Assessing and Ranking Vulnerabilities in Industrial Control Systems

A Structured Database Approach to Device Security Analysis

Matthew B. Collins

A report submitted as part of the requirements for the degree
of Master of Science in Computer Science
School of Science and Mathematics
Department of Cyber and Computer Sciences
The Citadel Military College of South Carolina

May 2024

Supervisor Shankar Banik, Ph.D.

Abstract

In the landscape of modern industrial operations, Industrial Control Systems (ICS) are crucial to the overall success and efficiency of the Operations Technology (OT) networks. The inherent high value of ICS devices makes them particularly desirable targets for attackers. These systems comprise both low-level devices like Programmable Logic Controllers (PLCs), which are inherently challenging to secure, and traditional Information Technology (IT) devices that often exhibit a range of security weaknesses. This thesis introduces a novel structured database approach to systematically analyze and rank these security challenges in ICS. By extracting vulnerability advisories, associating them with specific Common Vulnerabilities and Exposures (CVEs), and further linking these to Common Platform Enumerations (CPEs), the study constructs a relational database for comprehensive security analysis. This framework facilitates the quantitative evaluation and ranking of ICS vendors and their products based on their exposure to vulnerabilities. The approach, emphasizing automated data collection and processing, ensures that the assessment remains both current and scalable. While the initial findings identify pivotal trends in security issue distributions and underscore particularly susceptible areas within ICS, they primarily serve as a preliminary exploration into crown jewel vulnerabilities. This work not only opens new avenues for deeper investigation but also provides a foundational strategy for enhancing the cybersecurity resilience of industrial systems. The insights gained lay the groundwork for future researchers to further explore and mitigate these critical security issues.

Keywords— Industrial Control Systems (ICS), Operations Technology (OT), Common Vulnerabilities and Exposures (CVE), Common Platform Enumerations (CPE)

Acknowledgements

I am profoundly grateful to my Creator for placing me in an era where technological advancements occur at an unprecedented pace. This unique period in history has not only fueled my curiosity but also provided the fertile ground for my academic and professional growth.

My heartfelt thanks go to my wife, whose patience and support have been my anchor throughout this journey. Her unwavering belief in my dreams and her sacrifices have not gone unnoticed, and they have been critical in my pursuit of academic and professional excellence. Her strength and encouragement have made all the difference, allowing me to pursue my passion for learning and discovery.

I am also immensely thankful for the guidance and mentorship of my advisor, Dr. Banik. His exemplary dedication to hard work and leadership has not only shaped my research approach but also deeply influenced my personal development. Additionally, I owe a tremendous debt of gratitude to all my advisors and the computer science faculty. Their vast knowledge and rigorous academic standards have profoundly enriched my education. Each lecture and interaction has been a building block in the foundation of my understanding and skills in engineering and computer science.

Finally, I must express my sincere appreciation to my research partner, Richard Owing. His commitment, diligence, and camaraderie throughout this research have not only enhanced our work but made the process thoroughly enjoyable. His friendship and professionalism have greatly contributed to the success of our collaborative efforts.

To all who have guided, supported, and inspired me, thank you. This thesis not only stands as a testament to my academic efforts but also as a tribute to your generous contributions to my life and studies.

Faculty Approval

Signed .... Date ....
Shankar Banik, Ph.D.

Signed .... Date ....
Jacob Benjamin, PhD

Signed .... Date ....
William Johnson, PhD

Signed .... Date ....
Mohamed Baza, PhD

Declaration

I confirm that the work contained in this MSc project report has been composed solely by myself and has not been accepted in any previous application for a degree. All sources of information have been specifically acknowledged and all verbatim extracts are distinguished by quotation marks.

Signed .... Date ....
Matthew B. Collins

Contents

List of Tables

List of Figures

List of Algorithms

1 Introduction

The intersection of cybersecurity and industrial operations has become increasingly crucial as Industrial Control Systems (ICS) become more intelligent and networked. These systems, pivotal in Operational Technology (OT) networks, manage essential services such as electricity, water treatment, and public transportation. The effectiveness and safety of these services hinge on the resilience of ICS devices to defend against advanced cyber threats—threats that were traditionally reserved for sophisticated Information Technology (IT) networks.

The integration of advanced IT components with traditional industrial mechanisms has not only expanded the capabilities of ICS but also introduced significant vulnerabilities. As the attack surface broadens, the potential impacts of cybersecurity breaches in these systems not only pose risks to financial stability but also threaten human safety and national security. This duality of critical importance and inherent vulnerability makes ICS a prime target for cyber-attacks.

The urgency to secure ICS components is underscored by their operational necessity and the escalating sophistication of the threats they face. Traditional methods of securing these systems often fall short due to the unique and complex nature of ICS architectures, which include a mix of both outdated and modern interfaces. This complexity necessitates a strategic approach to understanding and mitigating vulnerabilities that can be systematically applied and scaled across diverse ICS environments.

This thesis addresses the pressing need for a structured methodology to assess and prioritize vulnerabilities in OT networks. By developing a relational database that captures and analyzes vulnerability advisories linked to specific Common Vulnerabilities and Exposures (CVEs) and Common Platform Enumerations (CPEs), this research provides a granular and actionable understanding of vulnerabilities. The approach leverages automated data collection and robust analytical frameworks to maintain an ongoing and dynamic assessment of vulnerabilities, enabling stakeholders to implement timely and effective security measures.

The primary goal of this research is to provide a data-driven approach for enhancing ICS security. By developing analytics that can be leveraged by engineers and researchers, this thesis aims to improve the understanding and management of ICS vulnerabilities. This focus on practical, actionable intelligence supports ongoing efforts to secure ICS environments against current and future cyber threats, directly contributing to the resilience and reliability of critical infrastructure operations.

2 Background

This chapter provides some background research on the project and examines some previous work.

2.1 The Purdue Model

2.1.1 Origin and Purpose

The Purdue Model was developed in the 1990s as part of a project led by Theodore J. Williams and other researchers for the Purdue University Consortium for Computer Integrated Manufacturing. It was initially intended to guide the integration of enterprise IT systems with manufacturing operations. Over time, it evolved to become a foundational model in understanding and securing industrial control systems [1].

2.1.2 Model Overview

The Purdue Model divides industrial control and IT systems into several hierarchical layers, each designated for specific types of processes and data handling within an industrial enterprise. The model is generally illustrated as a series of levels from 0 to 5, where each level represents a distinct function in the manufacturing and control process:

- **Level 0: Physical Process Layer** - This layer includes the actual physical processes involved in production. It consists of machinery and equipment that directly interact with the production operations, such as sensors and actuators.
- **Level 1: Control Layer** - This layer manages the real-time control of industrial processes through devices like Programmable Logic Controllers (PLCs) and Remote Terminal Units (RTUs).
- **Level 2: Supervisory Control Layer** - It involves the systems that oversee the control devices, such as Human-Machine Interfaces (HMIs) and Supervisory Control and Data Acquisition (SCADA) systems that provide operators with data visualizations and control capabilities.
- **Level 3: Manufacturing Operations Systems** - This layer handles operational management over production, including scheduling, batch processing, and manufacturing execution systems (MES). It acts as a bridge between the plant floor and enterprise systems.

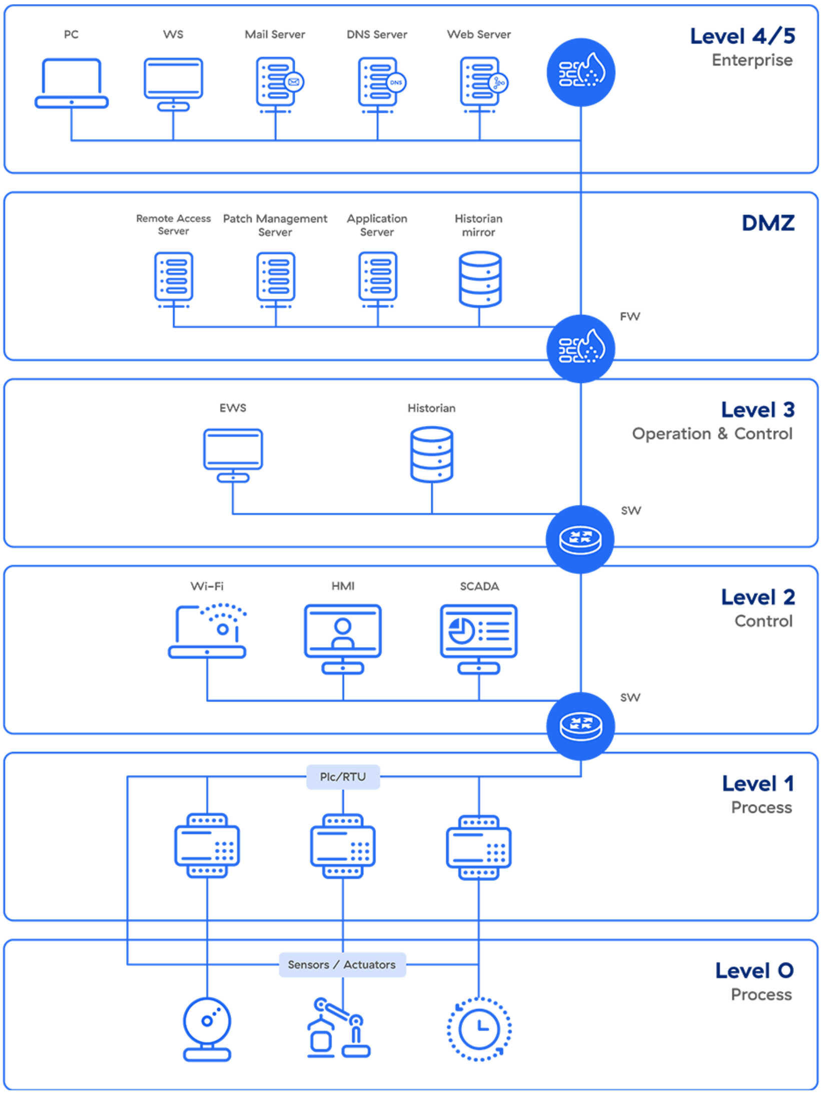


Figure 2.1: Purdue Model Diagram[2]

- **Level 4: Enterprise Management** - At this level, the focus shifts to business logistics and management systems, such as Enterprise Resource Planning (ERP) systems, which integrate various facets of an enterprise's operations.
- **Level 5: Business Planning and Logistics** - This top layer involves the corporate network and services, handling business planning, logistics, and other non-operational tasks.

