levels of success in detecting advanced threats, depending on the data quality and volume accessible in specific mining operations. Furthermore, this study may face challenges in replicating the full range of possible cyber-attack vectors in a controlled environment, which may limit the robustness of real-world application and implementation. Lastly, the scalability of the CTI system may be impacted by the computational resources required to process and analyze vast data sources, posing an additional limitation to the system's practical deployment in certain resource-constrained mining operations.

3.4 Ethical Considerations

Ethical considerations are paramount in the design and implementation of a Cyber Threat Intelligence (CTI) framework, especially in a critical industry such as mining. One of the foremost concerns is ensuring data privacy and confidentiality, particularly in handling sensitive operational and employee-related data. Adherence to data

protection regulations such as the General Data Protection Regulation (GDPR) and sector-specific cybersecurity regulations is essential to protect individual privacy and uphold legal standards. The CTI system's capacity to monitor network traffic and detect anomalous behavior may inadvertently collect personally identifiable information (PII) or proprietary operational data, necessitating robust data anonymization and encryption measures.

Moreover, transparency and accountability in the use of machine learning algorithms for threat detection are critical to address biases or misclassifications that could lead to unintended consequences, such as falsely attributing benign activities as malicious or disproportionately scrutinizing certain behaviors. To mitigate such risks, regular audits and model validation steps are incorporated to maintain fairness and accuracy in detection algorithms. Additionally, ethical oversight will be applied to ensure that the CTI system does not lead to invasive surveillance or monitoring practices that infringe upon employees' rights. Finally, collaboration with industry stakeholders and adherence to ethical guidelines ensure that the CTI framework is aligned with both industry standards and the ethical principles of autonomy, transparency, and justice in cybersecurity.

Chapter 4

Schedule of Work

This chapter outlines the schedule of work for the development of a Cyber Threat Intelligence (CTI) framework tailored to the mining industry. The research is planned over a 10-month period, starting from February 1, 2025, to November 30, 2025. Each phase consists of specific tasks aimed at meeting the objectives of this research.

4.1 Overview

The research schedule is divided into six phases: Initial Planning and Literature Review, Pre-Modelling and Data Preparation, Modelling and Machine Learning, System Evaluation and Optimization, Continuous Learning and Model Refinement, and Documentation and Final Reporting. Each phase addresses key tasks needed to develop, test, and validate a CTI framework designed for the unique cybersecurity needs of the mining industry.

4.1.1 Phase 1: Requirements Analysis and Literature Review

The initial phase involves a comprehensive literature review and requirements analysis to understand the existing cybersecurity landscape within the mining industry. Key activities include:

- **Stakeholder Engagement:** Conduct interviews and surveys with stakeholders, including cybersecurity professionals, to identify specific threats and security gaps in the mining sector.
- **Literature Review:** Examine existing CTI frameworks, cybersecurity models, and machine learning techniques relevant to critical infrastructure, particularly in industrial settings.
- **Gap Analysis:** Identify limitations in current CTI solutions and the specific needs of mining operations that require a tailored approach.

4.1.2 Phase 2: Data Collection and Preprocessing

This phase focuses on gathering and preparing data necessary for developing and testing the CTI framework. Key activities include:

- **Data Sources:** Collect data from multiple sources, including network logs, OSINT (Open Source Intelligence) feeds, vendor-provided threat reports, and historical attack data in the mining industry.
- **Data Preprocessing:** Clean and transform the data to ensure quality and consistency. This involves handling missing values, normalizing data formats, and applying feature extraction techniques such as Principal Component Analysis (PCA) for noise reduction.
- **Data Labeling:** Label data samples with assistance from cybersecurity experts to distinguish between benign and malicious activities, enhancing model training accuracy.

4.1.3 Phase 3: Framework Development and Modelling

In this phase, the CTI framework's architecture is designed and machine learning models are developed to detect, classify, and predict cyber threats in the mining sector. Key activities include:

- **Framework Architecture Design:** Create an architecture blueprint that integrates data collection, preprocessing, analysis, and decision-support components of the CTI system.
- **Model Selection:** Select machine learning algorithms suitable for anomaly detection and threat classification, such as decision trees, support vector machines (SVM), or neural networks.
- **Model Training:** Train and validate models using supervised and unsupervised learning techniques on labeled data to maximize threat detection accuracy.
- **Feature Engineering:** Conduct feature engineering and exploratory data analysis to identify the most relevant indicators of cyber threats.