2.1.3 Significance in Cybersecurity

The Purdue Model has become particularly valuable in cybersecurity for OT networks. By delineating clear demarcation lines between different network levels, the model helps in implementing appropriate security measures at each layer. This separation is crucial for protecting more vulnerable and critical lower layers from potential cyber threats that might infiltrate higher IT layers. For instance, it recommends the use of firewalls and DMZs (demilitarized zones) between levels to control and monitor data flow and to prevent unauthorized access [3].

2.1.4 Current Relevance

Today, the Purdue Model is instrumental in designing secure industrial environments, especially as the convergence of IT and OT systems increases with advancements in IoT and Industry 4.0 technologies. It helps organizations conceptualize and structure their control systems architecture to better manage risks, compliance, and cybersecurity strategies effectively. This framework has continued to influence how cybersecurity professionals approach the unique environments of OT and ICS, providing a structured way to address the complexities of modern industrial operations. As such, it remains a cornerstone in the study and implementation of cybersecurity measures within various sectors, including manufacturing, energy, and utilities [4].

2.2 Industrial Control Systems Advisory

2.2.1 Cybersecurity and Infrastructure Security Agency

The Cybersecurity and Infrastructure Security Agency (CISA) is a federal agency under the United States Department of Homeland Security (DHS). Established to enhance the cybersecurity and infrastructure resilience of the nation, CISA coordinates security efforts across critical infrastructure sectors, providing guidance, resources, and direct support to mitigate against evolving threats.

2.2.2 ICSA Background

CISA issues Industrial Control Systems Advisories (ICSA) as a key component of its strategy to protect the nation's critical infrastructure. These advisories are vital for organizations that operate industrial control systems, offering crucial information on vulnerabilities that might compromise these essential systems.

2.2.3 Purpose and Function

The purpose of ICSA is to alert, guide, and educate ICS operators and the broader cybersecurity community. By notifying stakeholders of vulnerabilities and potential cyber threats, providing detailed mitigation strategies, and discussing current cyber threat tactics, ICSA plays a critical role in the ongoing defense of critical infrastructure.

2.2.4 Content and Structure

ICSA typically includes detailed descriptions of vulnerabilities, complete with severity ratings using the Common Vulnerability Scoring System (CVSS). Each advisory identifies specific vulnerabilities by their Common Vulnerabilities and Exposures (CVE) identifiers, a system used to catalog publicly known information-security vulnerabilities and exposures in a standardized way. The advisories also list affected devices, software, or systems using Common Platform Enumerations (CPEs), which provide a standardized method to classify these elements across the IT industry [5]. Based on this information, the advisories outline recommended actions to mitigate risks, which may include applying patches, adjusting configurations, or enhancing network and access controls.

2.2.5 Impact and Importance

ICSA are indispensable for the security management of industrial control systems. They enable organizations to stay informed about new vulnerabilities, react promptly to security incidents, and standardize security practices across various sectors. This standardized approach is crucial for maintaining the safety and operational continuity of critical infrastructures.

2.2.6 Challenges and Considerations

Implementing ICSA recommendations can be challenging due to the high frequency of advisories and the complexity of the recommended measures. Smaller enterprises, in particular, may struggle with the resources required to keep pace with the advisories and implement the necessary security measures effectively.

2.3 Common Vulnerabilities and Exposures

Common Vulnerabilities and Exposures (CVE) is a catalog of publicly disclosed information security vulnerabilities and exposures, providing a standardized reference for identifying and discussing these issues within the cybersecurity community.

CVE aims to facilitate the exchange of information across different cybersecurity tools and services by standardizing the identification of vulnerabilities. Each CVE identifier ensures that all stakeholders refer to vulnerabilities in a consistent manner, enhancing clarity in communications across diverse systems and platforms.

2.3.1 Content and Structure

Each entry in the CVE database includes a unique CVE Identifier, a concise description of the security vulnerability, and references to detailed resources about the vulnerability, such as security advisories and patches. Additionally, most CVE entries are linked with corresponding Common Weakness Enumerations (CWEs), which describe the underlying software security weaknesses, and Common Platform Enumerations (CPEs), which specify the affected systems or components. This structured information helps in accurately identifying the nature of the vulnerability and the impacted environments.

2.3.2 MITRE Corporation

The CVE Program is managed by the CVE Program Secretariat, currently operated by the MITRE Corporation under the stewardship of the U.S. Department of Homeland Security. This arrangement ensures that CVE remains a vital resource for various cybersecurity measures, including vulnerability management and mitigation tools.

2.3.3 Significance in Cybersecurity

CVEs are vital for a unified approach to handling cybersecurity threats. They help organizations assess risks, coordinate responses to security advisories, and bolster overall security measures by providing detailed, standardized information about each vulnerability. This standardization is critical for effective communication and operational consistency in cyber defense efforts.

2.3.4 Challenges and Considerations

Despite their utility, CVEs face challenges such as incomplete coverage of all known vulnerabilities and delays in assignment, which can impede timely responses to emerging threats. Additionally, the severity indicated in a CVE does not necessarily reflect the potential impact on all systems, which can lead to misprioritization of threat management resources.

2.4 Common Platform Enumerations

Common Platform Enumerations (CPE) is a structured naming scheme for information technology systems, software, and packages. Maintained by the National Institute of Standards and Technology (NIST), CPE provides a standard format for identifying classes of applications, operating systems, and hardware devices within an IT environment. This universal identification is critical in managing security across diverse systems.

2.4.1 Purpose and Function

The primary purpose of CPE is to facilitate the clear and consistent naming of IT platforms in cybersecurity communications. By standardizing how products are identified, CPE enables cybersecurity professionals and tools to uniformly recognize and manage the security of various IT products. This uniformity is crucial for effective vulnerability management, security research, and regulatory compliance.

2.4.2 Content and Structure

CPE entries are formatted in a standardized way, including a unique identifier that makes referencing specific platforms straightforward. Each entry typically includes [6]:

- **Part** - Denoting the type of item (e.g., operating system, application, or hardware)
- **Vendor** - The creator or distributor of the product
- **Product** - The specific name of the product
- **Version** - The version of the product
- **Update** - Any updates to the product
- **Edition** - Specific editions if applicable
- **Language** - Valid language tag such as en-us for US English
- **SW_Edition** - This attribute characterizes how the product is tailored to a particular market or class of end users. It uses values selected from an attribute-specific valid-values list, which may be defined by specifications utilizing this specification. Any character string that meets the requirements for Well-formed Names (WFNs) can be specified as the value of this attribute.
- **Target_SW** - This describes the software computing environment within which

the product operates, such as a specific operating system or platform. Values should be selected from a defined list that matches the specifications in use, and any string meeting WFN requirements may be used.

- **Target_HW** - This attribute characterizes the instruction set architecture (e.g., x86) on which the product runs. Bytecode-intermediate languages, such as Java bytecode for the Java Virtual Machine or Microsoft Common Intermediate Language for the Common Language Runtime virtual machine, are also considered as instruction set architectures under this attribute. Values should come from an attribute-specific valid-values list.
- **Other** - Captures any other general descriptive or identifying information that is specific to the vendor or product and does not logically fit into any other attribute.

2.4.3 National Institute of Standards and Technology

CPE is part of the Security Content Automation Protocol (SCAP), a suite of standards supported by the National Institute of Standards and Technology (NIST) to facilitate automated vulnerability management, measurement, and policy compliance evaluation [7]. NIST plays a crucial role in the development, maintenance, and dissemination of CPE, ensuring that it aligns with broader cybersecurity frameworks and standards.

The NIST National Vulnerability Database (NVD) is an integral component of this ecosystem. It acts as a comprehensive repository for both CVE and CPE data, providing detailed information about security vulnerabilities and the products they affect. The NVD enhances the utility of CPE and CVE by adding depth with analysis, impact scores based on the Common Vulnerability Scoring System (CVSS), and links to corresponding CPE entries. This connection ensures that each CVE entry is tied to a standardized identifier for any affected platforms, facilitating a seamless approach to vulnerability management.

By managing the CPE specification and hosting the NVD, NIST provides a stable and scalable method to catalog and disseminate information about IT products that is universally accessible and useful for cybersecurity defense. This ongoing management involves regularly updating the CPE dictionary and NVD to accommodate new technology products and versions, reflecting the dynamic nature of the IT industry. NIST's involvement guarantees that the CPE standards and CVE entries remain robust and relevant, facilitating their integration into global cybersecurity practices.

2.4.4 Significance in Cybersecurity

CPEs are essential for accurately assessing and addressing vulnerabilities that affect various platforms. They allow cybersecurity tools and professionals to quickly determine which specific systems are impacted by a vulnerability, facilitating targeted and efficient security measures. This capability is vital in a landscape where timely and precise information can significantly influence security outcomes.

2.4.5 Challenges and Considerations

IT product versions and configurations, facilitating broader vulnerability management. Wildcards, represented by an asterisk (*), are essential for denoting unknown or variable aspects of a product's attributes, such as versions or updates that are either not fully known or too numerous to list individually [6]. This approach is particularly valuable when a security vulnerability affects multiple versions of a product, allowing for a generalized reference that encompasses all potentially impacted versions.

However, the necessity of wildcards introduces complexities that must be carefully managed. Achieving precision and accuracy in security reporting becomes more challenging with the use of wildcards. Maintaining consistency in how wildcards are applied across various databases and security systems presents a significant challenge. Inconsistencies in wildcard usage can lead to discrepancies in vulnerability assessments, undermining the effectiveness of security tools that depend on accurate CPE data for scanning and mitigation.

3 Literature Review

Throughout the research of ranking vulnerabilities across ICS devices and vendors, not much was found. Therefore, the literature review extends to how other researchers have done similar types of analyses.

3.0.1 Internet-facing ICS Devices

In the research paper, From Exposed to Exploited: Drawing the Picture of Industrial Control Systems Security Status in the Internet Age, Yixiong Wu, Jianwei Zhuge, and their team explore the vulnerabilities of Internet-facing industrial control systems (ICS) through a novel passive vulnerability assessment system called ICScope [8]. This system aims to assess and characterize vulnerabilities in a non-intrusive manner, which is crucial for environments where active scanning could disrupt operational stability. By utilizing data aggregated from various device search engines, ICScope identifies exposed ICS devices and assesses their vulnerability status based on available online data, thus avoiding the direct engagement of the targeted systems [8].

The study meticulously details the methodology of using ICScope for gathering and analyzing data. The researchers collated data from different sources that list exposed ICS devices and matched these findings against known vulnerabilities and patches. The analysis covered a substantial dataset of over 466,000 IP addresses spanning from December 2019 to December 2020. Their findings were alarming; nearly half of the analyzed devices (49.58%) were vulnerable to various exploits, highlighting significant security gaps in current ICS setups. The research also noted a slight but noticeable decline in the percentage of vulnerable devices, suggesting some improvements in security practices over the observed period [8].

Significantly, the study discusses the implications of these vulnerabilities in the context of real-world risks to critical infrastructure. The persistence of such vulnerabilities poses a severe threat, given the slow pace of updates and patches within industrial environments. These settings often prioritize operational continuity over system updates, which can close security gaps but require system downtime. This operational characteristic makes ICS particularly susceptible to long-standing vulnerabilities [8].

Furthermore, Wu et al. offer a critique of current security measures and propose more robust security frameworks that could better protect ICS from emerging threats. They suggest that improving the frequency and methods of updating and patching systems, coupled with a more proactive security stance, could mitigate many of the risks highlighted in their study [8].

Overall, the research by Wu and colleagues provides a crucial insight into the security landscape of Internet-connected ICS devices. By showcasing the effectiveness of passive vulnerability assessments through ICScope, they not only illuminate the extent of exposure of these systems but also advocate for systemic changes in how these critical systems are maintained and protected against cyber threats. This work serves as a significant reference point for further studies and developments in the field of ICS security, emphasizing the need for an evolution in security practices tailored to the unique requirements of industrial control systems [8].

3.0.2 Ontological Approaches

In the paper, Exploring Ontologies for Mitigation Selection of Industrial Control System Vulnerabilities, authors T. Heverin and M. Cordano delve into the intricacies of managing vulnerabilities in industrial control systems (ICS) by utilizing ontologies for mitigation. This innovative approach is centered around enhancing how vulnerabilities are addressed by improving the classification and retrieval of relevant mitigation strategies through the structured use of ontological models. The study emphasizes the significant role of CVE (Common Vulnerabilities and Exposures) entries, which include not just descriptions of vulnerabilities but also metadata about the affected firmware or software versions formatted as CPE (Common Platform Enumeration) entries [9].

The authors argue that current methods for vulnerability management in ICS often lack the structured, semantic relationships that can enable more effective mitigation strategies. By incorporating ontologies, the paper proposes a method to systematically link CVE entries with appropriate mitigation actions based on the specific characteristics of the vulnerability and the context of the ICS environment. This approach not only streamlines the selection of mitigation strategies but also aims to improve the accuracy and relevance of the strategies chosen [9].

Furthermore, the research discusses the development and application of a specific ontology model that categorizes vulnerabilities and their corresponding mitigations. This model serves as a decision-support tool that can aid cybersecurity professionals in quickly identifying and implementing the most effective responses to vulnerabilities that threaten ICS operations. The ontology-based model is designed to be dynamic, allowing for updates and modifications as new vulnerabilities and mitigation techniques are identified [9].

Heverin and Cordano also highlight the challenges of implementing such ontological systems, including the need for comprehensive and up-to-date data on vulnerabilities and mitigations, as well as the integration of this system within existing ICS security frameworks. The paper suggests that overcoming these challenges is essential for the success of ontology-based systems in improving the resilience of ICS against cyber threats [9].

Overall, the study by Heverin and Cordano offers a promising look at how ontological models can be applied to enhance the mitigation selection process in the field of ICS security. By providing a structured and intelligent approach to vulnerability management, their research contributes to the ongoing efforts to safeguard critical infrastructure from evolving cyber threats [9].

3.0.3 System-specific Risk Ratings

In the paper, System-specific risk rating of software vulnerabilities in industrial automation control systems, authors M. Maidl, D. Kröselberg, and T. Zhao explore the critical role of Common Vulnerabilities and Exposures (CVEs) in the vulnerability monitoring processes within industrial systems [10]. The study delves into the necessity of system-specific approaches to effectively rate and manage the risks associated with software vulnerabilities, which are prevalent and potentially devastating in industrial control environments.

The research underscores the inadequacies of traditional, generic risk assessment models that fail to account for the unique characteristics and operational contexts of industrial systems. Maidl, Kröselberg, and Zhao argue for a tailored risk rating approach that not only incorporates the standard metrics of vulnerability severity but also considers the specific operational and environmental factors of industrial automation systems. This method provides a more accurate reflection of the actual risk posed by each vulnerability in its specific industrial context [10].

Further, the paper discusses the development and implementation of a new risk rating model that integrates CVE data with system-specific parameters. This model is designed to enhance the precision of vulnerability assessments in industrial settings, enabling organizations to prioritize remediations based on the actual threat level rather than generic risk scores. The authors present case studies and simulations to demonstrate how their model more effectively identifies critical vulnerabilities that require immediate attention, thereby improving the overall security posture of industrial control systems [10].

Moreover, the paper highlights the importance of continuous monitoring and updating of the risk assessment models to adapt to the evolving landscape of threats and vulnerabilities in industrial systems. The dynamic nature of threat environments necessitates that risk models remain flexible and up-to-date to ensure they are capable of mitigating new and emerging threats [10].

Overall, Maidl, Kröselberg, and Zhao's research provides significant insights into enhancing vulnerability management in industrial automation and control systems through a system-specific risk rating approach. Their work contributes to a more strategic and effective framework for managing cybersecurity risks in industrial environments, emphasizing the critical role of CVEs in this process [10].

3.0.4 Risk Candidate Methods Analysis

In their paper, Industrial Control Systems Cyber Security Risk Candidate Methods Analysis, authors L.A. Dawson, C. Lamb, and A.J. Carbajal undertake a comprehensive examination of cybersecurity risk assessment methodologies applied specifically to industrial control systems (ICS) [11]. Their research focuses on evaluating the effectiveness of various risk assessment techniques that leverage Common Vulnerabilities and Exposures (CVE) scores alongside the Common Vulnerability Scoring System (CVSS) and their correlation with Common Platform Enumeration (CPE) names for precise vulnerability identification.

The authors analyze multiple methodologies to identify how well they accommodate the unique requirements of ICS security, emphasizing the need for tailored approaches that reflect the operational and security specifics of industrial environments. The study critiques existing methods for their often piecemeal approach to risk analysis, which typically addresses only single facets of risk without integrating a holistic view of the cybersecurity landscape within which ICS operate [11].

A significant contribution of Dawson, Lamb, and Carbajal's work is the critical evaluation of how CVE and CVSS are utilized in the context of ICS. They argue that while these tools are fundamental in identifying and scoring vulnerabilities, their application must be nuanced to accurately reflect the potential impacts on industrial systems. The research suggests that mismatches between CVE entries and their corresponding CPE names can lead to gaps in risk assessments, thereby complicating the mitigation process [11].

Furthermore, the paper explores the adaptation of these cybersecurity measures in the face of evolving threats, proposing a dynamic adjustment of risk assessment models to better suit the changing nature of threats and the technological advances in industrial controls. This involves a more integrated approach that not only assesses risks but also actively involves updating and refining risk models to maintain relevance over time [11].

Overall, Industrial Control Systems Cyber Security Risk Candidate Methods Analysis by Dawson, Lamb, and Carbajal provides a thorough exploration of risk analysis methods in the cybersecurity domain of ICS. The study underscores the necessity for methodological improvements that synchronize risk assessment with the operational realities of industrial environments, aiming to enhance the robustness of cybersecurity measures in this critical sector [11].

4 Design

This chapter elaborates on the research design and methodologies applied in systematically assessing and ranking vulnerabilities in Industrial Control Systems (ICS). Recognizing the integral role of ICS in Operations Technology (OT) networks and their susceptibility to cyber threats, the study introduces a novel, structured database approach to evaluate and rank the security posture of ICS vendors and their products comprehensively.