4.1.4 Phase 4: System Evaluation and Optimization

Once the CTI framework is developed, it undergoes rigorous testing and evaluation to ensure performance and reliability. Key activities include:

- **Model Validation:** Validate the model's performance on test datasets, evaluating metrics such as accuracy, precision, recall, and F1-score.
- **System Testing:** Conduct stress tests and simulate real-world threat scenarios in a controlled environment to evaluate the system's responsiveness and robustness.

- **Hyperparameter Tuning:** Optimize model performance through hyperparameter tuning, such as grid search or random search, to enhance detection accuracy and reduce false positives.
- **Scalability Assessment:** Assess the framework's scalability to ensure it can handle large volumes of data generated in mining operations.

4.1.5 Phase 5: Ethical Considerations and Compliance

This phase involves addressing the ethical and regulatory aspects of deploying a CTI system in the mining sector, focusing on data privacy and operational transparency. Key activities include:

- **Privacy Safeguards:** Implement data anonymization and encryption techniques to protect sensitive information and ensure compliance with data protection regulations.
- **Bias and Fairness Analysis:** Regularly audit machine learning models to detect and mitigate biases, ensuring fair and accurate threat detection across various operational scenarios.
- **Ethics Approval:** Seek ethical approval from relevant bodies and consult industry experts to ensure that the framework aligns with both legal standards and industry guidelines.

4.1.6 Phase 6: Documentation and Final Reporting

The final phase documents the research findings and provides guidelines for practical implementation of the CTI framework in the mining industry. Key activities include:

- **Results Documentation:** Summarize findings, including insights into CTI framework performance, challenges, and opportunities for improvement.
- **Recommendations for Implementation:** Provide recommendations for deploying the CTI framework in operational settings within the mining sector.
- **Final Report Preparation:** Compile and submit a comprehensive research report, detailing methodologies, results, and suggestions for future research directions.

Table 4.1: Research Schedule from February 2025 to November 2025

Phase	Activity Description	Timeline
Phase 1: Initial Planning and Literature Review		February - March 2025

Continued on next page

Phase	Activity Description	Timeline
Requirements Analysis and Stakeholder Engagement	Conduct initial meetings with stakeholders to identify specific cybersecurity needs and define the project scope.	Feb - Mar 2025
Literature Review	Review relevant literature on CTI frameworks, machine learning models, and cybersecurity within the mining sector to build a theoretical foundation.	Mar 2025
Ethical and Regulatory Review	Review data privacy regulations and ethical standards for cybersecurity in mining to ensure compliance.	Mar 2025
Phase 2: Pre-Modelling and Data Preparation		April - May 2025
Data Collection	Gather data from multiple sources, including network logs, OSINT feeds, and vendor-specific threat intelligence reports.	Apr 2025
Data Preprocessing	Clean, transform, and normalize collected data, applying techniques such as PCA for noise reduction.	Apr - May 2025
Goal Definition and Data Labeling	Define objectives for CTI (e.g., ransomware detection) and label historical data as normal or malicious based on prior incidents.	May 2025
Phase 3: Modelling and Machine Learning		June - August 2025
Model Selection and Initial Testing	Choose initial machine learning models and test them on sample data to assess suitability for mining-specific cybersecurity needs.	Jun 2025
Feature Engineering	Optimize feature sets for threat detection based on exploratory data analysis.	Jun - Jul 2025
Model Training and Evaluation	Train models using supervised and unsupervised learning techniques, evaluating performance with metrics such as accuracy, precision, and recall.	Jul - Aug 2025
Threat Classification	Classify network activity and system behavior into threat categories, such as malware or phishing attacks.	Aug 2025
Phase 4: System Evaluation and Optimization		September - October 2025
Model Validation	Validate models using test datasets, fine-tuning hyperparameters to maximize performance.	Sep 2025
Deployment Readiness	Prepare the validated model for deployment in a simulated mining environment and conduct stress testing.	Sep 2025