4.1 Problem Statement

Industrial Control Systems (ICS) are pivotal to the functionality and safety of modern industrial operations, yet they remain susceptible to sophisticated cyber threats. The Cybersecurity and Infrastructure Security Agency (CISA) issues ICS advisories that are critical in informing stakeholders about vulnerabilities by detailing associated Common Vulnerabilities and Exposures (CVEs) and Common Platform Enumerations (CPEs). However, these advisories are presented as discrete instances without relational links between them. This isolation poses a significant challenge: it is difficult to trace and correlate CVEs across different vendors and devices. As a result, identifying relationships and patterns among advisories—which is essential for comprehensive vulnerability assessment—is severely hampered. The absence of interconnected data makes it particularly challenging to discover which devices, often hidden within the swathes of data, are most at risk. These devices, known as the "crown jewels," represent the most critical vulnerabilities within an organization's ICS infrastructure due to their centrality or criticality to operations. Without the ability to cross-reference and analyze vulnerabilities across advisories, cybersecurity professionals are often unable to effectively pinpoint these vital assets, complicating efforts to prioritize and address the most significant threats. This research aims to bridge this gap by developing an advanced structured database system that consolidates and correlates data from various ICS advisories. This system is designed to enable dynamic linking and referencing of CVEs, CPEs, vendors, and devices, thus providing a holistic framework for the systematic tracking and analysis of vulnerabilities. By establishing a method to identify and connect the "crown jewels" of ICS vulnerabilities across different advisories, the study seeks to significantly enhance the capacity of security teams to predict, identify, and mitigate critical threats more effectively and efficiently.

4.2 Research Objectives

Addressing the challenges identified in the problem statement, this study sets forth specific objectives to enhance the security assessment capabilities within Industrial Control Systems (ICS). The revised research objectives include:

- **Develop a Comprehensive Database System** - Construct a relational database that systematically integrates and correlates data extracted from ICS advisories issued by CISA. This database will centralize various data elements including CVEs, CPEs, vendors, and devices, addressing the current limitations posed by isolated data sets.

- **Enable Cross-Advisory Analysis** - Create methodologies within the database system to allow dynamic linking of information across different advisories. This will facilitate a comprehensive analysis of vulnerability relationships and trends over time, enabling more effective and strategic cybersecurity interventions.

- **Identify and Prioritize 'Crown Jewels'** - Utilize the structured database to identify and rank the most critically vulnerable devices within ICS networks—those considered 'crown jewels' due to their high value and susceptibility. This objective is designed to help cybersecurity professionals quickly and accurately focus their defensive efforts on the most significant threats.

- **Ensure Continuous Data Updates** - Implement a system within the database framework that automatically incorporates new data from ongoing and future advisories. This system will ensure that the database remains current and relevant, reflecting the latest vulnerabilities and updates, thus maintaining its utility for continuous security assessment and proactive threat management.

These objectives are crafted to directly tackle the primary issues identified, ensuring that the research provides a robust tool for enhancing the detection, analysis, and management of vulnerabilities in ICS environments. Each objective is integral to building a comprehensive solution that addresses both the immediate and long-term security needs of industrial systems.

4.3 Research Design

This research adopts a mixed-methods approach to address the complexities involved in assessing and mitigating vulnerabilities within Industrial Control Systems (ICS). The design is exploratory and focuses on constructing a comprehensive relational database to facilitate the analysis of data extracted from multiple sources.

4.3.1 Data Collection Methods

Data collection for this study involves two distinct phases: initial historical data retrieval and ongoing updates. Initially, data is primarily collected from ICS advisories issued by the Cybersecurity and Infrastructure Security Agency (CISA) and vulnerability details from the National Institute of Standards and Technology (NIST) National Vulnerability Database (NVD). Automated scripts, predominantly written in Python, are developed to interact with public APIs to retrieve a comprehensive backlog of advisories and associated vulnerability data. This structured and consistent data collection method ensures that the initial dataset is robust and suitable for integration into the relational database.

Following the initial data setup, the research employs a dedicated worker program designed to monitor an RSS feed continuously for new advisories. This worker program, which operates autonomously, automatically updates the database with fresh information as soon as it becomes available. This ensures that the database remains current, capturing the latest cybersecurity threats and vulnerabilities as they are reported, thus maintaining the relevance and utility of the research over time.

4.3.2 Database Design and Development

At the heart of the research design is the relational database, constructed to store and efficiently manage data related to vendors, products, CVEs, and CPEs. The database schema is carefully designed to support complex queries, facilitating the identification of relationships and trends across different data points.

4.3.3 Analysis Methodologies

The core of the data analysis in this research is focused on identifying and ranking the most vulnerable vendors and devices within the ICS landscape. To achieve this, the study employs basic statistical analysis techniques that are tailored to assess the frequency and severity of vulnerabilities reported in the collected advisories. The analysis begins with the aggregation of vulnerability counts associated with each vendor and device, considering both the number of vulnerabilities and their impact scores as reported by the National Vulnerability Database (NVD). This approach allows for a comprehensive evaluation of which vendors and devices are most frequently targeted or contain the most severe security flaws. Subsequent ranking involves applying statistical measures to these aggregated data points to determine the relative vulnerability of each entity. The results of this analysis provide a quantifiable ranking system, highlighting those vendors and devices that represent the highest risk and therefore require the most immediate attention from cybersecurity professionals.

4.3.4 Tools and Software Utilization

The research incorporates a range of specialized tools and software, each selected for its ability to facilitate different aspects of the data collection, database management, and analytical processes.

Python: Central to this project, Python is utilized for both the backend RSS parser and the analytics component. For the RSS parsing that feeds data into the system, Python scripts utilize:

- **feedparser:** This library is used to parse the downloaded RSS feeds, allowing the extraction of necessary data from various advisory updates.
- **psycopg2:** This PostgreSQL adapter for Python enables the scripts to interact directly with the PostgreSQL database, facilitating the insertion and management of parsed data.
- **beautifulsoup:** Employed for scraping web pages that are linked within RSS feeds, this library helps in extracting detailed advisory information that is not directly available through the feed content.

For the analytical part of the project, Python provides a plethora of libraries for data analysis. The key libraries chosen for this project are:

- **pandas:** Used for its powerful data structures that simplify data manipulation and analysis, making it easier to clean, transform, and analyze large datasets.
- **scipy:** This library provides a broad range of tools for scientific computing, including functions for statistical testing and data exploration, which are essential for analyzing the patterns and trends in vulnerability data.
- **sklearn (scikit-learn):** Critical for implementing machine learning algorithms, scikit-learn is used for predictive analytics, such as predicting the likelihood of future vulnerabilities based on historical data.

PostgreSQL: The relational database system, PostgreSQL, is chosen for its robustness and capability to handle large volumes of data efficiently. It supports complex queries and extensive data handling, essential for managing the relational database aspects of this study.

Docker: Docker containers are used to encapsulate the Python environment along with PostgreSQL, ensuring that the application runs consistently across different systems. This approach simplifies deployment and scaling, making the research environment more robust and portable.

The integration of these tools and libraries forms a comprehensive technological framework that supports all phases of the research, from data ingestion and storage to complex data analysis and reporting. This setup ensures efficiency, scalability, and reproducibility throughout the research process.

4.3.5 Comparative Analysis of Contributions

Existing Research: Prior studies in Industrial Control Systems (ICS) cybersecurity have often focused on isolated incident reporting and reactive security measures. These methodologies generally lack a systematic framework for the aggregation and analysis of data, relying more on manual processes and less on comprehensive, automated systems. Although these efforts are valuable, they do not fully leverage the capabilities of integrated data analysis for identifying and mitigating vulnerabilities across ICS environments.

Contributions of This Study: As suggested in the literature review chapter of this thesis, a structured analysis approach of CVEs and CPEs applied to ICS is an important next step in ICS security. This thesis introduces a novel structured database system designed to automate the collection, storage, and correlation of extensive ICS-related data, marking a significant departure from existing practices. The primary contributions of this research, as compared to existing works, include:

1. **Systematic Data Integration:** Unlike previous efforts, this research systematically integrates data from multiple sources such as the Cybersecurity and Infrastructure Security Agency (CISA) advisories and the National Institute of Standards and Technology's National Vulnerability Database (NVD). This integration allows for robust vulnerability analysis by connecting discrete data points across various systems and vendors.

2. **Dynamic Linking and Referencing:** The developed database system employs dynamic linking of Common Vulnerabilities and Exposures (CVEs), Common Platform Enumerations (CPEs), vendors, and devices. This method significantly enhances the ability to trace and correlate vulnerabilities, facilitating more comprehensive assessments than typically possible in existing research.

3. **Automated Data Analysis:** By automating the analysis process, this thesis transcends the manual, often static methods of traditional studies. The use of advanced statistical techniques and machine learning algorithms enables ongoing, real-time analysis of emerging threats, which assists in predicting and mitigating potential vulnerabilities more effectively.

4. **Focus on Crown Jewels:** This research uniquely focuses on identifying and

protecting the 'crown jewels' of ICS networks—those assets most critical to operations and most vulnerable to attacks. The methodology not only identifies these assets but also prioritizes them based on empirical data, a strategy not commonly emphasized in previous research.

Advancements Over Existing Methods: This thesis provides a proactive and predictive approach to ICS cybersecurity, significantly advancing beyond the capabilities of existing research. The structured database and its analytical tools offer substantial improvements in how cybersecurity professionals can detect, analyze, and respond to vulnerabilities. These contributions represent a pivotal advancement in transforming reactive cybersecurity measures into a proactive, data-driven strategy within the ICS landscape.

5 Implementation

This chapter delves into the practical aspects of implementing the structured database system designed to assess and rank vulnerabilities in Industrial Control Systems (ICS). It provides detailed insights into the technical setup, the configuration of tools and software, and the operational procedures followed to ensure efficient data handling and analysis. Visual diagrams and figures are included to illustrate the architecture and workflow of the system, offering a clear view of how the theoretical design principles have been applied in practice.

5.1 System Architecture

The system architecture is designed to efficiently capture, process, and store data related to cybersecurity vulnerabilities in Industrial Control Systems (ICS). This section describes the components of the system and their interactions, facilitating a comprehensive understanding of the workflow and data management strategies employed in the research.

5.1.1 RSS Feed Monitoring

The architecture begins with an RSS feed provided by the Cybersecurity and Infrastructure Security Agency (CISA), which continuously releases ICS advisories. These advisories are crucial as they contain timely information about emerging vulnerabilities, including references to relevant Common Vulnerabilities and Exposures (CVEs).

5.1.2 Python Worker in Docker

A critical component of the architecture is the Python worker, which operates within a Docker container. This worker is programmed to monitor the RSS feed actively. Upon detecting new advisories, it parses the feed to extract key details, with a primary focus on identifying the CVEs mentioned in each advisory.

5.1.3 Data Enrichment via NIST NVD

Following the extraction of CVEs, the Python worker queries the National Institute of Standards and Technology's National Vulnerability Database (NVD) to retrieve comprehensive details about each CVE. This includes information such as the Common Vulnerability Scoring System (CVSS) scores and associated Common Platform Enumerations (CPEs). This step is crucial for enriching the advisory data with detailed vulnerability metrics and classifications, enhancing the dataset's utility for subsequent analysis.

5.1.4 Data Storage in PostgreSQL

Once the CVE details are enriched with the necessary information from the NIST NVD, the data is then stored in a PostgreSQL database, which also runs inside a Docker container. This setup not only ensures the data's integrity and security but also facilitates scalability and easy management. The database is designed to allow complex queries and to support the relational linking of data across different advisories, vendors, and devices.

5.1.5 System Diagram

Figure 5.1 provides a visual representation of the system architecture, illustrating the flow of data from the initial RSS feed through to the final storage in the PostgreSQL database. This diagram helps in visualizing the integration of various components and the data processing workflow employed in the research.

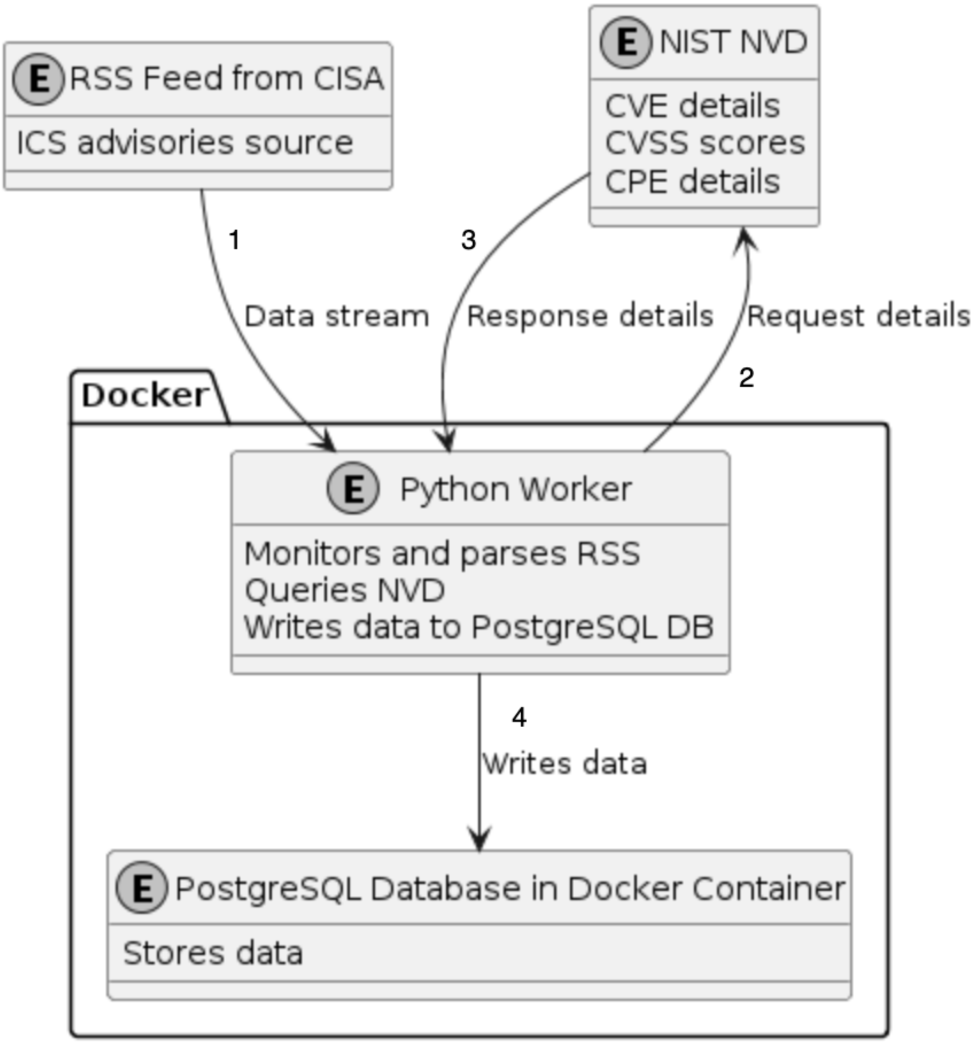


Figure 5.1: System Architecture Diagram

This architecture supports the research's goal to develop a comprehensive system capable of dynamically linking and referencing data across ICS advisories, facilitating a detailed analysis of vulnerabilities and their implications on the security of industrial systems.

5.2 Database Configuration

This section details the database schema configuration used to manage and analyze data collected from CISA advisories and the NIST National Vulnerability Database (NVD). The database is structured into several key tables, each designed to efficiently store specific types of data and facilitate the complex queries necessary for the research.

5.2.1 Advisory List

The *advisory_list* table stores information directly from ICS advisories issued by CISA. Each entry includes details such as the advisory's publication date, title, unique identifier (ICSA), hyperlink to the full text, directory where the advisory's HTML is stored, and the vendor concerned. This table acts as the primary entry point for advisory data, with unique constraints ensuring no duplicate entries and facilitating reliable joins with other tables. Table 5.1 shows the schema of the advisory_list table.

Column	**Type**	**Nullable**	**Constraints**
id	integer	not null	PRIMARY KEY, auto-increment
date	char var(255)		
title	text		
icsa	text		UNIQUE
link	text		UNIQUE
html_dir	text		
vendor	char var(255)		

Table 5.1: Schema of the advisory_list Table

5.2.2 CVE List

The *cve_list* table holds detailed information about CVEs fetched from the NVD. It includes the CVE identifier, publication dates, last modification dates, and various JSONB fields that store structured data such as descriptions, vulnerability metrics (both CVSS v3.1 and v2), and references. This table is central for linking CVE-specific information across the system. Table 5.2 shows the schema of the cve_list table.

Column	**Type**	**Nullable**	**Constraints**
cve_id	char var	not null	PRIMARY KEY
published	timestamp without time zone		
last_modified	timestamp without time zone		
descriptions	jsonb		
metrics_v31	jsonb		
metrics_v2	jsonb		
references	jsonb		
cvss_score	numeric		

Table 5.2: Schema of the cve_list Table

5.2.3 CPE Entries

The *cpe_entries* table catalogs all Common Platform Enumerations associated with each CVE, detailing the affected product, version, and other relevant metadata. This linkage is critical for identifying specific software or hardware that vulnerabilities affect. Table 5.3 shows the schema of the cpe_entries table.

Column	Type	Nullable	Constraints
cpe_id	int	not null	PRIMARY KEY, auto-increment
cve_id	char var(255)		FOREIGN KEY REFERENCES cve_list(cve_id)
part	char (1)		
vendor	char var(255)		
product	char var(255)		
version	char var(255)		
update	char var(255)		
edition	char var(255)		
language	char var(255)		
sw_edition	char var(255)		
target_sw	char var(255)		
target_hw	char var(255)		
other	char var(255)		