Continued on next page

Phase	Activity Description	Timeline
Experimental Setup	Establish a controlled testing environment, including data injection and monitoring for system performance assessment.	Oct 2025
Phase 5: Continuous Learning and Model Refinement		October 2025
Hyperparameter Optimization	Refine hyperparameters using techniques such as grid search or Bayesian optimization to improve model accuracy.	Oct 2025
Continuous Learning	Implement incremental learning methods to keep the model updated with new data and emerging threats.	Oct 2025
Phase 6: Documentation and Final Reporting		November 2025
Results Documentation	Summarize findings, including insights on CTI framework performance and areas for improvement.	Nov 2025
Recommendations and Practical Implementation	Provide guidelines for practical implementation of the CTI framework in the mining industry.	Nov 2025
Final Report Preparation	Compile and submit the final research report, detailing methodologies, results, and recommendations for future research.	Nov 2025

4.2 Detailed Timeline by Month

- **February - March 2025:** Requirements analysis, literature review, and ethical/regulatory assessment.
- **April - May 2025:** Data collection, preprocessing, and labeling for specific threat categories.
- **June - August 2025:** Selection, testing, and training of machine learning models, along with threat classification.
- **September - October 2025:** Model validation, deployment preparation, and experimental setup.
- **October 2025:** Hyperparameter optimization and continuous learning integration.
- **November 2025:** Documentation of findings, preparation of final report, and recommendations for implementation.

Chapter 5

Conclusion

The CTI System for the Mining Industry aims to bridge the gap between the mining industry's unique operational demands and the growing need for robust cybersecurity measures. By developing a tailored CTI System, the study seeks to provide an architecture design of a CTI System. The development of the CTI System will enhance the ability to detect and respond to cyber threats proactively in the mining industry. The CTI System's integration of advanced data analytics and machine learning will offer more precise and real-time action against evolving threats. Additionally, this research will address ethical and regulatory concerns and ensure the proposed solution is effective and compliant with mining industry standards. The findings from this research will contribute significantly to both the mining industry and the broader field of cybersecurity. By providing a specialized system for cyber threat intelligence, this study will help safeguard critical mining operations, protecting both assets and the environment. The lessons learned and the methodologies developed can serve as a model for other industries facing similar cybersecurity challenges, marking a step forward in the ongoing effort to secure vital industrial infrastructure.

Future Directions

While this research marks an important step in securing mining operations, there are several key areas for future research and development that will further enhance the capabilities of CTI systems in the mining industry.

- **Integration with Emerging Technologies:** Future work could explore the integration of **blockchain** and **quantum computing** to improve data integrity and secure threat intelligence sharing. Blockchain could enable tamper-proof threat data exchange, while quantum computing may present new challenges and opportunities for encryption and threat detection.
- **Advanced Threat Modeling:** As cyber threats evolve, there is a need for **dynamic and adaptive threat models**. Future research could focus on enhancing machine learning algorithms to predict complex, multi-stage cyberattacks and automate responses in real time.

- **Interoperability Across Industries:** Given the convergence of industrial technologies, future CTI frameworks could benefit from **cross-sector interoperability**, allowing mining companies to share threat intelligence with other critical sectors, such as energy and manufacturing, to improve collaborative defense.
- **Scalability in Remote Operations:** As mining operations expand in remote locations, research could focus on designing **scalable, decentralized CTI systems** that work effectively with low-bandwidth environments and limited access to centralized infrastructure.
- **Human-Centric Security:** The **human factor** in cybersecurity cannot be overlooked. Future research could explore **training and awareness programs** for miners and cybersecurity personnel, ensuring they can effectively interpret and act on CTI insights.

Appendix A

Extra Stuff

A.1 What is an appendix?

An appendix is useful when there is information that you need to include, but breaks the flow of your document, e.g. a large number of figures/tables may need to be shown, but maybe only one needs to be in the text and the rest are just included for completeness.

References

[Barnum 2012] Sean Barnum. Standardizing cyber threat intelligence information with the structured threat information expression (stix). *Mitre Corporation*, 11:1–22, 2012.

[bin Mohd Aziz 2024] Abdullah bin Mohd Aziz. Maximizing cyber threat intelligence (cti) in the financial sector: Benefits and implementation challenges. *Quarterly Journal of Emerging Technologies and Innovations*, 9(3):15–36, 2024.

[Böhm *et al.* 2018] Fabian Böhm, Florian Menges, and Günther Pernul. Graph-based visual analytics for cyber threat intelligence. *Cybersecurity*, 1(1):16, 2018.