Table 5.3: Schema of the cpe_entries Table

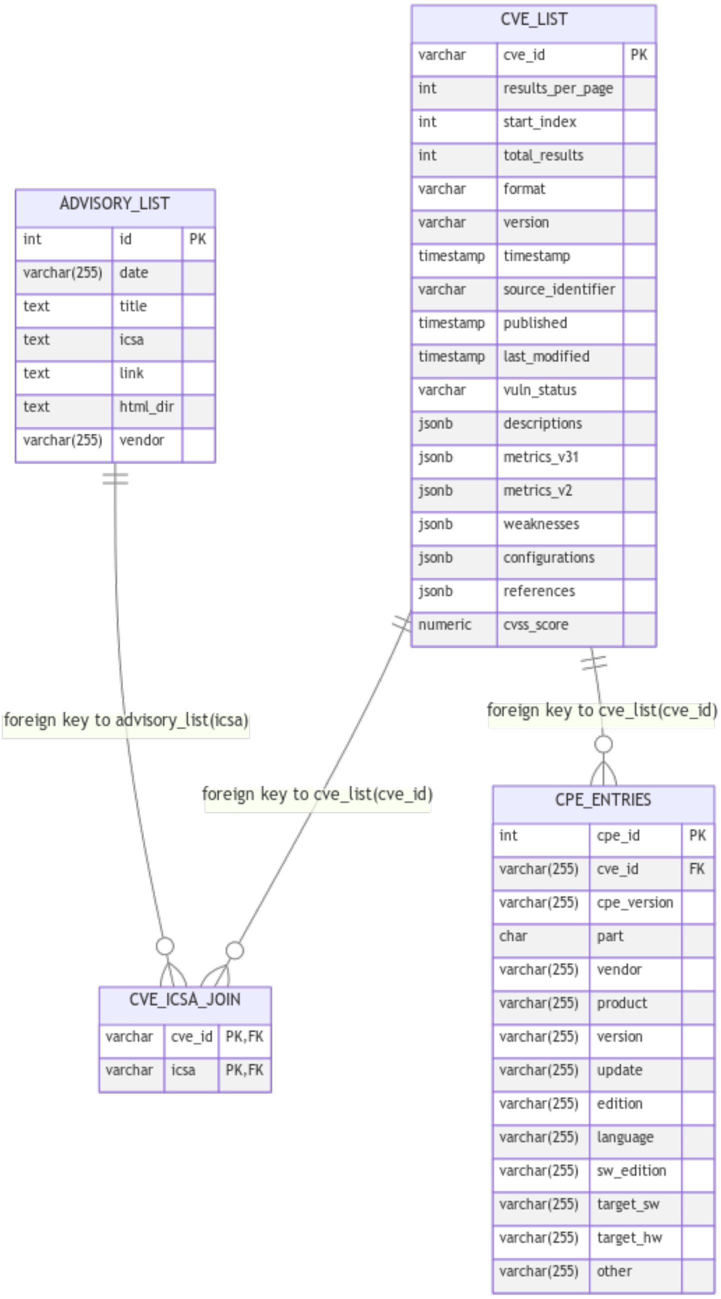
5.2.4 CVE-ICSA Join Table

The *cve_icsa_join* table is a junction table that creates a many-to-many relationship between the *cve_list* and *advisory_list* tables, linking CVEs to their respective advisories. This table is crucial for tracing which CVEs are mentioned in which advisories and supporting the system's ability to analyze the data across advisories dynamically. Table 5.4 shows the schema of the cve_icsa_join table.

Column	Type	Nullable	Constraints
cve_id	char var	not null	PRIMARY KEY, FOREIGN KEY REFERENCES cve_list(cve_id)
icsa	char var	not null	FOREIGN KEY REFERENCES advisory_list(icsa)

Table 5.4: Schema of the cve_icsa_join Table

This configuration supports the system's capability to dynamically link and cross-reference data across different advisories and vulnerabilities, facilitating comprehensive and actionable cybersecurity analyses. The use of foreign keys and indexing strategies enhances query performance and ensures data integrity throughout the system.



CVE_LIST		
varchar	cve_id	PK
int	results_per_page	
int	start_index	
int	total_results	
varchar	format	
varchar	version	
timestamp	timestamp	
varchar	source_identifier	
timestamp	published	
timestamp	last_modified	
varchar	vuln_status	
jsonb	descriptions	
jsonb	metrics_v31	
jsonb	metrics_v2	
jsonb	weaknesses	
jsonb	configurations	
jsonb	references	
numeric	cvss_score	

ADVISORY_LIST		
int	id	PK
varchar(255)	date	
text	title	
text	icsa	
text	link	
text	html_dir	
varchar(255)	vendor	

CVE_ICSA_JOIN		
varchar	cve_id	PK,FK
varchar	icsa	PK,FK

CPE_ENTRIES		
int	cpe_id	PK
varchar(255)	cve_id	FK
varchar(255)	cpe_version	
char	part	
varchar(255)	vendor	
varchar(255)	product	
varchar(255)	version	
varchar(255)	update	
varchar(255)	edition	
varchar(255)	language	
varchar(255)	sw_edition	
varchar(255)	target_sw	
varchar(255)	target_hw	
varchar(255)	other	

foreign key to advisory_list(icsa)
foreign key to cve_list(cve_id)
foreign key to cve_list(cve_id)

Figure 5.2: Database Schema Diagram

5.3 Data Collection Implementation

This section describes the operational steps and scripts implemented to automate the collection, parsing, and storage of data from cybersecurity advisories and vulnerability details.

5.3.1 RSS Feed Monitoring and Parsing

The system continuously monitors an RSS feed provided by the Cybersecurity and Infrastructure Security Agency (CISA), which lists the latest ICS advisories. Each advisory entry is parsed to extract critical details such as the publication date, advisory title, ICSA identifier, and the link to the full advisory text. This information is essential for initial data entry and subsequent cross-referencing with CVE details.

Algorithm 1: Data Collection and Processing Algorithm

```
Initialize database connection;
while true do
    Parse RSS feed from CISA;
    for each entry in RSS feed do
        Extract date, title, link, and ICSA number from entry;
        Attempt to extract CVEs from entry summary and title;
        if vendor information available in entry then
            Extract vendor information;
        end
        Store HTML content in designated directory;
        Prepare data tuple for database insertion;
        if new CVEs found then
            for each CVE found do
                Fetch detailed CVE data from NVD;
                Save CVE details to database;
                for each ICSA associated with CVE do
                    Insert mapping of CVE to ICSA in join table;
                end
            end
        end
        Insert advisory data into database;
    end
    Sleep for a defined interval (e.g., 1 hour);
end
```

5.3.2 CVE Data Retrieval and Processing

When CVE identifiers are mentioned within an advisory, the system fetches detailed information from the National Institute of Standards and Technology's National Vulnerability Database (NVD). This information includes vulnerability descriptions, scoring metrics (CVSS), and other metadata. Each CVE's details are stored in a structured format within the database, linked to the corresponding advisory via a relational join table.

5.3.3 Data Storage

Extracted data is systematically inserted into a PostgreSQL database. The database architecture is designed to facilitate efficient queries and maintain data integrity, allowing for the dynamic linking of advisories to their corresponding CVEs and other relational data. Data consistency and non-duplication are ensured through unique constraints and conflict resolution strategies during the insertion process.

5.3.4 Automation and Scheduling

The entire data collection process is automated using a script that runs in an indefinite loop, periodically waking up to check for new advisories. This setup ensures that the database remains up-to-date with the latest information without manual intervention. Error handling mechanisms are in place to manage potential disruptions in data access or processing.

The algorithm used for this implementation ensures comprehensive data capture and update, supporting the ongoing analysis and research needs related to cybersecurity vulnerabilities in ICS.

5.4 Analytical Tools Setup

5.4.1 Python Libraries

The project utilizes Python for its robust ecosystem of libraries suited for data analysis. Key libraries include:

- **Pandas:** Used for data manipulation and cleaning. Pandas provide powerful data structures to simplify the merging, reshaping, and aggregation of data, which is crucial for preparing datasets for analysis.

- **SciPy:** Utilized for scientific computing, including statistics. SciPy is instrumental in performing more complex calculations and statistical analyses on data.

- **Scikit-Learn:** Applied for machine learning tasks. This library is used to build predictive models based on the historical data collected, enabling the prediction of vulnerability trends and potential future security threats.

5.4.2 Data Analysis Workflow

Data analysis involves several stages, starting with data cleaning and preprocessing using Pandas. This is followed by statistical analysis or model building, where SciPy and Scikit-Learn are used to apply statistical tests and machine learning algorithms, respectively. The results of these analyses help identify patterns and insights into vulnerability exposures and their impacts.

5.4.3 Integration and Automation

The analytical tools are integrated into the data pipeline such that analysis can be run automatically on updated datasets. Scripts are scheduled to execute periodically, ensuring that the latest data is always analyzed and that insights are derived from the most current information available.

5.4.4 Visualizations and Reporting

Additionally, Python's Matplotlib libry is used to generate visualizations. These visualizations help illustrate the findings and are included in reports and dashboards for easy interpretation by cybersecurity professionals and stakeholders.

The setup of these analytical tools enables a comprehensive analysis of the data, ensuring that findings are not only accurate but also actionable for improving the security posture of ICS environments.

5.5 Containerization with Docker

The implementation of Docker containers to ensure a consistent and portable environment across different deployment scenarios is detailed. Configuration files and Docker-compose setups are discussed to illustrate how the application environment is defined and managed.

6 Evaluation and Testing

This chapter discusses the evaluation methods and testing procedures used to assess the effectiveness and accuracy of the data collection, storage, and analysis systems implemented in the previous chapters. The evaluation focuses on two main aspects: frequency analysis of Common Platform Enumerations (CPE) and weighted analysis based on Common Vulnerability Scoring System (CVSS) scores.

6.1 CPE Frequency Analysis

The frequency analysis of CPE entries helps in identifying the most frequently cited vendors and products in the vulnerability data. This analysis is crucial for determining which vendors and products are potentially more vulnerable or targeted more frequently, thus providing insights into possible security focus areas.

6.1.1 Methodology

The methodology involves extracting the frequency of occurrence for each vendor and product from the CPE entries stored in the database. The analysis is performed using statistical techniques such as Z-scores and Min-Max scaling to standardize the frequency counts and provide a clear comparison across different vendors and products.

6.1.2 Implementation

The analysis is implemented using a Python script that connects to the PostgreSQL database to fetch relevant data. The script calculates the frequency of each vendor and product, applies Z-score normalization to measure the relative significance of each frequency, and then scales the results using Min-Max scaling for better interpretability.