[Brooks 2024] Chuck Brooks. *Inside Cyber: How AI, 5G, IoT, and Quantum Computing Will Transform Privacy and Our Security*. John Wiley & Sons, 2024.

[Bugiardini *et al.* 2016] Raffaele Bugiardini, Lina Badimon, et al. The international survey of acute coronary syndromes in transitional countries (isacs-tc): 2010–2015. *International Journal of Cardiology*, 217:S1–S6, 2016.

[Caltagirone *et al.* 2013] Sergio Caltagirone, Andrew Pendergast, and Christopher Betz. The diamond model of intrusion analysis. *Threat Connect*, 298(0704):1–61, 2013.

[Cameron 2024] Michael Cameron. Sibanye cyber-attack lays risks bare. *Australia's Paydirt*, 1(330):16–17, 2024.

[Deliu *et al.* 2017] Isuf Deliu, Carl Leichter, and Katrin Franke. Extracting cyber threat intelligence from hacker forums: Support vector machines versus convolutional neural networks. In *2017 IEEE International Conference on Big Data (Big Data)*, pages 3648–3656. IEEE, 2017.

[Gately 2023] Hannah Gately. *Russian organised crime and Ransomware as a Service: state cultivated cybercrime*. PhD thesis, Macquarie University, 2023.

[Georgiadou *et al.* 2021] Anna Georgiadou, Spiros Mouzakitis, and Dimitris Askounis. Assessing mitre att&ck risk using a cyber-security culture framework. *Sensors*, 21(9):3267, 2021.

[Gong and Lee 2021] Seonghyeon Gong and Changhoon Lee. Cyber threat intelligence framework for incident response in an energy cloud platform. *Electronics*, 10(3):239, 2021.

[Islam *et al.* 2022] Chadni Islam, M Ali Babar, Roland Croft, and Helge Janicke. Smartvalidator: A framework for automatic identification and classification of cyber threat data. *Journal of Network and Computer Applications*, 202:103370, 2022.

[Janoti *et al.*] Nikhlesh Singh Janoti, Rida Rohan, and Neerja Negi. Strategic perspectives on cyber threat intelligence: A comprehensive analysis.

[Kayode-Ajala 2023] Olaolu Kayode-Ajala. Applications of cyber threat intelligence (cti) in financial institutions and challenges in its adoption. *Applied Research in Artificial Intelligence and Cloud Computing*, 6(8):1–21, 2023.

[Kim *et al.* 2016] Daegeon Kim, JiYoung Woo, and Huy Kang Kim. "i know what you did before": General framework for correlation analysis of cyber threat incidents. In *MILCOM 2016-2016 IEEE Military Communications Conference*, pages 782–787. IEEE, 2016.

[Klein and Celik 2017] R. Klein and T. Celik. The Wits Intelligent Teaching System: Detecting student engagement during lectures using Convolutional Neural Networks. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 2856–2860, Sep. 2017.

[Kotsias *et al.* 2023] James Kotsias, Atif Ahmad, and Rens Scheepers. Adopting and integrating cyber-threat intelligence in a commercial organisation. *European Journal of Information Systems*, 32(1):35–51, 2023.

[Krauss and Papesh 2022] Oliver Krauss and Konstantin Papesh. Analysis of threat intelligence information exchange via the stix standard. In *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, pages 1–6. IEEE, 2022.

[Lee and Shon 2016] Seokcheol Lee and Taeshik Shon. Open source intelligence base cyber threat inspection framework for critical infrastructures. In *2016 Future Technologies Conference (FTC)*, pages 1030–1033. IEEE, 2016.

[Lenka *et al.* 2023] Arpita Lenka, Madhurima Goswami, Harmandeep Singh, and Harsha Baskaran. Cybersecurity disclosure and corporate reputation: Rising popularity of cybersecurity in the business world. In *Effective Cybersecurity Operations for Enterprise-Wide Systems*, pages 169–183. IGI Global, 2023.

[Limited 2024] Evolution Mining Limited. *Cyber Security Incident*. https://evolutionmining.com.au/storage/2024/08/2759355-Cyber-Security-Incident.pdf, August 2024. Accessed: 2024-12-09.

[Liu *et al.* 2022] Jian Liu, Junjie Yan, Jun Jiang, Yitong He, Xuren Wang, Zhengwei Jiang, Peian Yang, and Ning Li. Tricti: an actionable cyber threat intelligence discovery system via trigger-enhanced neural network. *Cybersecurity*, 5(1):8, 2022.

[Medoh and Telukdarie 2022] Chuks Medoh and Arnesh Telukdarie. The future of cybersecurity: a system dynamics approach. *Procedia Computer Science*, 200:318–326, 2022.

[Naik *et al.* 2022] Nitin Naik, Paul Jenkins, Paul Grace, and Jingping Song. Comparing attack models for it systems: Lockheed martin's cyber kill chain, mitre att&ck framework and diamond model. In *2022 IEEE International Symposium on Systems Engineering (ISSE)*, pages 1–7. IEEE, 2022.

[Naseer 2023] Iqra Naseer. Machine learning applications in cyber threat intelligence: A comprehensive review. *The Asian Bulletin of Big Data Management*, 3(2):190–200, 2023.

[Noor *et al.* 2019] Umara Noor, Zahid Anwar, Tehmina Amjad, and Kim-Kwang Raymond Choo. A machine learning-based fintech cyber threat attribution framework using high-level indicators of compromise. *Future Generation Computer Systems*, 96:227–242, 2019.

[Papastergiou *et al.* 2021] Spyridon Papastergiou, Haralambos Mouratidis, and Eleni-Maria Kalogeraki. Handling of advanced persistent threats and complex incidents in healthcare, transportation and energy ict infrastructures. *Evolving Systems*, 12(1):91–108, 2021.

[Prakash *et al.* 2022] Ravi Prakash, VS Anoop, and S Asharaf. Blockchain technology for cybersecurity: A text mining literature analysis. *International Journal of Information Management Data Insights*, 2(2):100112, 2022.

[Provatas *et al.* 2023] Kimonas Provatas, Ioannis Tzannetos, and Vassilios Vescoukis. Standards-based cyber threat intelligence sharing using private blockchains. In *2023 18th Conference on Computer Science and Intelligence Systems (FedCSIS)*, pages 649–656. IEEE, 2023.

[Ravichandran *et al.* 2024] Naveesen Ravichandran, Tahrhunraj Tewaraja, Vishendraa Rajasegaran, Sri Sharvesh Kumar, Siva Kumar Livekha Gunasekar, and Siva Raja Sindiramutty. Comprehensive review analysis and countermeasures for cybersecurity threats: Ddos, ransomware, and trojan horse attacks. 2024.

[Ryandy *et al.* 2020] Ryandy, Charles Lim, and Kalpin Erlangga Silaen. Xt-pot: Exposing threat category of honeypot-based attacks. In *Proceedings of the 2020 International Conference on Engineering and Information Technology for Sustainable Industry*, pages 1–6, 2020.

[Sajid *et al.* 2016] Anam Sajid, Haider Abbas, and Kashif Saleem. Cloud-assisted iot-based scada systems security: A review of the state of the art and future challenges. *Ieee Access*, 4:1375–1384, 2016.

[Shahi 2018] Mohammad Ashraful Huq Shahi. Tactics, techniques and procedures (ttps) to augment cyber threat intelligence (cti): A comprehensive study. 2018.

[Song *et al.* 2022] Binghua Song, Rong Chen, Baoxu Liu, Zhengwei Jiang, and Xuren Wang. Time series attention based transformer neural turing machines for diachronic graph embedding in cyber threat intelligence. In *International Conference on Computational Science*, pages 17–30. Springer, 2022.

[Stafford 2020] V Stafford. Zero trust architecture. *NIST special publication*, 800:207, 2020.

[Tundis *et al.* 2022] Andrea Tundis, Samuel Ruppert, and Max Mühlhäuser. A feature-driven method for automating the assessment of osint cyber threat sources. *Computers & Security*, 113:102576, 2022.

[Wang and Lu 2013] Wenye Wang and Zhuo Lu. Cyber security in the smart grid: Survey and challenges. *Computer networks*, 57(5):1344–1371, 2013.

[Webb *et al.* 2014] Jeb Webb, Sean Maynard, Atif Ahmad, Graeme Shanks, et al. Information security risk management: An intelligence-driven approach. *Australasian Journal of Information Systems*, 18(3), 2014.

[Yu *et al.* 2023] Zhongkun Yu, JunFeng Wang, BinHui Tang, and Li Lu. Tactics and techniques classification in cyber threat intelligence. *The Computer Journal*, 66(8):1870–1881, 2023.