Algorithm 2: CPE Frequency Analysis Algorithm

```
Connect to PostgreSQL database;
Fetch vendor and product data from cpe_entries table;
for vendor and product data in cpe_entries table do
    Calculate frequency for each vendor and product;
    Compute Z-scores for frequency data;
    Apply Min-Max scaling to the Z-scores;
    Store and display the results in a structured format;
end
```

6.1.3 Results

The results of the frequency analysis are summarized in two tables, showcasing the top 10 vendors and products based on their frequency of occurrence, Z-scores, and Min-Max scaled values. These metrics provide insight into which vendors and products are most commonly associated with vulnerabilities, offering a window into areas where security efforts may be most needed.

Vendor Analysis

Table 6.1 lists the vendors most frequently mentioned in cybersecurity advisories, with Cisco leading, indicative of its significant presence in advisories. This suggests that the more prominent and widely used a vendor's products are, the more likely they are to be implicated in vulnerabilities.

	Frequency	Z-Scores	Min-Max Scaled
Vendor			
Cisco	7448	20.430537	1.000000
Siemens	5512	15.089267	0.740030
Intel	4238	11.574402	0.568954
Mitsubishi Electric	1742	4.688137	0.233785
Schneider Electric	930	2.447893	0.124748
Oracle	652	1.680913	0.087418
Rockwell Automation	618	1.587110	0.082852
Qualcomm	588	1.504343	0.078824
Moxa	581	1.485030	0.077884
Juniper	559	1.424334	0.074930

Table 6.1: Top 10 vendors based on frequency and their standardized scores in cybersecurity advisories

Similarly, Table 6.2 presents the most frequently mentioned products. The use of Z-scores and Min-Max scaling attempts to normalize the data, which was challenging due to the limited external data available from vendors about total products produced.

Figures 6.1 and 6.2 provide graphical representations of the frequency distributions of vendors and products. These distributions closely resemble a Pareto distribution, where a small number of entities account for a large proportion of occurrences. This pattern is typical in markets where large vendors dominate, leading to more reported vulnerabilities due to the higher usage of their products.

	Frequency	Z-Scores	Min-Max Scaled
Product			
iOS	4812	139.882879	1.000000
iOS_XE	714	20.712552	0.148202
Junos	428	12.395638	0.088755
NX-OS	257	7.422938	0.053211
OpenSSL	235	6.783175	0.048639
NTP	213	6.143413	0.044066
ESXi	193	5.561810	0.039909
Core i5	169	4.863888	0.034920
Core i7	168	4.834807	0.034712
Samba	164	4.718487	0.033881

Table 6.2: Top 10 products based on frequency and their standardized scores

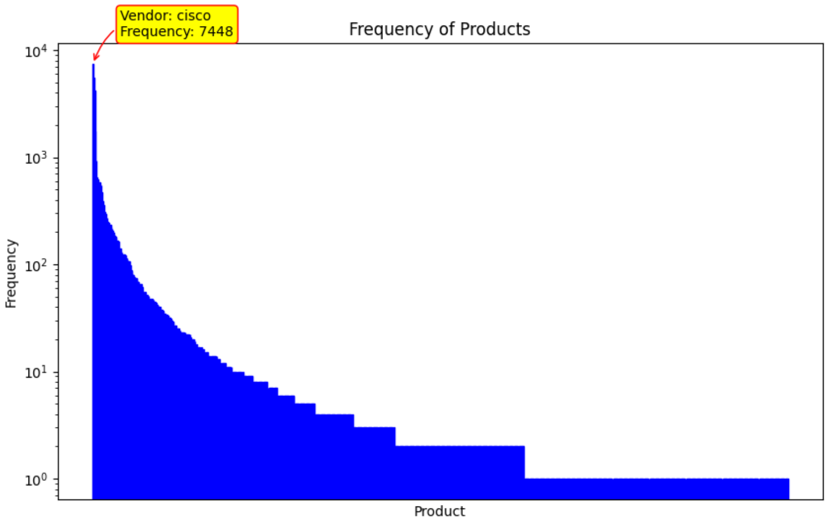


Figure 6.1: Frequency of Vendors

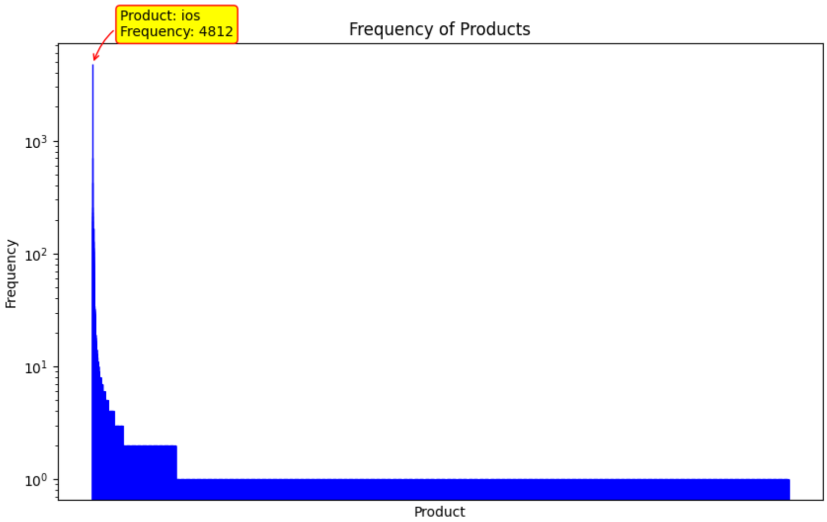


Figure 6.2: Frequency of Products

These findings underscore the challenges in data normalization and highlight the need for additional research into more effective normalization techniques that take into account the scale of production and distribution of products by vendors.

6.2 Weighted Analysis by CVSS Scores

This section delves into a deeper layer of analysis by incorporating the severity of vulnerabilities through CVSS scores, providing a weighted assessment of vendors based on the vulnerabilities associated with their products. This methodology assigns severity-weighted scores to each vendor, illustrating the potential impact and the urgency of addressing the vulnerabilities linked to them.

6.2.1 Methodology

The analysis involves merging CPE entries with CVE data to compile a comprehensive dataset that includes products and their respective CVSS scores. This dataset is then processed to aggregate the total CVSS score for each vendor, serving as a weighted measure of the severity of vulnerabilities associated with their products.

6.2.2 Implementation

Data extraction and processing are performed using a Python script, which efficiently retrieves the necessary data from the PostgreSQL database. The script calculates the total CVSS score for each vendor and ranks them accordingly. This automation ensures accuracy and timeliness in evaluations.

Algorithm 3: Weighted CVSS Analysis Algorithm

```
Connect to PostgreSQL database;
for vendor and product data in cpe_entries table do
    Group data by vendor and sum CVSS scores;
    Sort vendors by aggregated CVSS score in descending order;
    Display the ranked list of vendors based on CVSS scores;
end
```

6.2.3 Results

The results showcase the severity of vulnerabilities associated with different vendors, ranked by their total CVSS scores. This ranking not only highlights which vendors' products are more vulnerable but also reflects the relative severity of these vulnerabilities.

	Vendor	Total CVSS Score
1	Cisco	46696.9
2	Siemens	39513.2
3	Intel	24504.5
4	Mitsubishi Electric	14601.0
5	Schneider Electric	7347.8
6	Qualcomm	5684.2
7	Rockwell Automation	5043.8
8	Omron	4685.7
9	Oracle	4162.5
10	Moxa	4127.2

Table 6.3: Ranking of vendors based on total CVSS scores

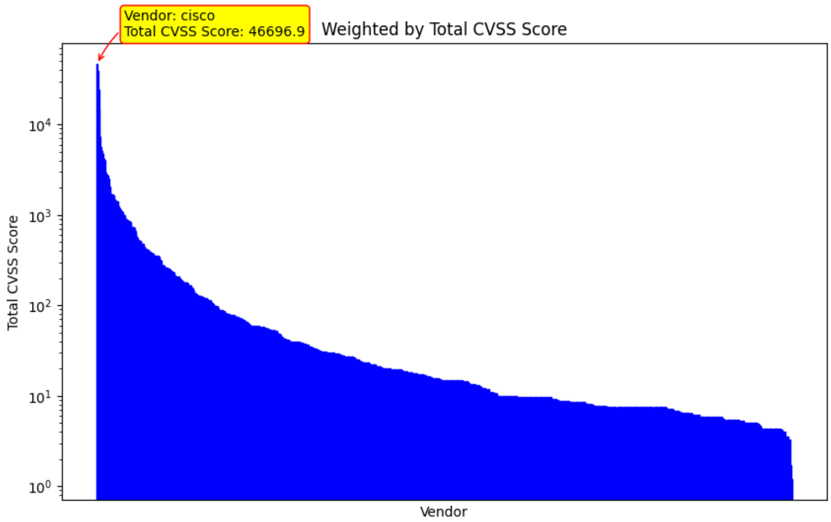


Figure 6.3: Graphical representation of vendors ranked by weighted CVSS scores

The graph in Figure 6.3 illustrates a clear variance when compared to the Vendor Frequency analysis, indicating the criticality of vulnerabilities rather than just their frequency. While the frequency table shows how often vendors appear in advisories, the weighted CVSS scores highlight the severity and potential impact of the vulnerabilities, revealing a more acute risk profile. This method, though basic, provides vital insights, suggesting that while some vendors may appear less frequently in advisories, the severity of the vulnerabilities associated with them can be greater, demanding prioritized attention. The preliminary nature of this weighted CVSS scoring underscores the need for developing more advanced and nuanced severity weighting methodologies to enhance security prioritization accuracy. Although effective in providing a basic insight into the security landscape of Industrial Control Systems, there is a compelling need for further research to refine these models. Such advancements will enable stakeholders to address the most critical threats more comprehensively and precisely, ensuring better-equipped mitigation strategies.

6.3 Challenges with Wildcards in CPE Data

The use of wildcards in Common Platform Enumerations (CPE) data introduces significant challenges in the frequency and severity analysis of vulnerabilities. Wildcards are used in CPE entries to denote unspecified or variable components within product identifiers, such as versions or updates. While this allows for a broader representation of affected products, it also introduces a level of imprecision that can skew the analysis of vulnerabilities.

6.3.1 Impact on Data Accuracy

Wildcards complicate the accuracy of vulnerability assessments because they can inflate the apparent scope of a vulnerability's impact. This is an inherent weakness of using the CPE data in the research of this thesis and all other research done by other researchers. For instance, a wildcard in the version field of a CPE may suggest that all versions of a product are vulnerable, which may not be the case. This can lead to overestimations of vulnerability exposure and severity, affecting prioritization and mitigation efforts.

6.3.2 Methodological Challenges

Analyzing data that contains wildcards requires careful consideration to avoid misleading conclusions. Standard statistical methods may not be sufficient to handle the ambiguities introduced by wildcards. For example, frequency counts derived from wildcard-inclusive CPEs might incorrectly suggest that a product is universally vulnerable, leading to skewed risk assessments.

6.3.3 Addressing Wildcard Issues

To mitigate these issues, more sophisticated data processing techniques are needed. These might include:

- **Estimation Techniques:** Developing statistical models to estimate the likely range of affected product versions based on available vulnerability and patch information.
- **Data Enrichment:** Augmenting CPE data with additional metadata from vendors or other trusted sources to narrow down the affected versions or configurations.
- **Sensitivity Analysis:** Conducting analyses to determine how changes in the interpretation of wildcard CPEs could affect the overall vulnerability assessment.

6.3.4 Linking Findings to the Purdue Model

The evaluation of Common Vulnerability Scoring System (CVSS) scores among vendors, particularly noted in the high vulnerability scores of vendors like Cisco, Intel, and Oracle, underscores their significant roles in the upper layers of the Purdue Model. These companies predominantly operate within Level 2 and above, where enterprise management and control systems occur. The critical systems at these levels often serve as the crown jewels of ICS environments, mirroring traditional IT systems in terms of their infrastructure components such as servers, routers, and switches.

The heightened CVSS scores associated with these vendors highlight the critical need for robust cybersecurity measures at these levels. The Purdue Model categorizes these layers as more susceptible to cyber threats due to their connectivity and the integration of complex IT systems. Many of the most significant vulnerabilities found in these systems are traditional vulnerabilities common in IT infrastructure, reflecting issues in operating systems and network configurations. By focusing cybersecurity efforts on these layers, organizations can better protect the operational technologies that are essential to their day-to-day functions, aligning with the model's recommendations for layered security and addressing vulnerabilities that could severely impact enterprise-wide operations.

7 Conclusion

This thesis presented the development and implementation of a sophisticated database system designed to systematically analyze and rank vulnerabilities in Industrial Control Systems (ICS). Through the innovative integration of data from the Cybersecurity and Infrastructure Security Agency (CISA) and the National Institute of Standards and Technology's National Vulnerability Database (NVD), this study has significantly advanced the understanding of the complex security challenges inherent in ICS.

7.1 Contributions of the Study

The primary contribution of this research lies in the creation of a dynamic system that efficiently processes, stores, and correlates extensive vulnerability data across various ICS components. The architecture of this system, as detailed in the Implementation chapter, enables comprehensive evaluations of vulnerabilities, significantly enhancing the capability to assess and prioritize cybersecurity threats based on empirical data. This is crucial for protecting higher-level systems that encompass critical operations, control, and crucial enterprise management and business planning functions similar to traditional IT systems in terms of their structure and vulnerability profiles.

The convergence of Information Technology (IT) and Operational Technology (OT) at these levels introduces complex cybersecurity challenges. These systems' critical roles and connectivity make them prime targets for cyber threats, as emphasized by the pronounced Common Vulnerability Scoring System (CVSS) scores associated with major vendors such as Cisco, Intel, and Oracle. This thesis has advocated for the development of future security models that address both common IT vulnerabilities and the unique requirements of operational technology, thereby proposing a refined focus on the most vulnerable yet crucial components of modern industrial systems.

7.2 Strategic Implications for Protecting Higher-Level Systems

The findings of this thesis demonstrate that vulnerabilities in higher-level systems, particularly those provided by major vendors, are significantly impactful within the framework of the Purdue Model. The pronounced CVSS scores associated with these vendors underscore the pressing need for robust security measures that not only protect against common IT vulnerabilities but also consider the unique requirements of operational technology. Future research should therefore focus on enhancing security models that integrate the protection of these crown jewels within ICS, developing methods that are tailored to the interconnected nature of these systems. Strengthening defenses at these critical points will provide a more resilient infrastructure, aligning with the Purdue Model's layered security approach and offering a fortified barrier against the evolving landscape of cyber threats.

7.3 Implications for Cybersecurity in ICS

The Evaluation and Testing chapter highlighted critical vulnerabilities and identified the most frequently targeted vendors and products. The analysis demonstrated the necessity for stakeholders in the ICS sectors to adopt mitigation strategies that address both the prevalence and severity of these vulnerabilities, which is essential for strengthening the resilience and security of critical infrastructure.

7.4 Future Research Directions

Despite the successes of the current system, the field of ICS cybersecurity is rapidly evolving, necessitating ongoing research to refine and expand the methodologies employed. Future research should focus on developing advanced severity weighting models that more accurately reflect the potential impacts of vulnerabilities. In this research, a very basic approach of summing the CVSS scores was taken. More advanced statistical or machine learning models should be used to analyze CVSS scores accross vendors and devices.

As discussed in the Evaluation chapter, further research is essential to develop methodologies that can accurately interpret and analyze wildcard-containing CPE data. This research could focus on machine learning models capable of predicting the actual impact of vulnerabilities based on historical data and trends, despite the presence of wildcards. Additionally, collaboration with standard-setting bodies and vendors to reduce the reliance on wildcards in vulnerability reporting could enhance the clarity and utility of CPEs. These enhancements would not only improve the accuracy of vulnerability databases but also support more precise security measures tailored to the actual risk landscape of Industrial Control Systems. Lastly, expanding the range of data sources to include international cybersecurity advisories would offer a broader perspective on global ICS vulnerabilities.

7.5 Concluding Remarks

The groundwork laid by this thesis is vital for the ongoing development of cybersecurity measures in industrial control systems. The system not only facilitates proactive vulnerability management but also serves as an essential tool for researchers, security professionals, and policymakers dedicated to safeguarding critical infrastructure. As ICS increasingly intersect with information technologies, the strategies developed through this research will play a crucial role in protecting the infrastructural backbones of modern society.

Bibliography

[1] T. J. Williams, "The purdue enterprise reference architecture," *Computers in Industry*, vol. 24, no. 2-3, pp. 141–158, 1994.

[2] Zscaler. "What is the purdue model for ics security?" (2024), [Online]. Available: https://www.zscaler.com/resources/security-terms-glossary/what-is-purdue-model-ics-security (visited on 04/13/2024).

[3] E. D. Knapp and J. T. Langill, *Industrial Network Security: Securing Critical Infrastructure Networks for Smart Grid, SCADA, and Other Industrial Control Systems*. Syngress, 2014.

[4] R. Malm, *Securing Operational Technology Networks: Best Practices for IT/OT Security*. Auerbach Publications, 2019.

[5] Cybersecurity and Infrastructure Security Agency (CISA). "Industrial control systems." Accessed on 13 April 2024. (2024), [Online]. Available: https://www.cisa.gov/topics/industrial-control-systems (visited on 04/13/2024).

[6] National Institute of Standards and Technology (NIST), "Ir 7695; polystyrene," U.S. Department of Commerce, Gaithersburg, MD, Technical Report, Jan. 2016.

[7] National Institute of Standards and Technology (NIST), "Ir 8085; polystyrene," U.S. Department of Commerce, Gaithersburg, MD, Technical Report, Jan. 2016.

[8] Y. Wu, J. Zhuge, T. Yin, *et al.*, "From exposed to exploited: Drawing the picture of industrial control systems security status in the internet age," in *Proceedings of the 7th International Conference on Information Systems Security and Privacy (ICISSP)*, SCITEPRESS, 2021, ISBN: 978-989-758-491-6. DOI: 10.5220/0010327902370248.

[9] T. Heverin and M. Cordano, "Exploring ontologies for mitigation selection of industrial control system vulnerabilities," in *Proceedings of the 17th International Conference on Cyber Warfare and Security*, Albany, New York, USA, 2022, pp. 72–80. [Online]. Available: https://papers.academic-conferences.org/index.php/iccws/issue/view/2.

[10] M. Maidl, D. Kröselberg, and T. Zhao, "System-specific risk rating of software vulnerabilities in industrial automation control systems," *IEEE Transactions on Industrial Informatics*, 2021.

[11] L. A. Dawson, C. Lamb, and A. J. Carbajal, "Industrial control systems cyber security risk candidate methods analysis," Department of Energy - Nuclear Energy (DOE-NE), Tech. Rep., 2018. DOI: 10.2172/1463794. [Online]. Available: https://www.osti.gov/biblio/1463794.